



# Focus

Volume 23

Number 1

Winter 2004

Temporary help employment and disadvantaged workers	1	How do welfare sanctions work? New findings from Wisconsin and Illinois	37
Welfare policy choices in the states: Does the hard line follow the color line?	9	Benefit-cost analysis in the evaluation of child welfare programs	44
The skills of American workers in today's labor market	17	A cost-benefit analysis of the Chicago Child-Parent Centers	50
The children of New Hope, five years later	26		
Respect, support, and accountability: Lessons in delivering the New Hope Project offer	33		

ISSN: 0195–5705

## Temporary help employment and disadvantaged workers

Carolyn J. Heinrich, Peter R. Mueser, and Kenneth R. Troske

Carolyn J. Heinrich is Associate Professor of Public Affairs, University of Wisconsin–Madison, and an IRP affiliate. Peter R. Mueser and Kenneth R. Troske are Associate Professors of Economics, University of Missouri–Columbia.

A disproportionate share of temporary help agency employees consists of young, female, predominantly minority workers, many of them recent or current welfare recipients. In 2001, for example, nearly 60 percent of temporary services workers were women, half of them between the ages of 16 and 34. And whereas the proportion of black and Hispanic workers in traditional employment—around 11 percent—generally reflected their share of the population, blacks and Hispanics accounted for 25 and 18 percent, respectively, of temporary agency employees.<sup>1</sup> In 1990, about 14 percent of women in temporary services supplemented their earnings with means-tested welfare benefits, compared to 3 percent of all women with permanent work and 6 percent of permanent part-time workers.<sup>2</sup>

The temporary services industry has become an increasingly important source of low-skilled work, and temporary employment grew five times faster than overall employment between 1972 and 2000.<sup>3</sup> In a number of states, temporary help service firms are registered providers of welfare-to-work services. Directly or indirectly, public policies may be encouraging the growth of temporary help service employment.

Workers hired through temporary help service firms are much less likely to receive fringe benefits than workers hired directly by the firms in which they are working. Temporary workers are more likely than others to work fewer hours, with less predictable schedules, and to be paid less than permanent employees in similar jobs. Some commentators have suggested that they have less control over working conditions and work assignments than others, and are less likely to receive job skills training or useful feedback on performance. Their social interactions in the workplace and sometimes their attachment to the workforce itself can be quite marginal.<sup>4</sup>

There is, however, a more positive view of the temporary services industry: that employment through labor market

**Table 1**  
**Characteristics of Individuals Entering Employment Programs in Missouri, 1997**

	Women			Men	
	TANF	JTPA	Employment Service	JTPA	Employment Service
Age (yrs)	28.1	37.2	34.5	38.9	34.1
Education (yrs)	11.3	12.4	12.3	13.0	12.3
High School Diploma (%)	57.8	86.9	87.2	86.6	87.2
College Degree (%)	1.1	5.5	7.8	16.9	7.8
Nonwhite (%)	38.1	32.9	26.6	28.0	23.2
Area of Residence (%)					
St. Louis Co. and St. Louis City	24.8	27.0	21.4	29.2	20.4
Kansas City central area (Jackson Co.)	16.1	13.5	10.9	14.8	11.0
Suburban areas	10.6	15.1	12.5	20.0	14.6
Small metro	12.1	9.9	12.5	8.1	13.1
Outside metro	36.5	34.3	42.0	27.7	39.9
Employment and Earnings History					
Working in previous 8 qtrs (%)	51.1	62.7	62.8	65.8	64.0
Working all of previous 8 qtrs (%)	17.4	35.9	39.0	37.8	40.7
No work in any of previous 8 qtrs (%)	19.3	15.9	17.9	14.3	17.6
Total annual earnings in prior yr (\$)	3,904	8,965	8,946	13,842	13,565
Total annual earnings 2 yrs prior (\$)	3,564	8,929	7,810	14,162	12,033
<i>N</i>	26,172	5,391	133,766	3,028	163,080

**Source:** Data from Missouri TANF files and Unemployment Insurance files.

intermediaries may provide many low-skilled workers with their best employment opportunity, providing some with their only feasible path to permanent and stable employment. By limiting the employer's commitment and hence the risk, such jobs may enable low-skilled workers to gain access to informal training and screening. And some proportion of temporary employees may prefer the flexibility and shorter hours that such jobs offer—those who value nonmarket time highly, those with young children or other family responsibilities.<sup>5</sup>

Weighing the costs and benefits of temporary employment for particular groups is not easy. In this article, we summarize our recent research that examines, in particular, the consequences of temporary employment for women who are or have recently been on public assistance (see box). Our ultimate aim has been to determine whether, in the long run, temporary employment helps or

The research summarized here is discussed in three papers: Carolyn J. Heinrich, "Temporary Employment Experiences of Women on Welfare," forthcoming in the *Journal of Labor Research*, and two papers by Carolyn J. Heinrich, Peter R. Mueser, and Kenneth R. Troske, "Welfare to Temporary Work: Implications for Labor Market Outcomes," forthcoming in the *Review of Economics and Statistics*, and "The Impact of a Temporary Help Job: An Analysis of Outcomes for Participants in Three Missouri Programs," working paper, Department of Economics, University of Missouri–Columbia.

hurts these low-skilled and otherwise disadvantaged workers. Earnings, although important, are only one item in the balance. Has entry in the labor market through a temporary job improved women's circumstances, as measured by wage growth, stability of employment, and welfare receipt, or do they remain mired at the lowest and most unstable employment levels? Who among welfare recipients goes to work for temporary services firms? Is there evidence that they would have preferred more permanent jobs, or does temporary employment fit well with their expectations and circumstances? How do they view their jobs and working conditions?

To explore questions like these, we followed several strategies. We examined the employment dynamics of welfare mothers who took temporary services jobs in two states, North Carolina and Missouri, where over 17 and 13 percent, respectively, of employed welfare recipients were working for temporary services firms in 1999, and where the proportion of recipients with such jobs more than doubled from 1993 to 1997. One study compared those working in temporary services jobs with similarly disadvantaged women who entered jobs in other employment sectors. A second study compared the experiences of Missouri welfare participants in temporary services employment with somewhat more advantaged workers who had sought employment assistance from the state Division of Employment Services and from programs under the Job Training Partnership Act (JTPA).<sup>6</sup> A third strategy explored women's own expectations and attitudes regarding temporary employment through a survey

In this issue of *Focus*, we include articles describing research by scholars who joined the core IRP faculty through awards made to departments under University-wide “cluster hiring” competitions. “New areas of knowledge and complex societal issues,” the University administration noted, “do not always fall neatly into departmental disciplines and structures,” and the cluster hiring initiative was designed to advance knowledge at interdisciplinary crossroads. Michael Handel, Assistant Professor of Sociology, came to UW in 2000 through the Economic Sociology initiative. Carolyn Heinrich, Associate Professor of Public Affairs, and Joe Soss, Associate Professor of Political Science, became faculty members and IRP affiliates in Fall 2003 through the Poverty Studies Cluster initiative.

of recipients of Temporary Assistance for Needy Families (TANF) in North Carolina. This information informs and strengthens our quantitative analysis.

In studying employment outcomes we made use of administrative data from state assistance programs (after 1996, TANF). In Missouri, data were available for all welfare recipients beginning in 1993 and through 2002; in North Carolina, data first became available for the population of welfare recipients in 1995.<sup>7</sup> We matched income assistance data with earnings and employment data from state unemployment insurance programs in North Carolina and (for Missouri) in Missouri and Kansas, thus ensuring coverage for TANF recipients in Kansas City, MO, who often work across state lines. In 2001, we undertook a telephone survey of a representative sample of North Carolina welfare recipients ( $n=74$ ) who had recently worked in temporary services (although they were not necessarily in a temporary services job at the time of the survey). We asked about their job search experiences and their expectations, the kinds of jobs they held and their satisfaction with these jobs, and their levels of well-being, and matched the interviews to individual data from administrative files.

### **The characteristics of the participants**

Welfare recipients in our two states differed substantially in race and residence. In 1997, the proportion of non-white welfare recipients was nearly 20 percentage points higher in North Carolina than in Missouri. Over 50 percent of Missouri welfare recipients lived in large metropolitan areas (St. Louis and Kansas City). In North Carolina, less than 15 percent lived in Charlotte, the state’s largest metropolitan area, and nearly 40 percent lived outside any metropolitan area at all. A much higher percentage of the women in Missouri were long-time welfare recipients, and their work experience and earnings were lower than those of the women in North Carolina.

Table 1 shows how these welfare recipients compare to a more general pool of low-wage workers. TANF entrants in Missouri were substantially more disadvantaged than applicants to the other two employment programs (JTPA and Employment Services). For example, only 58 percent had a high school diploma, compared to nearly 90 percent of both men and women in the other programs. TANF recipients were younger and more likely to be nonwhite. In all programs, women had lower average employment and earnings

than men, but the average prior earnings of TANF recipients were not even half those of women in the other programs. In large part, these differences reflect TANF program restrictions, which limit participation to those with very low incomes.

### **Who worked for temporary help agencies?**

Who, among these disadvantaged Missouri women, were most likely to end up in temporary help services jobs? Neither age nor education showed any particular relationship with a temporary job. Nor did previous earnings; indeed, some women with a history of higher earnings were less likely to be employed in temporary services than they were to be unemployed. These characteristics, combined with the state of the local economy, were important predictors of the likelihood of *any* employment, but none was specifically linked to employment in the temporary help sector.

We did not find that women with more or with younger children—and hence presumably greater responsibilities—were more likely to work for temporary service agencies. By far the most powerful predictor of temporary sector employment was race. In all groups, nonwhites were much more likely to be in temporary jobs. Another important predictor was region within the state: those in metropolitan counties, which offer a larger marketplace for temporary services firms, were much more likely to be in temporary jobs than those in nonmetropolitan counties.

These findings suggest that selection into temporary help jobs is not effectively explained by observed differences in human capital—what matters most is “race and place.” The importance of race is perhaps explained by employers’ difficulty in judging the ability of nonwhite workers, or their belief that nonwhite workers are less productive; they are thus less willing to hire them into regular jobs that imply long-term commitments. If this is the case, temporary help jobs may provide nonwhites with opportunities not otherwise available.

### **Earnings and welfare receipt of recipients working in temporary help services**

In both Missouri and North Carolina, those employed in temporary services resembled other employed welfare

**Table 2**  
**The Characteristics of TANF Recipients in Different Industries, 1997**

	No Job	Temporary Help	Manufacturing	Retail Trade	Service <sup>a</sup>	Other
<b>Missouri</b>						
Education < 12 yrs (%)	48.0	41.9	45.9	47.0	41.3	34.3
Nonwhite (%)	45.0	73.2	43.9	50.8	63.0	63.3
No. months on welfare in previous two yrs	16.6	16.4	13.2	14.9	16.1	15.8
% of previous 8 qtrs employed	25.0	54.5	49.4	53.6	54.5	55.4
<b>North Carolina</b>						
Education < 12 yrs (%)	41.6	33.8	41.2	36.2	28.8	28.9
Nonwhite (%)	66.9	81.1	68.7	64.3	77.1	66.6
No. months on welfare in previous two yrs	15.5	14.3	12.3	13.9	14.7	13.6
% of previous 8 qtrs employed	28.0	57.1	60.1	58.2	59.5	57.3

**Source:** Data from Missouri and North Carolina TANF files.

**Note:** Sample includes women aged at least 18 and younger than 65 in single-parent families, not in child-only cases. Sampling frame is quarter by welfare recipient.

<sup>a</sup>“Service” excludes temporary help employment.

recipients much more than they resembled those without jobs (Table 2). The only significant demographic difference between temporary services employees and other employed recipients, as already suggested, was race. The jobless, in contrast, were less educated, more likely to be white, and had spent fractionally more time on welfare but worked much less than those in temporary services. Nor does it appear, from the Missouri data, that TANF recipients were taking temporary services positions in much greater proportions than job-seekers through JTPA or the Employment Service; temporary employees in the TANF program were, however, much less likely to be

combining temporary help with other kinds of jobs (Table 3).

Patterns of welfare receipt and earnings in both Missouri and North Carolina confirm that TANF recipients in the temporary work sector resembled jobholders in other industries more than they resembled the jobless. Their current earnings are generally lower than workers in other sectors, always by at least 10 percent (Table 4). For those working in multiple sectors, the earnings disadvantage attached to temporary services diminished, or, in some cases, reversed. In both states over the next two years,

**Table 3**  
**Jobs of Low-Wage Workers in Missouri, at Entry into the Employment Program, 1997**

	Women			Men	
	TANF	JTPA	Employment Service	JTPA	Employment Service
No Job (%)	52.7	40.8	33.4	34.7	32.1
Job in Only One Sector (%)					
Temp help	4.6	4.8	4.1	5.3	4.3
Manufacturing	3.9	4.3	4.6	5.8	4.5
Retail trade	3.1	6.9	9.7	11.5	14.2
Service (excluding temp help)	14.7	19.9	19.0	12.9	8.6
Other	12.6	8.2	12.6	6.7	9.3
Jobs in Multiple Sectors (%)					
Temp help and any other industry	3.6	9.2	8.5	15.4	19.5
Any industry, not temp help	4.8	6.0	8.1	7.6	7.7
Employed 8 qtrs after qtr of entry (%)	57.0	68.7	63.9	65.1	62.9
N	26,172	5,391	133,766	3,028	163,080

**Source:** Data from Missouri TANF files and Missouri and Kansas Unemployment Insurance files.

**Table 4**  
**Distribution of Jobs among Industries, and Earnings of Welfare Recipients**

Industry Combinations	Missouri			North Carolina		
	Employed in Sector 1997 (%)	Earnings (\$)		Employed in Sector 1997 (%)	Earnings (\$)	
		Quarter of Entry <sup>a</sup>	Summed over Next 2 Yrs		Quarter of Entry <sup>a</sup>	Summed over Next 2 Yrs
<b>One Sector</b>						
Temp help	11.0	940	11,600	10.2	1,035	12,549
Manufacturing	4.9	1,565	13,391	9.5	1,604	14,444
Retail	25.3	1,090	10,705	30.1	1,128	11,329
Service <sup>b</sup>	34.1	1,461	13,798	26.6	1,413	14,218
Other	7.8	1,973	16,810	5.1	1,682	15,542
<b>Multiple Sectors</b>						
Temp help and any other industry	8.3	1,535	14,779	8.6	1,528	15,085
No jobs in temp help industry	8.6	1,615	13,981	9.9	1,652	14,569

**Source:** State Unemployment Insurance data from Missouri, Kansas, and North Carolina.

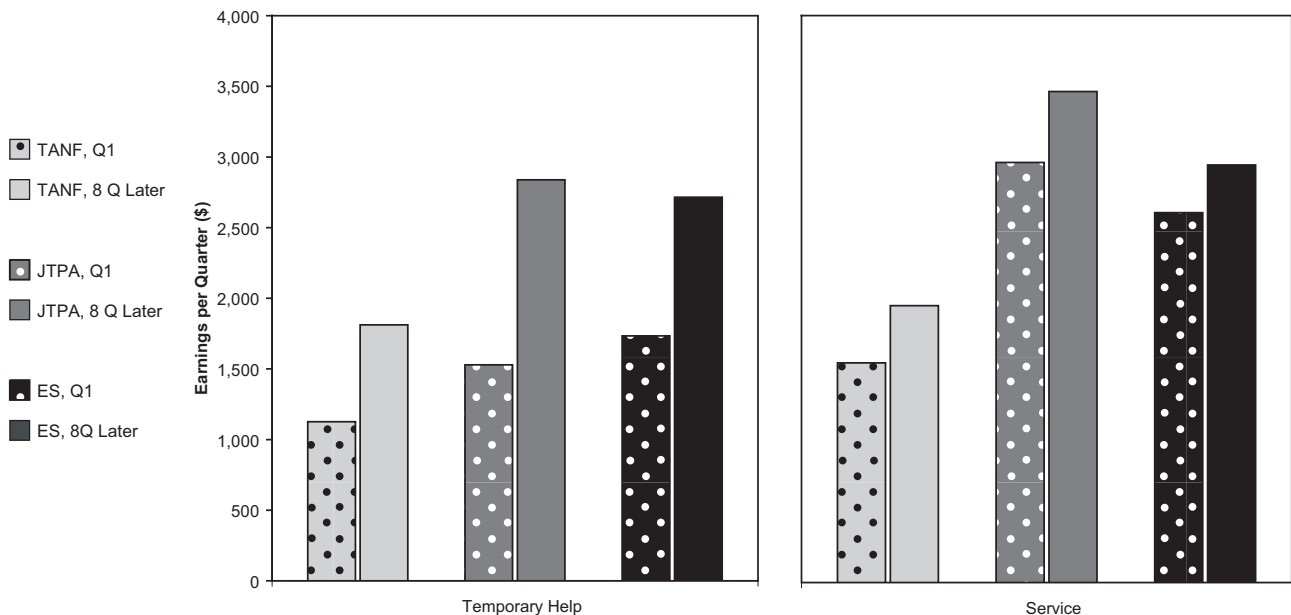
**Note:** Sample includes females aged at least 18 and less than 65 in single-parent families, not in child-only cases. Sampling frame is quarter by welfare recipient.

<sup>a</sup>Earnings in first job following observed quarter on welfare.

<sup>b</sup>Service excludes temporary help.

earnings for those originally in the temporary sector closed much of the disparity observed in current earnings (Table 4). Temporary sector earnings were, for example, only 14 percent less than manufacturing sector earnings and were higher than earnings in the retail sector. And among those with multiple jobs, women with both a temporary sector job and another job had higher earnings than those working more than one job in other sectors. Recipients working in the temporary help sector thus show higher rates of earnings growth than recipients employed in other industries, even after we have taken into account differences in workers' personal characteristics.

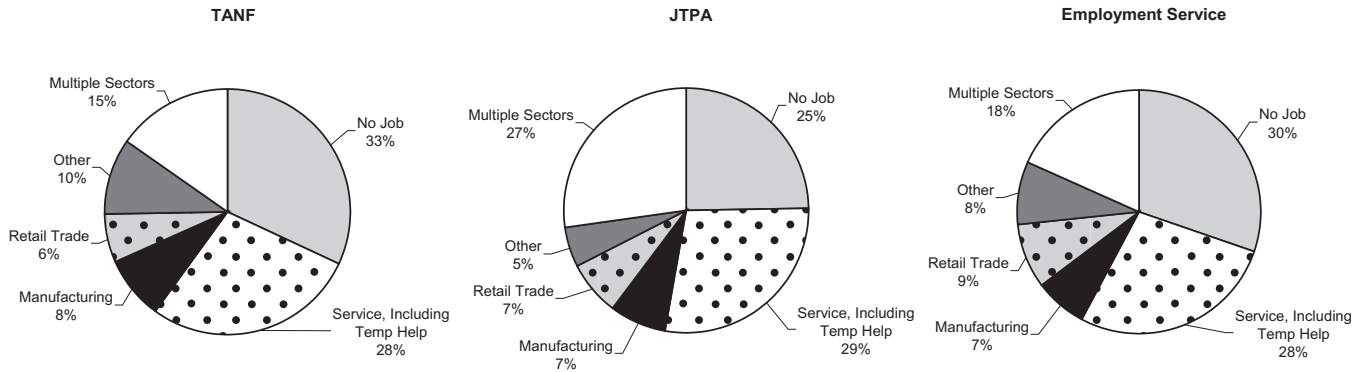
The evidence regarding comparative earnings from our Missouri study also suggests that the earnings disadvantage of other workers entering the temporary help sector declined over time.<sup>8</sup> The samples of women in the JTPA and Employment Service programs were quite heterogeneous, and their characteristics and employment histories substantially different from those of the TANF recipients, yet the role of temporary help employment appears remarkably similar. Figure 1 presents earnings in the quarter after program participation and earnings eight quarters later for women in temporary help industry jobs, comparing these earnings with earnings for women in



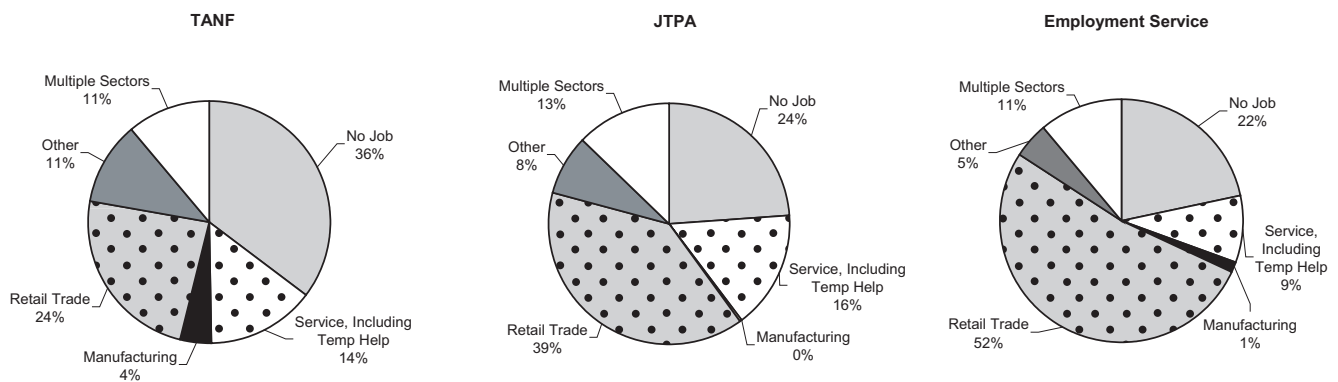
**Figure 1. Earnings Growth among Low-Wage Workers in Employment Programs in Missouri.**

**Source:** State Unemployment Insurance data from Missouri and Kansas.

Employment Sector after Two Years, for Those Who Entered in a Temporary Services Job



Employment Sector after Two Years, for Those Who Entered in Retail Trade



**Figure 2. Employment Two Years Later for Those in the Temporary Help and Retail Trade Sectors When They Entered the Employment Program.**

Source: State Unemployment Insurance data from Missouri and Kansas.

other service jobs (this is the sector in which the highest proportion of all women were employed at entry into the program). Once again, TANF participants earned least of all, and those differences hardly diminished over the next two years. Yet TANF recipients working in the temporary help sector also had the highest earnings growth, over 60 percent; the earnings of TANF workers in service jobs, by comparison, grew by just 25 percent. Among participants in both other programs, earnings growth in the temporary sector also outpaced growth in other sectors, and by large magnitudes.

**Job transitions**

What is the key to labor market advancement for disadvantaged workers who begin with a temporary job? Our evidence suggests that these workers improve their employment and earnings at least in part by moving to another job sector. Figure 2 offers a snapshot of the three groups of Missouri women in employment programs, comparing the employment experiences of those who

were in temporary jobs and in retail employment when they entered the program. There is little evidence here that those in temporary jobs were “stuck.” Over the next two years, women in temporary help positions were more likely to move to some other major sector than were individuals in the other sectors. For example, only 28 percent of TANF recipients were in temporary help services two years later, whereas 42 percent of recipients working in other service jobs were still working in some kind of service position (not in the figure).

**Women’s expectations and employment outcomes**

Evidence from our telephone survey of North Carolina welfare participants tends to confirm the positive effects of temporary services jobs, but also makes clear the difficulties faced by these women in obtaining satisfactory employment. Among the welfare participants, 30 reported that their current or most recent job was with a temporary help services firm. On average, they had worked 16.5 of the past 52 weeks



in such jobs, the vast majority for one or two firms. Almost all had worked on more than one assignment—most commonly clerical (47 percent) and industrial blue-collar (46 percent); 21 percent had worked in the services industry, and less than 10 percent in professional/managerial, sales, or technical services.

### **How did women find temporary services jobs?**

Although nearly three-quarters had participated in welfare-to-work program activities and another 18 percent in a workforce development program, 77 percent of the women had learned about temporary help services from other sources—by contacting the firm directly, from friends or word of mouth, from newspapers or other media. Information from the survey permits us to explore how the jobs they took compared with their expectations, and administrative data provide information about their earnings.

### **What sort of job did women expect to find, and what jobs did they take?**

Asked about their job hunting, about 90 percent of women said that they had been looking for full-time, permanent work. Their wage expectations were modest—on average, just under \$8 an hour. Over 80 percent wanted to work a 40-hour week, and another 8 percent 30–39 hours.

How did women's actual jobs over the past year measure up to their expectations? On average, the typical hourly wage was just 37 cents less than the desired hourly wage, and for nearly half of respondents, it was equal to or more than the desired wage. When asked how satisfied they were with their wages in the past year, almost three-quarters of the women reported themselves to be highly satisfied or satisfied. None reported being highly dissatisfied.

### **The difference between expectations and actual earnings**

If welfare recipients with temporary work experience had been able to achieve their goals of full-time work at about \$8 an hour, their annual earned income would have been about \$16,000. Yet in 2000, only 3 of the 74 women earned \$16,000 or more. Average earnings for 2000 were just over \$5,000. The average reported work week, 35.8 hours, explains only \$2,000 of the difference between what women hoped to earn and what they actually earned.

Our information suggests that time without work between assignments played a large role in the deficit between expected and actual earnings. Almost 40 percent of the women said that the typical temporary assignment lasted less than 3 months—most jobs, indeed, lasted less than a month. For almost half, the job assignment lasted 3–6 months. Only 9 percent found a temporary assignment that lasted a year or more. About half said they were without a job assignment for a month or less, and the remainder reported that the longest time without a job was 3 or more months. Reports of quarterly earnings show that 57 percent of the sample had at least one quarter with no reported earnings. And sector of employ-

ment matters: those whose temporary work assignments were in the services industry saw lower earnings than those who, for example, received clerical assignments.

Although these comparisons suggest that temporary help employees were frustrated in their goals, it is worth keeping these experiences in perspective. We know that, whatever their industry of employment, TANF recipients' earnings are low, since this is a requirement of participation. Other workers must face problems of similar import.

### **Work circumstances and benefits**

Despite the failure to meet earnings expectations, women expressed few major dissatisfactions with their current employment circumstances. Two-thirds thought it likely (though not "highly likely") that their current temporary assignment would lead to a permanent position; 57 percent were also "satisfied" with their opportunities for permanent employment. The greater part of the women expressed high levels of satisfaction with the kind of work they were doing, with their relationships with supervisors and coworkers, and, to a lesser degree, with their hours and work locations.

One important aspect of employment that evoked much criticism was the lack of employment benefits such as health insurance and paid vacations. More than two-thirds were dissatisfied or highly dissatisfied with this aspect of their work. Only about 8 percent received fully or partially paid medical insurance—a percentage consistent with national data.<sup>9</sup> About the same percentage received vacation or holiday pay, and only negligible numbers received any other kind of benefit. In our sample, over half said they did not receive any benefits in the past year because none were offered; another 35 percent said that they had not worked long enough to become eligible. Only 15 percent said that they could not afford their share of the cost or did not need or want the benefits offered.

### **What opportunities for training or advancement did women receive?**

About 80 percent of the women had received some kind of formal or informal training in the past year. This is very similar to the national average; one study has found that some 75 percent of temporary help services firms offer training and that it is typically offered before or in between assignments, on the worker's own time. Moreover, both our work and other national studies show that temporary firms favor those with high school degrees and more work experience.<sup>10</sup>

Women in the survey most commonly received occupational safety training, though this had been offered to only 43 percent of the women in the past year. Basic skills training—elementary reading, writing, math, and language—was offered to nearly 40 percent (among those surveyed, 38 percent had less than 12 years schooling). One-third or fewer of the group received any occupation-

specific training, mostly in clerical, computer, sales and customer relations, and communication skills.

Did the women see any future benefits from these training activities? About 70 percent were optimistic that training might lead to future advancement. Two-thirds said that training was mandatory to keep their current job, but also that they had gained valuable skills to improve their performance. But barely 40 percent identified tangible benefits from the training, such as a move from a temporary to a permanent position, more work hours, or a higher rate of pay.

## Is temporary employment a good thing or a bad thing for welfare recipients?

What then, can we conclude about the consequences of temporary services jobs for this group of disadvantaged women? The good news from the North Carolina survey appears to be that, in general, women who take such jobs are not being pushed into them and are not unhappy with their experiences on the job. Nor does the evidence suggest that those who take temporary jobs are stuck at the bottom of the wage ladder: earnings increases over two years are greater than in other low-wage job sectors, and job mobility is high and frequently positive. The TANF participants in the survey did not earn much less than comparable recipients in other types of jobs. Despite their greater disadvantages, their employment experiences and outcomes are not markedly inferior to those of welfare participants taking jobs in other sectors or more advantaged workers taking temporary sector jobs.

Among welfare recipients, those with jobs—whether in temporary help firms or in direct employment environments—face much better prospects than do those without jobs, in terms of both their future earnings and their chance of leaving welfare. If the availability of temporary help employment induces even a small growth in employment among recipients, it is clear that temporary help jobs are on net beneficial. Other studies we have conducted suggest that such jobs may provide employment opportunities for other disadvantaged workers as well. Nonetheless, it is useful to recognize that none of these jobs appear to offer a direct route out of poverty. TANF recipients who obtain employment remain subject to chronic problems of low wages and job instability. ■

---

<sup>1</sup>The research presented in these papers focuses specifically on employees in the temporary help sector, identified by SIC code=7363 in their unemployment insurance records. These employees are one category of contingent workers (including others such as part-time, short-term contract, seasonal employees, etc.), as defined by the Bureau of Labor Statistics. See, e.g., U.S. Department of Labor, Bureau of Labor Statistics, *Contingent and Alternative Employment Arrangements*, February 2001, Table 6; <<http://www.bls.gov/news.release/conemp.t06.htm>>

<sup>2</sup>K. Barker and K. Christensen, eds., *Contingent Work: American Employment Relations in Transition* (Ithaca, NY: Cornell University Press, 1998).

<sup>3</sup>Nevertheless, temporary services employees still accounted for only around 2 percent of total employment in 1996. See M. Pressler, "Temp Workers Credited with Cushioning Recession's Blow," *Washington Post*, July 7, 2002; L. Segal, "Flexible Employment: Composition and Trends," *Journal of Labor Research* 17 (Fall 1996): 525–41. A general review of the temporary help industry and the shifting emphasis on work for welfare recipients is D. Autor and S. Houseman, "The Role of Temporary Employment Agencies in Welfare to Work: Part of the Problem or Part of the Solution?" *Focus* 22, no. 1 (Special Issue, 2002):63–70.

<sup>4</sup>For discussions of these and other issues see R. Blank, "Contingent Work in a Changing Labor Market," in *Generating Jobs: How to Increase Demand for Less-Skilled Workers*, ed. R. Freeman and P. Gottschalk (New York: Russell Sage Foundation, 1998), pp. 258–94; L. Pavetti and G. Acs, "Moving Up, Moving Out, or Going Nowhere? A Study of the Employment Patterns of Young Women and the Implications for Welfare Mothers," *Journal of Policy Analysis and Management* 20 (2001): 721–36; L. Segal and D. Sullivan, "The Growth of Temporary Services Work," *Journal of Economic Perspectives* 11 (Spring 1997):117–36; J. McAllister, "Sisyphus at Work in the Warehouse: Temporary Employment in Greenville, South Carolina," in *Contingent Work*, ed. Barker and Christensen, pp. 221–42.

<sup>5</sup>S. Cohany, "Workers in Alternative Employment Arrangements: A Second Look," *Monthly Labor Review* 121 (November 1998): 3–21; M. Morris and A. Vekker, "An Alternative Look at Temporary Workers, Their Choices, and Growth in Temporary Employment," *Journal of Labor Research* 22 (Spring 2001): 373–90.

<sup>6</sup>The federally financed JTPA provided job skills training with an emphasis on disadvantaged workers. (It was replaced with a similar program, the Workforce Investment Act, in 1998.) Participants in JTPA registered for services as adult or dislocated workers. The adult program was means tested and limited to individuals whose income in the previous six months was below specified levels. Dislocated workers have generally lost employment through firmwide layoffs. Most individuals who receive unemployment insurance are required to register for job services under the Employment Service, which also serves employed individuals seeking a better job. In 1998, state Employment Services were required by federal law (the amended Wagner-Peyser Act) to become part of the One-Stop services delivery system, with the goal of assisting job seekers and employers alike and, in some areas, providing job training and related services.

<sup>7</sup>The 1993 findings are not reported in this article.

<sup>8</sup>In Figure 1, we compare only women workers in the three employment programs. Earnings for males are somewhat lower in the temporary sector. Nevertheless, for every male who holds a job in the temporary sector, there are four or five who hold a temporary help job and a job in another sector, and earnings for these men are higher than earnings in almost every other category. Temporary help employment appears usually to be supplementary for men, either as a second job or because it is held for only a portion of the quarter, followed by a rapid move to another job. These data suggest that, for men, temporary sector employment may facilitate continued involvement in other forms of employment.

<sup>9</sup>Cohany, "Workers in Alternative Employment Arrangements."

<sup>10</sup>Survey questions about training used categories established by the Bureau of Labor Statistics. On training in general, see D. Autor, "Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing," *Journal of Labor Economics* 21 (January 2003): 1–42; and "Why Do Temporary Help Firms Provide Free General Skills Training?" *Quarterly Journal of Economics*, forthcoming; H. Frazis, M. Gittleman, and M. Joyce, "Correlates of Training: An Analysis Using Both Employer and Employee Characteristics," *Industrial and Labor Relations Review* 53 (April 2000): 443–68.



# Welfare policy choices in the states: Does the hard line follow the color line?

Joe Soss, Sanford F. Schram, Thomas P. Vartanian, and Erin O'Brien

---

Joe Soss is Associate Professor of Political Science at the University of Wisconsin–Madison and an IRP affiliate; Sanford F. Schram is Visiting Professor of Social Work and Social Research and Thomas P. Vartanian is Associate Professor of Social Work and Social Research at Bryn Mawr College; and Erin O'Brien is Assistant Professor of Political Science, Kent State University.

---

In the 1990s, federal welfare reform shifted control over many aspects of social provision down to the state level. Some observers predicted the states would act as “laboratories of democracy” seeking out new and better ways to solve problems related to poverty. Others suggested the states would “race to the bottom” in an effort to pare costs and avoid becoming a comparatively generous “magnet” for low-income families. State policy choices, however, are more than just efforts to solve problems, and they can rarely be explained solely by the urge to minimize costs. They are, at root, *political* decisions, reflecting the wide range of values, beliefs, demands, and constraints that shape governance in a diverse democracy.

Under welfare reform, state lawmakers and their constituents confronted fundamental questions about how and when government should extend aid to the poor. As states responded to the new policy prescriptions dictated by the federal welfare law, the political process that sets the terms of relief for poor families was replicated many times, in different places, under different configurations of political forces. In the research summarized here we explore the sources of these decisions, taking advantage of the opportunity to observe how state responses differed within a single time period under a single federal mandate.<sup>1</sup>

Until recently, most quantitative research on the political roots of state welfare policies sought to explain differences in benefit levels and spending patterns.<sup>2</sup> In contrast, we analyze variation in the *institutional form* of welfare provision—the rules and penalties that condition access to resources and structure the treatment citizens receive in government programs. In the 1990s, public officials showed renewed interest in using program rules as tools to modify poor people’s behaviors. The 1996 federal law emphasized such aid requirements, and as states responded, their policy changes focused less on the amounts of relief offered than on the terms on which aid

would be given. Some states adopted a moderate course in this area of reform; others used their enlarged discretion to pursue relatively stringent program rules, with punitive measures for those who did not comply. Here we seek to understand why some states were more likely than others to “use their new authority to limit access to social provision and, most especially, to shift the balance in welfare policy design from rights to obligations.”<sup>3</sup> Our findings, as we describe below, suggest that state choices can be traced in significant ways to race-related factors. To a degree that some may find surprising, the “hard line” in welfare reform appears to have followed the color line.

## Ending permissiveness, getting tough: Policy choice in the states

With welfare reform, states gained more authority over eligibility rules and administrative procedures than they had enjoyed since the 1960s. Proponents acclaimed the new, less regulatory environment, seeing Temporary Assistance for Needy Families (TANF) as a “devolution revolution” that would liberate states from constricting federal rules and allow them to create more effective poverty policies. In a sense, however, the term “revolution” is a bit misleading: states did not gain unprecedented freedom to shape policy. Rather, they recouped many forms of discretion they had lost to the federal government during the welfare rights era of the 1960s. In addition, the federal law imposed new mandates and constraints on the states. For example, it set quotas on the percentage of adult recipients who must participate in “work-related activities,” and defined these activities rather narrowly. Likewise, it imposed limits on the length of time states could provide cash assistance to residents.

In principle, states can now make benefits more accessible to poor families and enable clients to pursue new opportunities. Indeed, most states have passed policies that offer clients new services, supply transitional benefits, and allow clients to keep higher amounts of earnings. The political momentum toward welfare reform, however, was fueled largely by the belief that the existing program, Aid to Families with Dependent Children (AFDC), was too permissive. As reform advocate Lawrence Mead puts it, “Today ‘welfare reform’ largely means that the government seeks to supervise poor citizens”; public officials have embraced the idea that welfare provision is partly about “telling the poor what to

do.”<sup>4</sup> The federal legislation of 1996 reflected and reinforced this view, with its emphasis on ending welfare dependency and encouraging marriage. In consequence, when it came to program requirements, policy innovation in the states leaned in a restrictive direction. Many states stuck close to the basic federal rules, but the states that deviated from the baseline requirements (for work, time limits, and so on) mostly used their new authority to limit rather than expand access to cash assistance, though some states did increase “work supports” such as child care and transportation assistance.

In selecting policy choices for analysis, we emphasized two goals: covering the major domains of “get-tough” welfare reform (i.e., rules and penalties intended to counter permissiveness) and identifying the policies most salient in public debate and most widely considered by state governments. Surveying scholarly books and articles, policy reports, legislative materials, and mass media, we found a consistent emphasis on four key areas in which federal lawmakers sought to end permissiveness. Each defined a specific policy choice for state governments.

- Imposing obligations in exchange for assistance: states decided whether to demand work from recipients earlier than the federal requirement of 24 months.
- Ending long-term program usage: states decided whether to adopt a lifetime eligibility limit shorter than the federal requirement of 60 months.
- Changing social behavior, especially reproductive choices: states decided whether to impose a family cap denying additional benefits to children conceived by women receiving assistance.
- Imposing meaningful penalties: states decided whether to choose a weak, moderate, or strong sanctions policy for recipients’ infractions.

As a group, these program rules define the key terms of participation for citizens seeking aid under TANF; they also capture some of the fundamental goals of 1990s welfare reform. But there are differences among them. Family caps, time limits, and work requirements are widely viewed as complementary tools for combating permissiveness; each is tied to a distinct goal: deterring childbirth among recipients, combating welfare dependency, and demanding work. Sanctions, in contrast, are punitive tools for enforcing a broad range of program rules that may have diverse political constituencies. Because of this greater reach and ambiguity, one might expect sanction choices to be subject to a wider range of political influences.

## The forces that shape state policy choices

The roots of policy choice may be traced in many directions. In the case of welfare reform, we confront a range

of plausible explanations. Our analysis tests hypotheses derived from six approaches that offer contrasting, though not mutually exclusive images of welfare policy as a forum of moralistic problem-solving, a site of ideological conflict, an arena for policy innovation, an outcome of electoral politics, a mechanism of social control, and a domain of racialized politics.

### Problem-solving and morality politics

Critics of welfare in the 1980s and ’90s echoed a long tradition of observers who argued that permissive policies invited personal irresponsibility. AFDC regulations allowed poor women to spend too long on the welfare rolls, discouraged the formation of two-parent families, and perhaps encouraged childbirth among unmarried women and teens. More paternalistic TANF policies aimed to achieve two key goals: changing “deviant” behavior and affirming majoritarian social norms regarding work, marriage, and family.

Political discourses about “dependency” and “illegitimacy” frequently have a weak connection to actual patterns of behavior. Policymakers concerned about such issues, however, may respond at least partly to real conditions, adopting tougher policies when confronted with behavior patterns that deviate more sharply from prevailing values. To counter long-term welfare usage, for example, states with higher caseload-to-population ratios under AFDC might adopt more restrictive policies under TANF. And to counter what is seen as irresponsible reproductive behavior, states with a high rate of births to unmarried mothers might adopt more restrictive TANF policies.

### Welfare liberalism: Ideology and practice

Beliefs about the proper role of government define an important cleavage in U.S. welfare politics. Relative to conservatives, liberals have historically favored a larger government role in protecting vulnerable people against the vagaries of the labor market. Accordingly, liberals have tended to support more generous benefits and more inclusive eligibility standards in public assistance programs. Although some liberal policymakers in the 1990s were dissatisfied with the existing system—and some even joined calls to “end welfare as we know it”—conservative officials took the lead in promoting the toughest new policies.

Two hypotheses are suggested by the durable relationship between welfare liberalism and specific policy preferences. First, states that earlier had adopted relatively liberal AFDC policies might be expected to continue on a more liberal path after 1996, whereas states that worked to keep their caseloads down under AFDC might simply deepen their efforts under TANF. In contrast to the dependency hypothesis, this “continuity hypothesis” predicts that states with higher caseload ratios under AFDC might adopt *less* restrictive TANF policies.

Second, previous research has shown that state governments vary significantly in their ideological orientation and that conservative states are more likely to reduce benefit packages and to restrict eligibility for public assistance.<sup>5</sup> To the extent that state policy choices are shaped by the ideologies of current elected officials, we might hypothesize that states with more liberal political representatives would adopt less restrictive TANF policies.

### **Policy innovation**

Research suggests that states can be characterized by a fairly consistent orientation toward policy innovation: some tend to be leaders, some laggards. TANF work requirements, time limits, family caps, and sanctions share a get-tough quality, but within the constraints imposed by federal legislation they also represent real opportunities for policy innovation. Thus our hypothesis suggests that states with a stronger propensity toward innovation will be more likely to adopt restrictive TANF policies.<sup>6</sup>

### **Electoral politics**

V. O. Key's classic analysis of Southern politics suggests that two features of state electoral systems can have an important influence on policy choice. The first is the degree of interparty competition. When political parties are more evenly matched and forced to compete with one another for voters, the policy process will be more responsive to the needs of the disadvantaged. Second, in states where higher proportions of low-income voters go to the polls, politicians will be more responsive to the needs of the poor and working class.<sup>7</sup> If these arguments hold, we would expect less restrictive welfare policies in states with higher levels of interparty competition as well as in states with higher turnout rates among low-income voters.

### **Controlling the poor**

Complex societies rely on a range of instruments to maintain social order, combating disorder through a mix of informal controls, in families, neighborhoods, and communities, and formal controls deployed by the state. One strand of social control theory identifies welfare systems as auxiliary institutions that serve the broad needs of primary institutions related to states and markets.<sup>8</sup> When hard economic times combine with civil unrest, relief is readily extended to mollify the poor and maintain legitimacy for the state. Under better economic conditions, access to public aid is restricted in order to push potential workers toward jobs and ease the pressure of tight labor markets on employers. Given the strong economy of the late 1990s and the relative absence of civil unrest, states with tighter labor markets would, by this hypothesis, be more likely to enforce work and limit access to welfare benefits. Accordingly, we would expect states with lower unemployment rates to adopt more restrictive TANF policies, especially for work requirements.

Social control theory also suggests that TANF policies may reflect a general preference for the use of formal mechanisms to enforce order. Heavier reliance on institutional tools such as incarceration may indicate that a state is more willing, politically, to crack down on marginal or deviant groups, in general adopting "tougher" solutions to community problems such as drug abuse. In the United States during the 1990s, state officials passed stiffer penalties for criminal behavior and increased funding for prison construction and maintenance; incarceration rates soared, though the rate of increase varied considerably among the states.<sup>9</sup> Thus our hypothesis predicts that states with larger increases in incarceration from 1990 to 1996 will make more restrictive TANF policy choices.

### **Race and ethnicity**

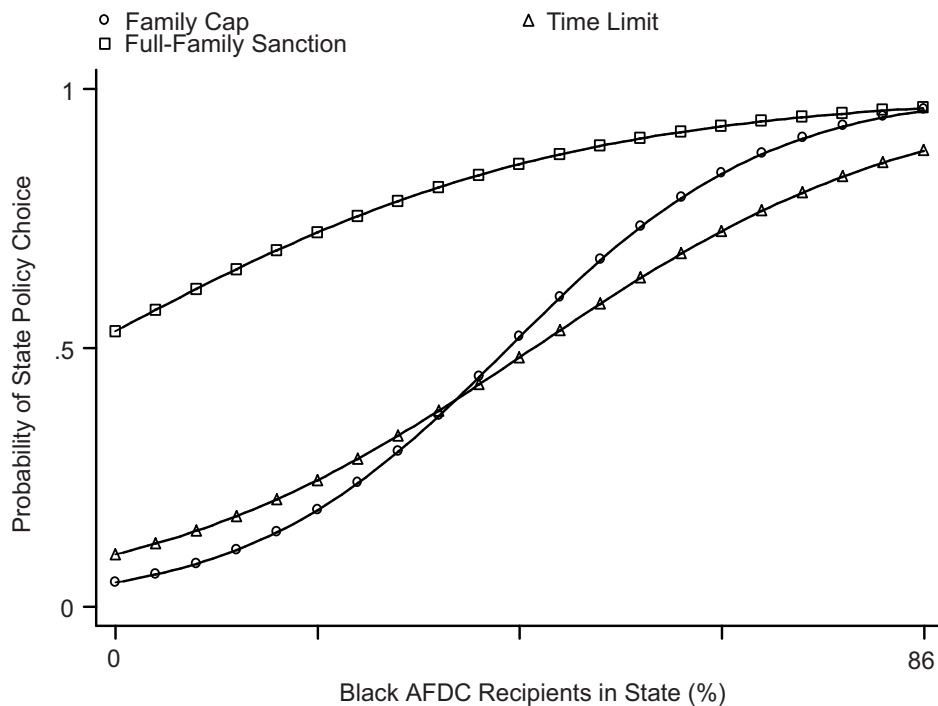
The entwining of race relations and welfare provision has a long and troubled history in the United States. In debates over the Social Security Act of 1935, southern members of Congress managed to exclude domestic and agricultural workers from social insurance coverage, effectively channeling people of color into public assistance programs controlled at the state level. Many scholars argue that state administration of these programs continued to be shaped by race into the 1990s, and that programs associated with nonwhite clients were more likely to be saddled with popular hostility and punitive rules. Research on welfare spending and benefit levels in the states provides mixed but suggestive evidence. Cash welfare benefits, for example, were systematically lower in states where black recipients made up a higher percentage of the caseload, even when other relevant factors are taken into account.<sup>10</sup>

Racial differences might, then, have played a key role in shaping the terms of public relief after 1996. Most existing research has focused on blacks as the group most likely to be targeted by less generous policies. But as the percentage of Hispanics in the U.S. population continues to rise, attitudes toward welfare and policy might become associated with perceptions of Hispanics as well as blacks. Our hypothesis, therefore, is that tougher TANF policies were likely to be adopted in states where blacks *or* Hispanics made up a higher proportion of the welfare caseload in 1996.

### **State policy choices: Testing the hypotheses**

Based on the hypotheses just described, we employed a set of 10 independent, state-level variables to structure our analysis: the unmarried birth rate, the caseload-to-population ratio, government ideology, interparty competition, low-income voter turnout, the unemployment rate, change in the incarceration rate, the percentages of Hispanics and blacks, and propensity to welfare innovation.

We began our analysis by assuming that TANF policy choices constituted a single "package" of stringent poli-



**Figure 1. The effect of black caseload percentage on welfare policy choices in a hypothetical “average” state.** The black percent of AFDC recipients is shown for the full observed range.

**Source:** J. Soss, S. Schram, T. Vartanian, and E. O’Brien, “Setting the Terms of Relief: Explaining State Policy Choices in the Devolution Revolution,” *American Journal of Political Science* 45, no. 2 (April 2001): 378-95.

cies driven by a coherent set of political forces. Our empirical analysis based on a single equation did, indeed, offer some explanatory power. It told a relatively simple story based in race and ideology: states were significantly more likely to make restrictive policy choices if they had conservative governments and if blacks made up a higher percentage of AFDC recipients. No other factors stood out.

As our earlier discussion suggests, however, the policies we selected express somewhat different political motives and so may actually be influenced by different configurations of political forces. This speculation is strengthened by the fact that states did not in general adopt the entire package of restrictive policies. The median state, indeed, adopted only one, and only five states adopted all four. We thus undertook separate analyses for each of our four policies. Because our outcome measures were dichotomous for work requirements, family caps, and time limits, we employed logit analysis in each instance. By contrast, our sanctions measure, with its ascending values of weak, moderate, and strong, made an ordered logit model more appropriate.

The analyses showed that, in each area, restrictive policy choices were systematically related to the state-level characteristics we identified; 9 of our 10 independent variables were statistically significant in at least one area.

For some of the individual hypotheses, the patterns of influence appeared quite strong. To begin with, family caps and strict time limits seemed to arise from virtually identical processes, with each being closely tied to race. All else equal, these policies were significantly more likely in only two kinds of states: those with a higher percentage of blacks and those with a higher percentage of Hispanics in their AFDC caseloads. Our analysis of work requirements also revealed a tight cluster of determining factors—in this case related not to race but to our social control hypotheses. States with larger increases in incarceration from 1990 to 1996 were significantly more likely to adopt strict work requirements, as were states with tighter labor markets.

Sanctions, as we expected, evoked a more complex set of relationships. Race was again prominent: states with larger proportions of blacks in their AFDC caseloads were significantly more likely to adopt strict sanctions. So too were states with conservative governments, less vigorous party competition, higher unmarried birth rates, and smaller AFDC caseloads, and states that were policy innovators. The only characteristic that appeared to have no significant effect on any area of welfare policy was low-income voter turnout—a finding we return to below.

Thus, in three of our four policy domains, the racial composition of welfare recipients turned out to be a sig-



nificant predictor of state choices, even after the effects of other state differences were taken into account. To gauge the scale of these effects, we made use of an interpretive procedure developed by Gary King and colleagues.<sup>11</sup> First, we created a “hypothetically average” state by setting all our independent variables (except the racial composition of the welfare rolls) at their mean value. We then estimated how the probability of this state making a particular policy choice would change as we shifted the black percentage of its welfare rolls from a low to a high value—assuming that no other state characteristic changed at all. Figure 1 shows that as the percentage of black recipients rises across its full range, so too does the probability that states will adopt strong (full-family) sanctions, institute strict time limits, and introduce a family cap policy. The estimated effects of having more Hispanics on the rolls are similarly large. Most dramatically, as the percentage of black recipients rises from moderately low to moderately high, this “average” state’s chance of instituting a family cap rises from 9 percent to 75 percent; the same shift for Hispanics produces a parallel change from 19 percent to 63 percent.<sup>12</sup>

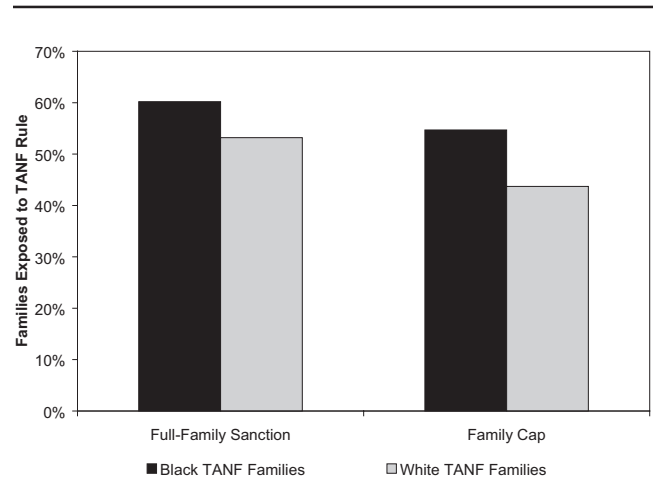
This analysis suggests that far from being a pure technocratic search for policy solutions, the construction of public assistance policy in the new era of welfare reform continues to be shaped by forces that are familiar in the history of American welfare politics—particularly race, ideology, and the control of representative institutions. At the same time, it also underscores that important insights may be lost if diverse policy choices are lumped together in a single index and analyzed as an undifferentiated move away from permissive program rules.

### The importance of race

To the question of whether “hard line” policy choices under welfare reform have followed the “color line” in the states, our answer must be a qualified—but unequivocal—yes. In two of our four policy areas, time limits and family caps, we find very strong connections to the black and Hispanic proportion of state welfare rolls, and we find no relationship to any other factor. These results focus attention squarely on race as a central problem for contemporary welfare reform.

In contrast, and despite considerable evidence that welfare politics is bound up with racial stereotypes regarding work effort,<sup>13</sup> our analysis provided no evidence that work mandates have been linked to the racial composition of the rolls. Instead, it pointed to two factors suggested by social control theory: the tightness of state labor markets and the rate of increase in state incarceration rates.

Sanctions policy provided the best example of how racial effects may intersect with other forces to shape state policy choices. Strong sanctions were especially likely in states with large numbers of black welfare recipients. But



**Figure 2. Interstate variation and welfare inequity: Exposure to strict TANF rules and race of the family, 2000.** (Authors’ analyses.)

they were also significantly more likely in states with several other characteristics: conservative governments, less vigorous party competition, higher unmarried birth rates, a history of policy innovation, and smaller AFDC caseloads. We may speculate that the popularity of sanctions is due to their versatility: they raise the stakes for participants who fail to follow any of a variety of new welfare initiatives, and in legislative debates they may not be tied to any specific program goal other than achieving “compliance.”

On balance, these results suggest that policy devolution created new openings for racial distortions in U.S. welfare policy. Indeed, because state policy choices have tracked so closely with the racial composition of welfare rolls, black recipients nationwide are now more likely than their white counterparts to be participating under tough welfare policies. For example, a black recipient who conceives a child is now more likely than a white recipient to live in a state that offers no additional aid for the child. Likewise, a black recipient who misses a meeting with a caseworker is now disproportionately likely to live in a state where this single infraction causes family benefits to be terminated (Figure 2). It is hard to know how such different rates of *exposure* to sanction policies actually translate into rates of *being sanctioned*. After all, many intermediary processes are involved within each state. But the numbers do suggest that, if full-family sanctions were applied equally to blacks and whites in each state, national rates of sanctioning would be higher for black recipients than for their white counterparts.

### Reflections and extensions: Revisiting our findings

The research reported in this article was conducted immediately after the passage of welfare reform and has circulated among political scientists for some time now. It is



well worth asking how the findings have fared and, as well, whether our findings for race might have represented an evanescent phase—an echo of past policies that disappeared as the system matured. Our central findings regarding the connection between racial composition and hard-line welfare policies proved to be very robust in the face of our own efforts to dislodge them. Statistical tests indicate that they do not, for example, simply reflect the distinctiveness of the American South. Nor do they fade when subjected to additional controls for social behaviors that vary across racial groups.

Over the past few years, the race-related findings have been corroborated and extended by a number of different studies. An analysis by Kent Weaver and Thomas Gais underscored that racial effects are tied to restrictive and punitive welfare policy “sticks,” not beneficial “carrots”; Richard Fording showed that similar connections to race could be found in state policy choices during the AFDC waiver era; and most recently, Matthew Fellowes and Gretchen Rowe have produced a striking extension of the results to state choices regarding TANF eligibility rules and the degree of flexibility in state welfare-to-work requirements.<sup>14</sup> Indeed, Fellowes and Rowe’s analysis offers a persuasive demonstration of racial effects in the one area, work requirements, where we observed none.

Beyond the issue of race, it is also worth asking about findings for other political factors reported here—especially our lack of significant results for some variables we expected to play a role. First, our analysis suggested that turnout rates among low-income voters had no discernible influence on TANF policy choices in the states. Yet in a country with very high wage and income inequality, it struck us as unlikely that income and class divisions played no role. Recent evidence suggests we had the right expectations but the wrong measure. Fellowes and Rowe’s analysis demonstrates that it is not absolute levels of low-income voting participation, but rather *class bias* in participation (low-income turnout relative to high-income turnout) that actually matters. And it matters in precisely the direction one would expect: restrictive TANF eligibility and work policies are significantly more likely in states where electoral turnout is more biased toward high-income, as opposed to low-income, residents.

Second, in supplemental analyses reported in our 2001 article, we found that none of our four TANF policies were significantly affected by state resource levels, as measured by per capita income, or by the policies of neighboring states. Although brimming state coffers do not guarantee generous welfare policies, states with more resources may provide higher benefits than less wealthy states.<sup>15</sup> With the shift to block-granted funding, however, many observers became concerned that states could compete to avoid becoming a relatively generous “welfare magnet” (attractive to low-income people from other states), and accordingly would engage in a “race to the

bottom.” This concern suggested wealthy states might not use their resources to offer greater benefits and would, instead, try to keep pace with the least generous policies of their neighbors. By contrast, our results suggested that neither state resources nor neighboring states’ policies exerted a discernible effect on TANF policy choices. Similarly, Fellowes and Rowe find little evidence that state policies responded either to abundant resources in their own state or to the restrictiveness of policies in neighboring states—although their evidence does suggest that states with more slack resources may be more likely to allow clients to continue receiving benefits, even if they are not fully meeting traditional work requirements.

“Implementation of TANF may have radically changed the power structure of welfare politics by shifting a substantial amount of policy authority to the states,” Fellowes and Rowe conclude, “but it has not changed the politics of welfare policy. . . . For welfare clients, this empowerment of state representatives has meant that their experience on welfare will vary even more widely between the states.”<sup>16</sup>

## The future of state policy experimentation

Freedom from the tether of federal regulation hitches welfare policy to the social and political forces that operate in each of the American states. In the first half of the 20th century, such state discretion was used for a variety of social purposes. Welfare policies were used to control women’s sexual and parental behaviors (e.g., the “man in the house” rule), and to regulate the labor activities of the poor, absorbing them during slow economic times and impelling work when more hands were needed in the factories or on the farms. Likewise, many states administered benefits in a racially biased manner, withholding aid from people of color and using program rules to punish those who violated race-specific, segregation-enforcing norms of social conduct.

This era largely came to an end with the welfare rights victories of the late 1960s. Today, our post-civil-rights political and legal context makes it unlikely that the TANF system could replicate the worst of earlier practices. Yet TANF policies remain deeply entwined with the politics of gender, class, and race. Recent policies include rules that explicitly target women’s sexual and familial behaviors, primarily related to childbearing among unmarried women. Meanwhile, work enforcement remains central. Finally, our central conclusion in this article is that race and ethnicity continue to be major influences on the terms of relief state governments set for poor families. Indeed, as caseloads have become slightly more skewed toward people of color under the TANF program, people of color (as shown in Figure 2) have become more concentrated in states adopting the stricter policies.

Based on our own study and the research that has followed, we are convinced that the racial composition of the welfare rolls has a significant relationship to state policy choices that “get tough” on TANF recipients. But we also believe this statistical correlation raises more questions than it answers. What does it mean? Do lawmakers operate from different assumptions when they see their policy targets as people of color? Do public preferences change in response to the composition of a state’s welfare rolls, and do these constituent views shape policy outcomes? Does the observed relationship between race and TANF policies reflect divergent patterns of political development in states with different levels of racial-ethnic diversity? The search for a compelling causal account continues. What seems more certain, however, is that the “problem of the color line” and troubling questions of racial justice remain very much with us as we move into the second decade of America’s new era of welfare provision. ■

---

<sup>1</sup>This article summarizes research presented in greater detail in two articles by J. Soss, S. Schram, T. Vartanian, and E. O’Brien, “Setting the Terms of Relief: Explaining State Policy Choices in the Devolution Revolution,” *American Journal of Political Science* 45, no. 2 (April 2001): 378–95, and “The Hard Line and the Color Line: Race, Welfare, and the Roots of Get-Tough Reform,” Chapter 9 in *Race and the Politics of Welfare Reform*, ed. S. Schram, J. Soss, and R. Fording (Ann Arbor: University of Michigan Press, 2003).

<sup>2</sup>See, e.g., C. Howard, “The American Welfare State, or States?” *Political Research Quarterly* 52 (1999): 421–42; R. Plotnick and R. Winters, “A Politico-Economic Theory of Income Distribution,” *American Political Science Review* 79 (1985): 458–73.

<sup>3</sup>S. Mettler, “States’ Rights, Women’s Obligations: Contemporary Welfare Reform in Historical Perspective,” *Women & Politics* 21 (2000): 1–34.

<sup>4</sup>L. Mead, “Telling the Poor What to Do,” *Public Interest* 132 (Summer): 97–112.

<sup>5</sup>M. Rom, “Transforming State Health and Welfare Programs,” in *Politics in the American States*, ed. V. Gray and H. Jacobs (Washington, DC: CQ Press, 1999).

<sup>6</sup>As our measure of state propensity toward welfare policy innovation, we employ the first year in which each state requested a policy waiver of any kind under the AFDC program. Although many of these early waivers leaned in a direction more liberal than the choices we examine here, a policy-innovation perspective suggests that the states submitting early waiver requests would be more likely to institute stringent rules and sanctions under TANF.

<sup>7</sup>V. Key, Jr., *Southern Politics in State and Nation* (New York: Alfred A. Knopf, 1949); F. Piven and R. Cloward, *Why Americans Don’t Vote* (New York: Pantheon, 1988).

<sup>8</sup>F. Piven and R. Cloward, *Regulating the Poor: The Functions of Public Welfare*, new ed. (New York: Vintage, 1993).

<sup>9</sup>D. Rose and T. Clear, “Incarceration, Social Capital, and Crime: Implications for Social Disorganization Theory,” *Criminology* 36 (1998): 441–79; J. Lynch and W. Sabol, “Did Getting Tougher on Crime Pay?” in *Crime Policy Report* (Washington, DC: Urban Institute Press, 1997).

<sup>10</sup>See the following chapters in *Race and the Politics of Welfare Reform*, ed. Schram and others: R. Lieberman, “Race and the Limits of Solidarity,” R. Hero, “Racial/Ethnic Diversity and States’ Public Policies,” and M. Johnson, “Racial Context, Public Attitudes, and Welfare Effort in the American States.” See also J. Quadagno, *The Color of Welfare: How Racism Undermined the War on Poverty* (Oxford: Oxford University Press, 1994); C. Howard, “Field Essay: American Welfare State or States?” *Political Research Quarterly* 52 (1999): 421–42; L. Orr, “Income Transfers as a Public Good: An Application to AFDC,” *American Economic Review* 66, no. 3 (1976): 359–71.

<sup>11</sup>G. King, M. Tomz, and J. Wittenberg, “Making the Most of Statistical Analysis: Improving Interpretation and Presentation,” *American Journal of Political Science* 44 (2000): 347–61.

<sup>12</sup>“Moderately low” to “moderately high” are defined as one standard deviation below the mean and one standard deviation above the mean, respectively. As Figure 1 suggests, shifting the percentage across a wider range of values would produce even larger estimates of the predicted change in probability.

<sup>13</sup>See M. Gilens, “How the Poor Became Black,” in *Race and the Politics of Welfare Reform*, ed. Schram and others.

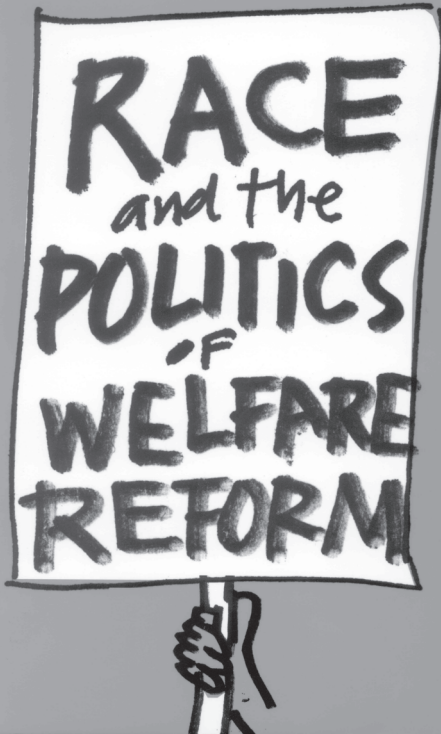
<sup>14</sup>T. Gais and R. Weaver, *State Policy Choices under Welfare Reform*, Policy Brief No. 21, Brookings Institution, Washington, DC, April 2002; R. Fording, “‘Laboratories of Democracy,’ or Symbolic Politics?” in *Race and the Politics of Welfare Reform*, ed. Schram and others; M. Fellowes and G. Rowe, “Politics and the New American Welfare State,” *American Journal of Political Science*, April 2004 (forthcoming).

<sup>15</sup>J. Tweedie, “Resources Rather than Needs: A State-Centered Model of Welfare Policy-Making,” *American Journal of Political Science* 38, no. 3 (1994): 651–72.

<sup>16</sup>Fellowes and Rowe, “Politics and the New American Welfare State.”

---

Sanford F. Schram, Joe Soss, and Richard C. Fording, Editors



## *Race and the Politics of Welfare Reform*

Sanford F. Schram, Joe Soss, and Richard C. Fording, editors

*Race and the Politics of Welfare Reform* investigates race in the United States and its distinctive effects on contemporary welfare politics. Over the last three decades, an impressive body of scholarship has brought the interplay of race and poverty politics into sharper focus. By bringing together diverse scholars with overlapping, substantive concerns, the book aims to foster a more integrated understanding of how race has shaped the past and present of U.S. social policy. Contributors explore the role of race in the historical roots of contemporary welfare dilemmas; questions of racial bias in the ways welfare is reported in the mass media and evaluated by the public at large; the racial dynamics of welfare policy discourse; racial influences on state welfare policy choices and implementation patterns; and the ways racial factors intersect with social policy developments beyond "welfare reform."

### **Contributors:**

**James Avery, Michael Brown, Richard Fording, Martin Gilens, Susan Gooden, Rodney Hero, Robert Lieberman, Martin Johnson, Erin O'Brien, Mark Peffley, Frances Fox Piven, Sanford Schram, Joe Soss, Holloway Sparks, Gertrude Steuernagel, Caroline Tolbert, and Thomas Vartanian.**

**Published 2003 by the University of Michigan Press, 839 Greene Street, Ann Arbor, MI 48104-3209  
Phone: 734-764-4388; Fax: 734-615-1540**

---

# The skills of American workers in today's labor market

Michael J. Handel

---

Michael J. Handel is Assistant Professor of Sociology at the University of Wisconsin–Madison and an IRP affiliate.

---

Are the education and skills of American workers adequate to the demands of today's jobs?<sup>1</sup> For at least three decades, academics and policymakers have heatedly debated this question. First came the perception, current in the 1970s, that the American labor force was overeducated relative to labor market needs. By this view, American workers faced a rough future in which their rising educational levels and aspirations for meaningful work had outstripped the demands of jobs available to them.

In the 1980s, however, the “skills glut” seemed rather rapidly to become a severe “skills deficit.” Labor economists, observing that the wage premium for a college education had reached record highs, concluded that the growth in the demand for skills had outrun the supply, resulting in a large growth in wage inequality.<sup>2</sup> Sociologists such as William Julius Wilson argued that the same labor market changes were contributing to the problems of the urban underclass, as the skills of minority workers lagged behind rising employer requirements.

A vociferous sector of popular and academic opinion has argued that the skills mismatch can only grow worse as the pace of economic change accelerates, the workplace becomes more dependent on information technology, and an ever larger share of jobs requires workers with higher levels of literacy and technical expertise and with greater capacity for flexibility and change. The blame for this perceived crisis was laid squarely upon the country's educational system.<sup>3</sup> The best known and perhaps the most overheated expression of these concerns was the 1983 report of the National Commission on Excellence in Education (1983), which declared, “Our Nation is at risk. . . . If an unfriendly foreign power had attempted to impose on America the mediocre educational performance that exists today, we might well have viewed it as an act of war.”<sup>4</sup>

The rapid about-turn in opinion regarding workers' skills and labor market needs is reason enough to look closely at the evidence. There appear, indeed, to be clear grounds for some skepticism about the current notion of a skills deficit. While education continued to be pilloried in the political and public discourse, many of the economic

problems that had fueled concern in the 1980s had diminished significantly by the mid-1990s. Economic growth was more robust before the recent recession than at any time in the previous thirty years, and U.S. economic dominance remains unchallenged internationally, though trade with low-wage countries continues to prompt anxiety over the shrinking numbers of middle-income jobs available to less-educated workers. The source of these current anxieties is in striking contrast to the 1980s, when concern focused on Japan and its highly skilled and motivated workers—all distant memories today. The late 1990s boom was built largely around new computer technologies, despite fears of a shortage of capable workers. In the late 1990s, before the recent recession, the wages of those in the lowest wage percentiles began rising for the first time in decades. None of this was predicted by theories of skills mismatch that emerged in the 1980s and remain the dominant discourse today.

Throughout, one sector of economic opinion—economists working in the segmented labor markets tradition—remained skeptical about the “skills decline” as a cause of labor market problems. They have always emphasized the importance of institutional conditions and employer characteristics in the wage determination process, rather than simply the human capital and other individual characteristics of workers. The real causes of inequality growth, they argued, lay in free-market government policies, declining manufacturing employment, deunionization, and managerial shortcomings in product quality, capital investment, work organization, and worker training that hindered competitiveness. As unionized manufacturing jobs that had provided middle-class incomes for less-educated workers were replaced by low-end service jobs, and as contingent work, outsourcing, and offshore production increasingly intruded into the primary labor market jobs, the quality of jobs declined and wage inequality grew. The skills mismatch discourse, from this perspective, was a classic example of “blaming the victims”—the workers who bore the brunt of low-road management strategies. In this view, the sometimes heated rhetoric over skills decline appears as a kind of “moral panic” that was entirely disproportionate to a sober estimate of the evidence.<sup>5</sup>

The speed with which the labor market reversed its course suggests, indeed, that swings in macroeconomic forces had a far greater effect on the nation's fluctuating fortunes in the 1980s and 1990s than did the modest trends in school quality or individual attainment. Were schools made the scapegoats for poor economic performance



whose real sources lay elsewhere? Any theory that attributes the 1980s decline to a skills mismatch or poor school quality must also account for the 1990s boom. In fact, even though many of the economic problems that helped fuel concern in the 1980s and early 1990s have stabilized or receded significantly since then, the sense of alarm regarding school quality has, if anything, grown. In this article, I explore some unresolved issues regarding whose education—and what skills—are believed to be deficient. I also examine the ways in which particular skill trends have been characterized and examine some explanations of the underlying causes offered by economists and sociologists. Because existing research and data sources leave so many questions about job skill requirements unanswered, I have also begun a new project to collect data on trends in the skills, technology, and work organization of U.S. jobs (see box).

### **Who lacks the necessary skills, and what skills are in short supply?**

There are clear conceptual as well as evidential difficulties in the skills mismatch thesis. The diagnoses its proponents offer have frequently been unclear or ambiguous.

Different groups of workers have at one time or another been cited as lacking appropriate work skills: cohorts educated since the 1960s, young workers, older workers, disadvantaged minorities, job-seekers with a high school education or less, and even the college-educated without a technical background.

The range of skills identified as deficient is correspondingly wide: eighth- or tenth-grade reading, writing, or math skills; college-level reasoning skills; underspecified “problem-solving” abilities; computer skills; “soft skills” such as teamwork or other interpersonal skills; or work-related attitudes, such as low work effort, poor demeanor, and unwillingness to follow direction, which are more personality variables than skills, strictly speaking. The vague but often heard complaint about inadequate “problem-solving” skills conflates both these views: employers seem to use the term to express dissatisfaction with workers’ cognitive skills *and* their lack of interest in exercising them.

One version of skills mismatch argues that there are serious deficits in basic or intermediate-basic (eighth to tenth grade) skills among those educated since the 1960s, when test scores fell and the quality of schools is believed to have declined. This is an implied cohort thesis: according

#### **New Labor Force Panel Survey to Track Changing Job Demands**

Researchers across several disciplines, including sociology, labor economics, education, and public policy analysis, have shown keen interest in questions relating to the changing nature of work. There is a widespread belief that the skill requirements of jobs are growing significantly as a result of the spread of computers and increased employee involvement in workplace decision-making. These developments are believed to have significant implications for a wide range of issues, such as the growth of wage inequality and the labor market prospects of less-educated workers, including those making the transition from welfare to work.

Progress on these questions remains limited by the lack of any detailed, nationally representative data on what people actually do at work. Researchers have only crude information on levels and kinds of job skill requirements, rates of change, and the dimensions along which job skills are changing.

To address this gap in knowledge, Michael J. Handel, assistant professor of sociology at the University of Wisconsin–Madison and an IRP affiliate, is conducting a nationally representative, two-wave panel survey of wage and salary workers (N=2,500). The surveys will collect information on various cognitive skills required on the job (reading, writing, math, problem-solving), interpersonal skills, manual skills, and physical demands. The surveys will also collect detailed information on computer use at work and participation in high-performance workplace practices, such as self-directed teams, as well as standard labor force variables.

Most of this information has never been collected before or has been collected only for special samples, such as surveys of particular industries or cities, although the United Kingdom has sponsored two different series of somewhat similar surveys.

The goal of the project is to understand the incidence of various skill and technology-related job requirements and workplace practices, the interrelationships among these three sets of variables, and their relationship to various measures of job quality (e.g., wages, layoffs, job satisfaction). The use of a refreshed panel will also shed light on rates of change, a longtime subject of speculation. If future waves receive funding, the data will be the first consistent, long-time series on the changing nature of work in the United States.

The project is funded by the National Science Foundation, Russell Sage Foundation, and the University of Wisconsin Graduate School/Wisconsin Alumni Research Foundation.



to this view, poorer skills should be persistent characteristics of the affected age groups, and should be apparent when we compare adults educated before and after the 1960s. Evidence discussed below does not support this view and raises the possibility that much of the literature mistakenly interprets transitory difficulties among youth in their adjustment to the labor market—an age effect—for permanent intercohort declines in skills.

Complaints about social skills and motivation also suggest that what employers perceive as a cohort effect may actually be an age effect. Young adults—and not only in recent years—may pass through a phase characterized by low effort and weak attachment to career employment, attitudes reinforced by a perceived scarcity of jobs offering career opportunities. As workers age and shoulder more adult responsibilities, their casual work attitudes yield to the workplace norms they find attached to the kinds of jobs they consider worth keeping. Indeed, the cohorts whose deficiencies were considered so alarming in *A Nation at Risk* are now in middle age, and are themselves among those who complain about the declining skills of the young. If the problem is the attitudes of young workers, time is the remedy, and there is no progressive problem facing the workforce overall.

Indeed, to the extent that the issue is the increased importance of computer skills and related technological competencies, younger workers should, in theory, be better placed than older workers, whose skills may have become obsolescent and who may have difficulty in retraining for a computerized workplace.

Much of the anxiety about skills mismatch focuses on one sector of the labor market—those with a high school diploma or less, or, more narrowly still, less educated, disadvantaged minorities. If a skills mismatch exists only for these groups, this problem must be distinguished from the idea that schools are failing to teach sufficient skills to the general population of school children. Similarly, the alleged declines in the quality of higher education or in the number of college graduates, especially in science and technology, must be distinguished from poor math and reading skills among high school graduates. In each case the scale of the problem and the potential consequences are very different, but they are often folded into the same rhetoric over inadequate skills.<sup>6</sup>

In sum, existing notions of skills mismatch are a confused jumble. Any satisfactory argument needs to specify whether the problem is a shortage of cognitive skills or a surplus of youthful attitudes, whether it is too many workers with inadequate basic English and math literacy or too few with sophisticated technological expertise, and whether all high school graduates are inadequately prepared for employment or whether the problem is confined to high school dropouts and certain other disadvantaged groups.

### **Skill trends: Declining/stagnating supply or greater demand?**

The trends in skills are as much a subject of debate as the nature of skills mismatch. Is the workforce experiencing an *absolute* decline in skills across cohorts or large groups of workers, as many of those concerned about the school system insist? Or is the issue a *relative* deficit in skills? And if so, is it because growth in the supply of human capital has slowed or demand has accelerated? Labor economists studying inequality are still undecided.<sup>7</sup>

The difference between these questions has clear implications for where we might look for evidence. Absolute declines or slower growth in the supply of human capital suggest problems with the education system and with workers' behavior—failing schools, declining test scores, “underclass” or even mainstream youth values that diverge from mainstream work norms. Accelerating demand for human capital suggests employer-side changes—the spread of computer technology and a workplace that is more participatory and less hierarchically organized.<sup>8</sup>

### **The evidence for a skills decline: Education and test scores**

The most readily available measures of workers' skills are educational attainment, measured by years of schooling, and educational quality, measured by test score trends.

#### **Educational attainment**

In 1964, before the perceived deterioration of public education, 47 percent of all Americans and 31 percent of young adults (ages 24–29) were high school dropouts. In 1997, only about 13 percent of either group had dropped out of high school.<sup>9</sup> Clearly, any view of the pre-1960 period as a golden age of public education is an exercise in nostalgia.

How much educational attainment has risen depends on the measure (is it mean years of education or categories of attainment?), the time range (the rate of growth has varied in different periods), and the population chosen (is it all workers or only young workers?). For the entire population, educational attainment grew most rapidly through the mid-1970s and slowed thereafter through the 1990s. Attainment among young adults rose most rapidly between 1965 and 1975, partly because of rapidly declining high school dropout and rising college attendance (boosted temporarily by Vietnam War draft deferments). Inequality in attainment declined by 25–30 percent for both all workers and young adults between 1962 and 1982; thereafter it did not change.

If educational attainment is the measure of skill, the workforce today is more skilled than ever, although the trend was flat for young workers during 1975–1990, when anxiety about education and the economy both were at their peak.

### Test scores

Schools vary in quality, and educational categories such as “high school graduate” are imprecise measures of skill. Thus test scores have been used as an arguably more precise measure of cognitive skills. Economists and policy analysts whose focus is education are especially inclined to see the trends in public school test scores as evidence that the quality of the workforce is declining. This dissatisfaction underlies much of the current emphasis on high stakes testing and school vouchers as a way to improve the quality of schools and thereby enhance the skills of American workers. In contrast, labor economists concerned about wage inequality have been more skeptical about blaming declining educational quality, partly because older high school graduates who completed their education before 1960 experienced the same relative wage decline as younger workers.

The three measures most relevant for assessing any trend in cognitive skills are intelligence (IQ) tests, college entrance exams (the Scholastic Aptitude Test, SAT, and the American College Test, ACT), and, most complete and representative, the U.S. Department of Education’s National Assessment of Educational Progress, NAEP.

*Intelligence tests* show large gains for Americans in every decade of the twentieth century, and there is no obvious, recent change in the rate of growth.<sup>10</sup> Indeed, the gains in the United States and other industrialized countries are so large that they have generated intense controversy over their meaning. But even if the large gains in IQ scores do not signify commensurate gains in intelligence, they are certainly not evidence of decline.

The decline in *college entrance exam* scores, beginning in the mid-1960s, initiated the recent concern over the state of U.S. public education. It has been less widely reported that math SAT scores started rising around 1980 and by the mid-1990s exceeded 1971 levels, though verbal scores did not recover. The ACT trends were the reverse: English scores have exceeded earlier levels in recent years and math has not fully rebounded. Contrary to the popular impression, then, the SAT test score decline ceased (verbal) or reversed (math) 20 years ago. Other test scores confirm that any downward trend was highly cohort-specific, that is, it was restricted to those educated in the 1960s and 1970s.

However, the usefulness of college entrance exams for tracking cohort-specific skill trends is highly problematic. The population of students who take such tests is not random and has changed so greatly over time that some researchers believe that no credible conclusions can be drawn from these data.<sup>11</sup> Others argue that a genuine decline can be inferred at least for the 1970s. But these scores offer no more definitive evidence of long-term deterioration than IQ tests do.

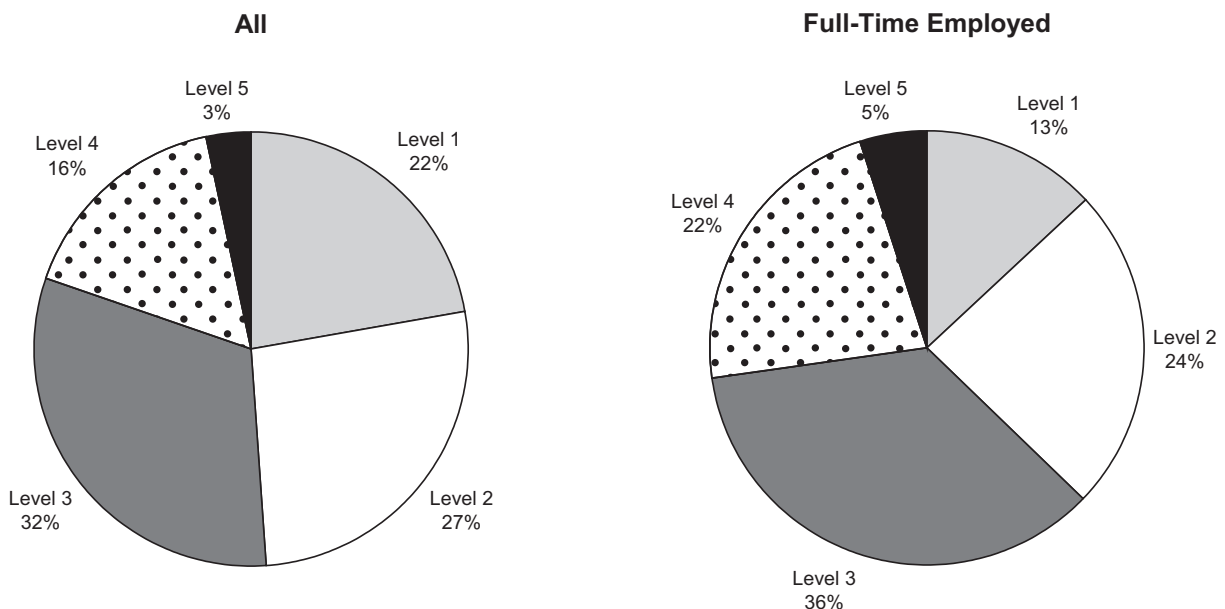


Figure 1. National Adult Literacy Test Scores, 1992.

**Note:** Percentage of entire sample and of full-time employed participants only in each level.

**Scores:** Level 1, 0-225; Level 2, 226-275; Level 3, 276-325; Level 4, 326-375; Level 5, 376-500.

The best data are provided by the *NAEP*, which has a continuous series of reading and math scores for representative samples of 17-year-olds, beginning in the early 1970s; the test itself has remained relatively unchanged over the last 30 years. From 1971 to 1999, reading scores did not change significantly, though their general trend was upward, unlike SAT scores. From 1973 to 1982 math scores fell; then they rose almost continuously and surpassed 1973 levels in 1999. Even critics of American education acknowledge that the changes are small. However, stable achievement scores at a time of rising demand for higher skills may in themselves be a reason for concern.

Because raw NAEP scores mean little to policymakers and the public, the NAEP scale was divided into five performance categories that in principle offer some indication of the tasks that students at each level can perform. The proportion scoring in the highest level is generally small—a statistic that has received great attention. But some research shows that as cohorts of students grow older, the proportion performing at higher levels rises. For example, 21 percent of one sample of young adults who took the NAEP scored at the highest performance level, although only 5 percent of the same cohort had done so at age 17.<sup>12</sup> Evidence of this kind suggests that the NAEP scores of 17-year-olds cannot be taken at face value as measures of the competency of young adults. Other problems in the interpretation of test scores are discussed below.

### Adult reading and math scores

There is little complete or representative information on adult reading and math scores over time. The richest source of data is the cross-sectional National Adult Literacy Survey (NALS) administered in 1992 to a representative national sample of Americans by the Educational Testing Service (which also writes the SAT) for the U.S. Department of Education. In 2003, the National Center for Education Statistics embarked on a second round of this survey, now called the National Assessment of Adult Literacy (NAAL).<sup>13</sup> This round includes new components that are designed to measure the literacy of those adults with the poorest comprehension skills, and an enlarged background questionnaire to provide more demographic and other information about the factors associated with literacy.

The NALS asked respondents to answer questions based on text, forms, and math tasks they would be likely to encounter in their daily lives. How well could the test-takers read newspaper articles or product instructions? Could they figure out documents such as payroll forms and bus timetables? Could they calculate a tip, balance a checkbook, determine the interest rate in a loan advertisement? Figure 1 shows calculated performance levels, on a scale of 1 to 5, for the 1992 sample, and Table 1 the average scores for various groups.

**Table 1**  
National Adult Literacy (NALS) Test Scores, 1992

	Mean
All	270
In Labor Force	283
Employed Full-Time	287
Unemployed	258
Not in Labor Force	241
Employed Full-Time	
<1.25 poverty level	251
Not poor	298
Out of Labor Force	
<1.25 poverty level	213
Not poor	265
Some High School	228
GED	267
High School	268
2-Year College Degree	305
Bachelor's Degree	319
Postgraduate	332
Manager/Professional/Technical	320
Clerical/Sales	291
Craft	267
Operator/Laborer	251
Services	262
Age (high school dates)	
19–24 (1986–1991)	279
25–39 (1971–1985)	283
40–54 (1956–1970)	283
55–64 (1946–1955)	257
65+ (before 1945)	225
White	284
Black	230

**Source:** Kirsch and colleagues, *Adult Literacy in America*, pp. 17, 26, 31, 33) and Sum, *Literacy in the Labor Force*, pp. 24, 32ff., 62, 76ff.

**Note:** All values are simple means of prose, document, and literacy scores. Some occupational means are weighted averages of means for narrower occupational groups. Unless otherwise noted figures refer to all Americans, not simply workers.

The most widely reported—and deplored—result was the large number of Americans, 22 percent, placed in the lowest literacy level, Level 1. But the implications of this finding for the quality of public education or of the labor force are hardly clear. Those categorized in Level 1 were a demographically heterogeneous group. One-third were over 65, a category that includes disproportionate numbers of less educated and retired individuals for whom negative aging effects on cognitive performance are well documented. About a quarter scoring in Level 1 were foreign-born, including many with limited English skills and less schooling from their country of origin than natives. About two-thirds in Level 1 had not finished high school; a third had gone no further than eighth grade. One quarter reported a disability that precluded full-time work.<sup>14</sup>

Only 13 percent of those working full time had Level 1 NALS scores, 9 percentage points below the overall rate for the population. This has been used to suggest that the labor market filters out many low-scoring individuals, but clearly many Americans in the overall population performing at Level 1 are out of the labor market for reasons unrelated to low skills, such as age or physical disability. Insofar as foreign-born workers contribute to the numbers of Level 1 scorers in both the workforce and the overall population, the cause of this situation cannot be readily ascribed to the U.S. education system. Indeed, many in Level 1 who are employed have been drawn into the U.S. labor market from abroad, working for employers who are happy to trade off these workers' lower English literacy skills for the low pay they will accept. Clearly, the reported numbers in Level 1 cannot be used in a straightforward manner to draw conclusions about the number of native-born potential job-seekers who are hard to employ because of low skills.

The NALS results also lend no support to the idea that younger Americans have poorer literacy than older Americans. Indeed, the reverse seems true: among age groups with high probability of labor market participation, younger groups score better than those in the oldest category (aged 55–64). And although the race gap is large, it is smaller for younger adults; the gap for those aged 19–24 in 1992 was about a third smaller than for those aged 40–54 (not shown).

### **What do test scores mean for the skills mismatch debate?**

The significance of test scores is widely debated, and the differing views are reflected in the skills mismatch debate. A sizable group of skeptics believes that the tests underestimate the real-world skills of minorities and low-scoring individuals and that they are biased either in their content or in the way they are used to select employees.<sup>15</sup> Others, especially those in education testing, industrial/organizational psychology, and traditional intelligence psychology, believe that test scores are among the strongest predictors of outcomes, including job performance. Those who score higher, they argue, can perform many jobs faster and more accurately, require shorter training and less on-the-job assistance, and can generalize their knowledge better to unfamiliar circumstances. Yet supporters of the predictive value of test scores do acknowledge that scores account for only a modest amount of the overall variance in job performance and wages, that personality traits and work attitudes are also important predictors of performance, and that the less complex the job, the smaller the association between test scores and job performance.<sup>16</sup>

However, what is particularly problematic for the present discussion is the strong impulse to use test scores to infer

absolute levels of cognitive skills and real-world abilities that might be matched against the skills demanded by employers. All the tests described previously may be reasonably valid for ranking individuals relative to one another, but policymakers and other consumers of these data seek to use them as absolute measures of the tasks people can or cannot perform, a more difficult standard for a test to meet. To use the language of educational psychology, observers mistakenly treat *norm-referenced* tests as if they were *criterion-referenced* tests. There are good reasons to believe that scores on test like the NAEP or NALS are not easily matched with what people can or cannot do outside the test situation.

Because raw scores do not mean a lot by themselves, the five discrete performance categories established for the NAEP are central to drawing connections between test scores and the real world. But the reliability and validity of these categories are dubious. Indeed, evaluators have recommended that they be discontinued, among other reasons because the assignment of test items to performance levels by raters is unreliable and the competency descriptions attached to each level tend to give a lower impression of students' abilities than their scores on other tests warrant.<sup>17</sup> The NALS has not received the same scrutiny, but sample items and their assigned performance levels also suggest that individuals' capabilities would be underestimated if their performance on NALS tasks were interpreted literally as reflecting what they could accomplish in their daily lives. For example, it would not necessarily be expected that people in Level 1 read a newspaper daily. Nevertheless, about 35 percent reported that they did, though this is significantly lower than the average of 50–60 percent in higher levels.

The probability criteria that assign performance levels in the different tests are themselves arbitrary and controversial. The NAEP math test assigns scores based on a 65 percent probability of answering correctly the items classified at that level, but uses an 80 percent probability for reading. NALS uses an 80 percent criterion for all scales. Other tests differ, some setting the criterion as low as 50 percent. There seems to be no strong theoretical or conceptual reason for choosing one standard rather than another, and individuals may be placed in different literacy categories, depending on the test they take and the corresponding probability criterion used to assign scores to performance levels.

Also, it is often forgotten that assignment to a given level does not mean that individuals *cannot* perform at a higher level, merely that they have a lower probability of doing so. For example, someone at the mid-point of Level 2 in NALS, a bit below the average for high school graduates, has about a 30 percent probability of performing tasks at Level 4, which is considered above average for those with at least a bachelor's degree. *Education Week*, reporting the large proportion in NALS Levels 1 and 2, declared, "Nearly half of all adult Americans cannot read, write,



and calculate well enough to function fully in today's society."<sup>18</sup> One of the most widely cited implications of the distribution of NALS scores across performance levels was that nearly 80 percent of adults do not have the skills to calculate a restaurant tip or understand a bus schedule. From these and similar claims, one might conclude that a large proportion of Americans are not only unfit for moderately skilled jobs but might even have trouble finding their way to work, although millions of Americans, disproportionately lower income and likely to have lower test scores, negotiate the public transportation system daily.

These problems with inferring real-world cognitive skills from test performance point to other limitations of such tests. Completing paper-and-pencil exercises in a solitary context is an unrealistic model of how people actually function. In the give-and-take of work, for example, people can ask for clarification or assistance when they find something confusing or unclear; this is not allowed in test situations.

The distinction between academic or test-taking skills and real-world competencies is reflected in an alternative conception of skills known as "practical intelligence" or "situated cognition." Traditional testing and intelligence psychology presuppose that individual tests using prestructured, often abstract tasks in a formal setting generally provide the best general measure of skills. The alternative approach argues that people display greater skills when performing tasks in natural settings, learning from others and through daily experience. Case studies in the workplace and elsewhere support this view.<sup>19</sup> Delivery drivers with near-perfect performance on daily multi-plex tasks at work made many mistakes on paper-and-pencil tests with similar problems. Studies of civilian

jobs consistently find that the correlation between cognitive tests and job performance declines with experience.<sup>20</sup>

In general, then, it appears that the skills workers can develop and for which they are rewarded are partly a function of the jobs employers offer, and that the intrinsic capacities of individuals do not operate as a hard constraint. The NALS analysts acknowledged as much when they cautioned, "These results do not answer the question, 'Are the literacy skills of our nation's workers adequate?'" but they offered a message that was, in the end, mixed: "For an economy that . . . is becoming dependent on high-performance workplaces to spur economic growth, competitiveness, and productivity, many members of the existing labor force appear ill-equipped with respect to key literacy proficiencies."<sup>21</sup>

### U.S. test scores in an international context

Those worried about the skills of U.S. workers argue that the nation's exceptionally low rankings in international comparisons threaten its economic competitiveness. To assess this claim, we can turn to the International Adult Literacy Survey (IALS), modeled on the NALS.<sup>22</sup> From 1994 to 1998, the IALS was administered to representative samples of adults aged between 16 and 65 in 21 countries, though unfortunately not including Japan. As Table 2 shows, the United States is among a lower-scoring group of advanced industrialized nations that includes most other English-speaking countries. The average American is at the 53rd percentile of the pooled sample of all high-income countries—hardly dire, though some have argued that the scores do not match the higher levels of schooling and per capita expenditure in the United States.<sup>23</sup>

#### Access to IRP information via computer: The World Wide Web site

IRP has a World Wide Web site that offers easy access to Institute publications. From the Web site, recent publications are available for immediate viewing and for downloading. Publications available on the Web site include *Focus* articles, recent Discussion Papers and Special Reports in Adobe Acrobat (.pdf) format. Order forms for printed copies and instructions for downloading and printing these files are given on the Web site.

The IRP Web site also provides information about the Institute's staff, research interests, and activities such as working groups, conferences, workshops, and seminars. The Web site also includes an annotated list of affiliates, with their particular areas of expertise. It offers an extensive set of links to poverty-related sites and data elsewhere.

IRP's home page on the Web can be found at: <http://www.ssc.wisc.edu/irp/>



**Table 2**  
**International Adult Literacy (IALS) Test Scores, 1994–1998**

	1 Mean	2 5th pct.	3 95th pct.	4 95/5 ratio	5 GDP/U.S.
1. Ireland	263	151	353	2.34	0.66
2. United Kingdom	267	145	360	2.48	0.70
3. Switzerland	271	143	352	2.46	0.90
<b>4. United States</b>	<b>272</b>	<b>133</b>	<b>371</b>	<b>2.79</b>	
5. New Zealand	272	158	361	2.28	0.62
6. Australia	274	146	359	2.46	0.76
7. Belgium	277	163	359	2.20	0.79
8. Canada	280	145	372	2.57	0.81
9. Germany	285	208	359	1.73	0.76
10. Netherlands	286	202	355	1.76	0.75
11. Finland	288	195	363	1.86	0.72
12. Denmark	289	213	353	1.66	0.89
13. Norway	294	207	363	1.75	0.95
14. Sweden	304	216	386	1.79	0.71

**Source:** Organisation for Economic Cooperation and Development and Statistics Canada, *Literacy in the Information Age: Final Report of the International Adult Literacy Survey* (Paris: OECD, 2000), pp. 135ff., 176; CIA World Factbook.

**Note:** Values in Column 5 are per capita GDP divided by value for United States. Values for Switzerland are averages for three language groups weighted by their share of the population speaking them: German, 70%; French, 21%; Italian, 9%.

Perhaps the most notable feature of these international comparisons is the high U.S. inequality in cognitive skills (Table 2, column 4), especially compared to those countries with the highest mean scores. But when immigrants are excluded from the samples, differences in test score inequality shrink or disappear. In one study of eight countries that excluded immigrants, and that measured inequality as the difference between scores at the 50th and the 10th percentiles, test score inequality in the United States disappeared for women, and shrank by 55 percent for men.<sup>24</sup>

The implications of these results for international economic competitiveness are not obvious. High scorers like the Scandinavian countries and the Netherlands are not usually considered a serious economic threat to the United States. Even when we adjust for the fact that U.S. employees work more hours than employees in the other countries, there seems to be no strong relationship between skill rankings and GDP. Despite the popular assumption that there is a close relationship between the two in wealthy industrialized nations, the links appear, instead, to be weak. U.S. economic performance remains comfortably ahead of most other nations, after a quarter-century of concern over skills deficiencies in the U.S. workforce. ■

<sup>1</sup>This is a summary of Michael J. Handel, “Skills Mismatch in the Labor Market,” *Annual Review of Sociology* 29: 135–65 (2003). A more extended treatment of these issues will be published by the Economic Policy Institute.

<sup>2</sup>Examples of the views described here are D. Bell, *The Coming of the Post-Industrial Society: A Venture in Social Forecasting* (New York: Basic Books, 1976); W. J. Wilson, *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy* (Chicago: University of Chicago Press, 1987); L. Katz and K. Murphy, “Changes in Relative Wages, 1963–1987: Supply and Demand Factors,” *Quarterly Journal of Economics* 107 (1992): 35–78. See U.S. Department of Health, Education, and Welfare, *Work in America* (Cambridge, MA: MIT Press, 1973).

<sup>3</sup>H. Levin, “Schools – Scapegoats or Saviors?” *New Political Economy* 3(1998): 139–43.

<sup>4</sup>U.S. National Commission on Excellence in Education, *A Nation at Risk*, report commissioned for the U.S. Department of Education (Washington, DC: GPO 1983), p.1.

<sup>5</sup>E. Goode and N. Ben-Yehuda, “Moral Panics: Culture, Politics, and Social Construction,” *Annual Review of Sociology* 20 (1994): 149–71.

<sup>6</sup>Examples of these and similar discussions are National Research Council, *Building a Workforce for the Information Economy* (Washington, DC: National Academy Press, 2001); D. Card and T. Lemieux, “Can Falling Supply Explain the Rising Return to College for Young Men? A Cohort-Based Analysis,” *Quarterly Journal of Economics* 116 (2001): 705–46; R. Murnane and F. Levy, *Teaching the New Basic Skills: Principles for Educating Children to Thrive in a Changing Economy* (New York: Free Press, 1996).

<sup>7</sup>See, e.g., Card and Lemieux, "Falling Supply"; D. Autor, L. Katz, and A. Krueger, "Computing Inequality: Have Computers Changed the Labor Market?" *Quarterly Journal of Economics* 113 (1998): 1169–1213.

<sup>8</sup>For discussions of employer demands see Murnane and Levy, *Teaching the New Basic Skills*; Autor and colleagues, "Computing Inequality"; Harry Holzer, *What Employers Want: Job Prospects for Less-Educated Workers* (New York: Russell Sage, 1996).

<sup>9</sup>These are my calculations, based on the March Current Population Survey.

<sup>10</sup>J. Flynn, "IQ Gains over Time: Toward Finding the Causes," in *The Rising Curve: Long-Term Gains in IQ and Related Measures*, ed. U. Neisser (Washington, DC: American Psychological Association, 1998).

<sup>11</sup>R. Hauser, "Trends in Black-White Test Score Differentials: Uses and Misuses of NAEP/SAT Data," in Neisser, ed., *The Rising Curve*.

<sup>12</sup>I. Kirsch and A. Jungblut, *Literacy: Profiles of America's Young Adults* (Princeton: Educational Testing Service, 1986); J. Campbell, C. Hombro, and J. Mazzeo, *Trends in Academic Progress: Three Decades of Student Performance*, National Center for Education Statistics, U.S. Department of Education, Washington, DC, 2000.

<sup>13</sup>For an overview of the 2003 NAAL, see the World Wide Web site of the NCES, <<http://www.nces.ed.gov/naal/design/about02.asp>>.

<sup>14</sup>I. Kirsch, A. Jungblut, L. Jenkins, and A. Kolstad, *Adult Literacy in America: A First Look at the Results of the National Adult Literacy Survey*, National Center for Education Statistics, U.S. Department of Education, Washington, DC, 1993.

<sup>15</sup>J. Hartigan and A. Wigdor, eds., *Fairness in Employment Testing: Validity Generalization, Minority Issues, and the General Aptitude Test Battery* (Washington, DC: National Academy Press, 1989); C. Jencks, "Racial Bias in Testing," in *The Black-White Test Score Gap*, ed. C. Jencks and M. Phillips (Washington, DC: Brookings Institution, 1998). This is, of course, in addition to the simmering debate over the value of test scores as a high school graduation requirement (so-called "high-stakes testing").

<sup>16</sup>L. Gottfredson, "Why g Matters: The Complexity of Everyday Life," *Intelligence* 24 (1997): 79–132.

<sup>17</sup>See, e.g., J. Pellegrino, L. Jones, and K. Mitchell, *Grading the Nation's Report Card: Evaluating NAEP and Transforming the Assessment of Educational Progress* (Washington, DC: National Academy Press, 1999).

<sup>18</sup>*Education Week*, quoted in R. Stiles, "'How Much Literacy Is Enough?' Issues in Defining and Reporting Performance Standards for the National Assessment of Adult Literacy," Working paper No. 2000-07, National Center for Education Statistics, U.S. Department of Education, Washington DC, 2000.

<sup>19</sup>C. Stasz, "Assessing Skills for Work: Two Perspectives," *Oxford Economic Papers* 53 (2001):385–405.

<sup>20</sup>See, e.g., S. Scribner, "Thinking in Action: Some Characteristics of Practical Thought," in *Practical Intelligence: Nature and Origins of Competence in the Everyday World*, ed. R. Sternberg and R. Wagner (London: Cambridge University Press 1986), pp.13–30; E. Hunt, *Will We Be Smart Enough: A Cognitive Analysis of the Coming Workforce* (New York: Russell Sage, 1995).

<sup>21</sup>A. Sum, *Literacy in the Labor Force*, National Center for Education Statistics, U.S. Department of Education, Washington DC, 1999 (quotations on pp. xvi and 33, respectively); The designers of the NAAL were warned against "the temptation to directly address this question" (Stiles, "How Much Literacy Is Enough?").

---

FOCUS is a Newsletter put out up to three times a year by the

Institute for Research on Poverty  
1180 Observatory Drive  
3412 Social Science Building  
University of Wisconsin  
Madison, Wisconsin 53706  
(608) 262-6358  
Fax (608) 265-3119

The Institute is a nonprofit, nonpartisan, university-based research center. As such it takes no stand on public policy issues. Any opinions expressed in its publications are those of the authors and not of the Institute.

The purpose of *Focus* is to provide coverage of poverty-related research, events, and issues, and to acquaint a large audience with the work of the Institute by means of short essays on selected pieces of research. Full texts of Discussion Papers and Special Reports are available on the IRP Web site.

*Focus* is free of charge, although contributions to the U.W. Foundation–IRP Fund sent to the above address in support of *Focus* are encouraged.

Edited by Jan Blakeslee.

Copyright © 2004 by the Regents of the University of Wisconsin System on behalf of the Institute for Research on Poverty. All rights reserved.

---

<sup>22</sup>The IALS was sponsored by the Organisation for Economic Cooperation and Development, Statistics Canada, the ETS, and other government agencies and national study teams.

<sup>23</sup>A. Sum, I. Kirsch, and R. Taggart, *The Twin Challenges of Mediocrity and Inequality: Literacy in the U.S. from an International Perspective* (Princeton, NJ: Educational Testing Service, 2002).

<sup>24</sup>F. Blau and L. Kahn, "Do Cognitive Test Scores Explain Higher U.S. Wage Inequality?" Working Paper 8210, National Bureau of Economic Research, Cambridge, MA, 2001.

# The children of New Hope, five years later

The New Hope Project, designed and implemented in Milwaukee, Wisconsin, was founded on two basic principles: that people who are willing to work full time should have the opportunity to do so, and that people who work full time should not be poor.<sup>1</sup> Conceived by a nonprofit, community-based organization and funded by a consortium of foundations and by the state and federal governments, New Hope was intended as a demonstration of work supports that could be replicable as government policy.

## The New Hope offer to participants

New Hope recruited participants in two inner-city areas of Milwaukee in 1994–95. There were only four eligibility requirements: that applicants live in one of these areas, be age 18 or over, have a household income at or below 150 percent of the federal poverty level, and be willing and able to work at least 30 hours a week.

To support work, New Hope offered a varied menu of benefits and services:

- *Job access.* Participants who were unemployed or wanted to change jobs received individual job search assistance. If they failed to find work after an eight-week search, they could apply for a community service job (CSJ) in a nonprofit organization. CSJs paid the minimum wage and might be either full or part time.
- *Earnings supplements* were offered monthly to qualifying participants (those working at least 30 hours a week) whose earnings left the family below 200 percent of the poverty level. CSJ earnings were counted toward the 30-hour work requirement, and also qualified participants for the federal and Wisconsin Earned Income Tax Credits.
- *A health insurance plan* was offered to qualifying participants who were not covered by an employer plan or Medicaid. Participants contributed to the cost on a sliding scale that took into account their income and household size.
- *Child care assistance* for children under 13 was offered to qualifying participants. Again, parents paid part of the cost, depending on their income and family size. For participants to qualify for the subsidy, their children's care had to be provided in state-licensed or county-certified homes or centers.
- *Staff support.* All participants were assigned to project representatives who could provide advice and information about employment, child care, and other

topics. The New Hope model emphasized respect and helpfulness in staff interactions with participants.

Qualifying New Hope participants could use any number or combination of program benefits or services they chose, or needed. Eligibility for earnings supplements, health insurance, and child care assistance extended for three years from the date a participant entered the program.

New Hope was designed to provide information to policymakers on a number of issues. Would the program succeed in boosting employment, increasing economic security and family well-being, and lowering the use of public assistance? Would it affect the lives and long-term development of children? To answer these and other questions, the effects of the program on parents, children, and families were intensively evaluated two and five years after New Hope was initiated. The evaluation was carried out by an independent organization, MDRC, using a sophisticated random assignment methodology to select a study sample. For the purposes of the evaluation, applicants were randomly assigned to be in the program group, which was eligible to receive New Hope services, or the control group, which was not. Both groups could receive all other community programs.<sup>2</sup>

This article discusses the effects of the program on children at the five-year point—that is, two years after participants' eligibility for the program had ended. The findings are drawn from a new report (see box) about effects on a subgroup of the sample, the Child and Family Study (CFS), that was selected to evaluate the program's effects on children and families. The CFS sample included all 745 adult New Hope participants who had one or more children between the ages of 1 and 11 at the time participants were randomly assigned either to the program or to the control group (these constituted 55 percent of the entire sample). If the family had more than one child in that age range, two children were identified as "focal children." Thus the evaluation began with 1,140 children.

This article summarizes findings from *New Hope for Families and Children: Five-Year Results of a Program to Reduce Poverty and Reform Welfare* (New York: MDRC, 2003). Authors of this MDRC report are Aletha C. Huston, Cynthia Miller, Lashawn Richburg-Hayes, Greg Duncan, Carolyn A. Eldred, Thomas S. Weisner, Edward Lowe, Vonnie C. McLoyd, Danielle A. Crosby, Marika N. Ripke, and Cindy Redcross.

The CFS evaluation drew information from state employment, public assistance, and tax records, and from New Hope's own administrative records. Parents and children were interviewed two and five years after random assignment about their receipt of New Hope services, their economic circumstances, parent-child relations, children's participation in after-school activities, and children's social and academic development. Mail surveys asked teachers of the school-age children about their school performance and social behavior. At the five-year mark, 561 parents responded to the survey; 840 children from these families, aged between 6 and 16, were interviewed. Surveys were also returned by 547 teachers to whom parents gave permission to respond. A three-year ethnographic study (1998–2000) of 44 families from the CFS sample (both program participants and controls) provided context and deeper understanding of data from administrative sources and surveys.

At the beginning of the study period, the majority of parents in the CFS sample had never been married; just over 10 percent were married and living with a spouse. Slightly over half of parents were African American, and over one-quarter were Hispanic. Almost half had three or more children, and over three-quarters of them had children under the age of 5. Furthermore, many had responsibility for the children of a partner or of other family members. Around 40 percent had neither completed high school nor received the GED. As a group they faced a number of other barriers to employment: only 44 percent had access to a car, and nearly 20 percent had been arrested at some point since their 16th birthday. When they applied for New Hope, over half were not employed and about 80 percent were receiving Aid to Families with Dependent Children (AFDC), general assistance, food stamps, and/or Medicaid.

### **Why New Hope might have lasting effects**

New Hope's designers originally conceived of the program as a set of work supports that would be in place as long as individuals needed them; the three-year eligibility limit it imposed was a function of financial constraints. The evaluation discussed here, two years after participation ended, was particularly concerned with whether three years of benefits would have lasting effects on parents, children, and family life.

There are reasons to expect that they might. First, if persons gained job experience and confidence in their ability to earn a living, the employment and income gains of New Hope might continue, especially because the EITC continued to be available as an important earnings supplement. (Evidence from the ethnographic survey suggests that families chose lump-sum EITC payments to provide a form of savings, for big-ticket purchases such as cars or appliances, and to pay down debt.)

Second, children's experiences in formal child care and structured out-of-school activities might occur during "sensitive developmental periods," in which experiences have formative and enduring effects. For example, if formal child care provides children with basic academic preparation, they may begin school already on a trajectory to success. Teachers may perceive them as more skilled and may provide more opportunities. Organized after-school activities may contribute to academic and social skills; children taking part in such activities under New Hope may later continue them. Thus advantages accrued during New Hope may lead to an upward developmental spiral, and program-induced changes in the child's behavior may elicit particular reactions from those around or lead the child to seek out different contexts.

Finally, the changed contexts brought about by New Hope—effects on employment, income, and family routines—may endure. Parents may have acquired greater skills in the workplace, in negotiating bureaucracies, and in finding community resources for themselves and their children.

This said, the rapid changes in federal, state, and local policies affecting poor parents from 1995 to 2000 unquestionably "raised the bar" for showing the effects of New Hope. These changes diminished the difference between what New Hope offered and what was available outside, for instance, through enhanced EITC, child care subsidies, health insurance for low-income children, and the state's Wisconsin Works (W-2) program.<sup>3</sup>

### **What New Hope benefits did parents actually use?**

New Hope designers expected that many participants would not need all the program's benefits in every month. The health insurance, for example, would be of little interest to participants who were already covered by Medicaid. Parents with long-standing child care arrangements might not wish to disrupt them to qualify for the subsidy.

The vast majority (almost 88 percent) of CFS program group members received at least one of the benefits during their three-year eligibility period. Almost all of these received at least one earnings supplement. Slightly more than half made use of the health insurance and child care subsidies. Benefit usage remained relatively stable over the last 18 months, though it rose slightly in the third year, particularly for health insurance. There is some evidence that this increase is due to the larger numbers of families leaving welfare and so needing additional supports.

On average, families received the earnings supplement and child care subsidies for about 15 months, and health



insurance for just over a year. Those who had lower incomes and larger families received the most substantial supplements; some 15 percent received an average supplement of over \$200 a month. The average monthly payment for health insurance on behalf of those using that benefit was \$278 a month, and the average child care subsidy paid by New Hope was around \$700 a month. At the end of eligibility, 45 percent of the sample were still receiving some type of New Hope benefit. New Hope's staff had worked hard to prepare people for the impending loss of benefits and to ensure smooth transitions out of the program. Nonetheless, about 20 percent of participants reported that the end of eligibility was a serious problem for them.

The gross costs of providing these services were approximately \$5,270 per CFS program family per year. Child care subsidies accounted for 38 percent of these costs, and case management, benefit administration, and the development and management of the CSJs accounted for another 23 percent. Health benefits absorbed 11.7 percent, and the remainder was about equally divided among the earnings supplement, CSJ wages, and program administration.

### **Setting a benchmark: The effects of New Hope two years after entry**

In exploring whether the effects of New Hope persisted after participation ended, we start with the changes observed while families were still participating. At the two-year point, when focal children were between 3 and 12 years old, parents in the New Hope program, especially those not employed full time when they entered, had significantly higher rates of employment and higher earnings than did those in the control group. There were no noticeable changes in measures of mental health (e.g., depression), but parents did report less stress, fewer financial worries, and a greater sense that they could take action and achieve their goals. They also reported greater time pressures.

At this point, New Hope had strong effects on children's experiences outside their homes but little measurable effect on the home environment or on parent-child relations. Most important, the child care subsidies encouraged greater use of formal center-based care and after-school care among preschool and elementary school children. Children aged 9 to 12 also participated more in structured lessons, sports, clubs, and religious groups. In the program group families, boys especially were making better academic progress and displaying more positive social behavior than children in control group families. The picture for girls was more mixed, less generally positive.

Teachers' higher ratings of boys' classroom behavior and work habits, and the boys' own higher expectations about

their future educational and occupational attainment, are particularly noteworthy. The ethnographic interviews suggest that parents were well aware of the risks of delinquency and school failure confronting boys and may have invested greater resources in ensuring that boys had alternative activities to hanging out with unsupervised peers after school. There is no evidence that girls assumed greater household responsibilities than boys while their mothers worked, but girls may have responded differently to the role models they observed as their mothers entered the world of low-wage work (90 percent of parents in the sample were women).

In sum, there appears to be no single, predominant pathway through which the program supported families. In effect, New Hope offered a "cafeteria" of supports, which participants could tailor to their own needs. Different families might use the program in very different ways. The ethnographic evidence suggests that most commonly, New Hope was helpful when it offered benefits that fitted into the family's already functional daily routine. For example, child care vouchers helped parents who wanted to provide better-quality care or to disentangle themselves from social networks that were not providing care reliably or well. Some parents found New Hope case representatives to be valuable allies in finding jobs and services. Others simply found New Hope a support in dealing with the cascading problems that most working-poor families faced. Income supplements, a CSJ to bridge over a patch of unemployment, or some combination of New Hope and state benefits could increase family stability and improve daily lives.

### **Benefit use after the end of New Hope eligibility**

Five years out, any significant differences in benefit use between participating and nonparticipating ("control group") families in the experimental sample had disappeared. Receipt of cash welfare (AFDC or W-2) declined dramatically over the study period for both program and control families, perhaps because W-2 made it more difficult and less desirable for both groups to remain on the welfare rolls. Nor did New Hope have significant effects on the dollar amount of welfare or food stamp receipt. For example, in the first year, the program group received average annual welfare benefits of \$3,496, and the control group received \$3,583. By Year 5, these amounts had fallen to \$476 and \$466, respectively. Around the same percentage of both groups had health insurance for adults (controls, 88 percent; program families, 86 percent), from roughly the same sources: just over 30 percent had coverage through their employer, 13–14 percent had another private or family employer plan, and 44–48 percent were covered by Medicaid or the state's BadgerCare plan. About 16 percent of both groups were receiving a child care subsidy from the welfare agency or some other organization.

**Table 1**  
**Summary of New Hope's Effects at the Five-Year Mark**

Outcome for Program Group vs. Control Group

**Parents' employment and income**

- Modestly higher income
- Less poverty
- More stable employment
- Higher wages

**Parents' well-being**

- No difference in material or financial well-being
- Slightly better physical health
- Fewer depressive symptoms
- Better awareness of public and community resources
- Better able to sustain daily routine

**Parenting**

- Few overall effects
- Fewer problems with control in discipline situations
- Boys:* More positive parent relations
- Adolescents:* More effective child management

**Child care**

- More center-based care
- More after-school programs
- Less home-based care
- Less unsupervised care
- Fewer changes in arrangements

**Children's out-of-school activities**

- More participation in religious activities and organizations
- Adolescents:* More participation in structured activities (for example, sports, lessons, community centers)

**Children's academic achievement**

- Better scores on standardized reading achievement test
- Better reading performance (as reported by parents)
- Boys:* Better academic skills (as reported by teachers)

**Children's motivation and well-being**

- No overall impacts
- Boys:* Higher educational expectations
- Greater school engagement
- Adolescents:*
- Higher educational expectations
- Greater school engagement
- Increased feelings of efficacy to reach goals
- Greater future community involvement

**Children's social behavior**

- More positive social behavior (as reported by parents)
- No difference in risky, delinquent behavior
- Boys:* More positive social behavior (as reported by teachers)
- More appropriate classroom behavior (as reported by teachers)
- Less hostility in provocation situations
- Girls:* Less positive social behavior (as reported by teachers)
- More problem behavior (as reported by teachers)

**Children's health**

- No impacts

**Parenting, child care, and children's activities at the five-year mark**

**Parenting**

Measures of parenting included parents' own reports, children's reports, and ratings by interviewers. These were grouped into composite scores: effective child management, positive and negative youth-parent relations, and warm and structured parenting. In light of the meager effects of New Hope on parenting behavior while families were still participating, the program was not expected to have robust effects at the five-year point, and these low expectations were fulfilled. There was, indeed, a significant effect on only one of the many measures included—program group parents reported fewer problems controlling their children than did control group parents. These findings were not much affected by the parent's employment or race, but were strikingly linked to the child's age. New Hope parents reported that they could more effectively manage their 13-to-16-year-old children; they had higher levels of control and greater confidence in their ability to keep their children from harm, and less frequently resorted to punishment or discipline.

At the two-year point, program group boys, though not girls, had reported more positive relations with their parents than control group boys; these effects, though modest, held after five years. These longer-term program effects may reflect, in part, parents' responses to the improvements in children's behavior and school performance that were a likely consequence of the increased time in structured before/after-school programs made possible by New Hope subsidies. Parents working full time at baseline appear also to have exhibited warmer and more effective parenting approaches as a result of the New Hope experience.

**Child care**

The effects of New Hope on the type of child care experienced by children in participating families were large and consistent. Children from program families spent more time in center-based care and less time in home-based care, both during the school year and in the summer (Table 2). Over the whole year, the effect of the program can be measured as an additional month of formal care, one month less of home-based care, and about two-thirds of a month less of unsupervised care. Despite the differences in the type of care that program and control children received, both sets of families spent about the same out-of-pocket amounts for child care and received about the same amounts of public child care assistance. Between ages 9 and 12, when most children discontinue formal child care, children in the program spent significantly less time in unsupervised care or in taking care of younger children. The use of formal care increased for children from African American, Hispanic, and white families, but there were some differences: use of home-

Table 1 summarizes differences between the program and control group families five years after they entered New Hope—two years after their eligibility for the program ended.

**Table 2**  
**Effect on Child Care Use and Costs, by Child's Age**

Child Care Used during the Prior Year	Program Group	Control Group	Difference
<b>Aged 6–8</b>			
Any formal care (months)	5.1	3.6	1.4**
Any home-based care (months)	5.0	6.3	-1.3*
Any unsupervised care (months)	1.7	1.6	0.2
Out-of-pocket care costs in month before survey (\$)	63.8	69.4	-5.5
<b>Aged 9–12</b>			
Any formal care (months)	3.7	2.6	1.1*
Any home-based care (months)	5.3	6.6	-1.2*
Any unsupervised care (months)	2.2	3.8	-1.7***
Out-of-pocket care costs in month before survey (\$)	45.6	24.3	21.3*
<b>Aged 13–16</b>			
Any formal care (months)	1.9	1.5	0.4
Any home-based care (months)	4.5	5.1	-0.5
Any unsupervised care (months)	4.2	4.3	-0.1
Out-of-pocket care costs in month before survey (\$)	13.4	18.1	-4.7

**Source:** Five-year report, Table 5.5, pp. 111–113.

**Note:** Difference between program and control group scores significant \* at the 10% level \*\* at the 5% level, \*\*\* at the 1% level.

based and unsupervised care diminished significantly only among white families, whereas Hispanic children in the program spent slightly more time in unsupervised settings.

Why would families continue to use more formal types of care two years after the end of the program (and the generous subsidy), even though their earnings and income were not markedly greater? The qualitative data suggest, first, that low-income parents liked the stability and predictability of formal care when they were working. Indeed, there is a feedback loop. Stable employment makes it possible to sustain formal care, and reliable child care may contribute to the ability to maintain stable employment. Some parents clearly thought that formal child care contributed to children's academic skills, and actively sought ways to maintain it, with the help of the New Hope program representatives. These may have helped them access other community sources of child care assistance; two years after eligibility for New Hope ended, program group parents were no more likely than control group parents to be receiving public child care assistance, nor were they paying significantly more out of pocket for the higher amounts of formal care their children received.

### Children's activities

Parents and children aged 9–16 provided reports about their participation in lessons, organized sports, clubs and youth groups, and religious classes and events, and about their attendance at recreation or community centers—all structured activities that afforded opportunities for adult supervision, acquiring skills, and socializing with peers. Families were asked about before- and after-school pro-

grams, day camps, and summer school, about volunteer or service activities, about social activities (shopping or eating out) with adults and peers, and about television viewing habits.

For children in late childhood and adolescence, structured and organized activities can promote positive psychological, intellectual, and social development, and provide protection from risk—advantages that may be especially pertinent for low-income children. Continuing a pattern observed three years earlier, adolescent children in the program group participated more frequently in organized activities than did control group children. But the primary difference between program and control children lay in the significantly greater amount of time program children of all ages spent in classes or activities sponsored by religious institutions, both during the school year and in the summer. They also spent more time in service or volunteer activities. Otherwise, the effects of the program on children's structured and social activities appear to be small and inconsistent; there are few differences between the two groups.

### Children's educational performance, motivation, and expectations

The school achievement of New Hope children was assessed in three different ways—through standardized achievement test scores that measure reading and mathematical skills, through parents' ratings, and through teachers' ratings that assessed children's current performance and behavior and teachers' expectations for the children's educational attainment.<sup>4</sup>

By all three measures, many of the positive effects evident after two years persisted at the five-year point. Across all age groups, children from program group families were performing better academically. Test scores and parents' reports suggest that the effects were more pronounced for reading than for math. For example, on a standardized measure of reading achievement, program group children had an average reading score of 98.1, compared to 96.0 for the control group. The average child in the United States population attains a score of 100. New Hope children scored higher than about 45 percent of the U.S. population, and the controls scored higher than about 39 percent of the population. Adolescents in the New Hope families were less likely to be retained in grade, to be receiving remedial services, or to be receiving poor grades. These children, who were in elementary school when their parents entered the program, were more engaged with school, and had higher hopes about their futures and higher educational expectations.

At the five-year evaluation, as earlier, there were insignificant differences by ethnicity or age, but large differences by gender. Teachers rated program group boys significantly higher than control group boys on both academic performance and classroom behavior. The large advantages of program boys on measures of positive social behavior (self-control, autonomy, sensitivity) had faded by the fifth year, though mostly because control group boys had improved on these measures.

The effects for girls remained weaker and less consistent at this point. As before, teachers rated program group girls lower than control group girls on some areas of achievement and classroom behavior. These less favorable ratings carried over into the girls' own motivation and engagement in school—program group girls were less engaged and had lower expectations of graduating from college than did girls in the control group.

The random assignment of families to program and control groups makes it unlikely that there were systematic differences between these two groups at the beginning. Moreover, all analyses took into account the characteristics of families when they entered the program. The large gender differences in experimental impacts appeared mostly in teachers' reports; for the most part, the program had similar impacts on achievement when measured by parents' assessments and test scores for boys and girls. Teachers were given no information about children's participation in New Hope, so the program-control group differences they observed are unlikely to have been affected by knowledge of the intervention. The few other studies that have included teachers' ratings, such as the New Chance Study, have shown a similar divergence—positive effects for boys in the program, negative effects for girls.<sup>5</sup>

The gender differences in effects should be considered in light of boys' higher risk of school failure. In the control

group, boys had lower scores than girls on academic and social behavior. In effect, New Hope raised teachers' ratings of program group boys to be approximately equal to those of girls in the control group families. Other research suggests that teachers generally rate girls more favorably than boys.<sup>6</sup>

## How important are the effects of New Hope on children?

Are the effects of New Hope on children socially and economically significant? After all, as Table 1 notes, the effects of the program for the parents were quite modest. But these modest direct effects appear to have produced real and lasting benefits for children. This is the more impressive in that all effects of New Hope on children were, by the nature of the program, indirect, mediated through effects on the parents. The New Hope Program provided no intensive early childhood interventions, and all decisions about child care, after-school care, and other activities were made by the parents. And the effects of New Hope on children are consistently present in measures obtained from multiple sources—teachers, parents, and the children themselves. Given the time that had elapsed between the end of the program and the measured outcomes, the differences summarized in Table 1 are large.

How do the New Hope effects compare with the effects of intensive early interventions designed specifically to improve the school performance of low-income children? The Abecedarian program, for example, provided full-day, high-quality child care from infancy until school entry.<sup>7</sup> At age 12, the children in the program scored 5 to 7 points higher than control group children on the Woodcock-Johnson scales of reading and math achievement. The children in New Hope scored about 1.5 to 3 points higher than control group children. Thus New Hope produced about a third of the gain that occurred in this expensive, long-run, and high-quality program.

In short, program effects are real, and they are large enough to be socially significant. The annual costs of around \$5,300 per family are not trivial. But nor are the benefits of the program. ■

---

<sup>1</sup>See "The New Hope Project," *Focus* 18, no. 1 (Special Issue 1996): 82–85.

<sup>2</sup>"The New Hope Project: Two-Year Results of the MDRC Evaluation," *Focus* 20, no. 2 (Spring 1999): 49.

<sup>3</sup>For a comparison, see T. Kaplan and I. Rothe, "New Hope and W-2: Common Challenges, Different Responses," *Focus* 20, no. 2 (Spring 1999): 44–50.

<sup>4</sup>The achievement tests used were drawn from the Woodcock-Johnson Achievement Battery; teachers used the academic subscale of the Social Skills Rating System (1 = child is in lowest 10 percent of the class, 5 = child is in highest 10 percent of the class) and the Classroom



Behavior Scale, and also produced a mock “report card” on current school performance adapted from the Study of Early Child Care and Youth Development by the National Institute of Child Health and Human Development. For summaries of these test results, see Huston and colleagues, *New Hope for Families and Children*, Table 6.2.

<sup>5</sup>J. Quint, J. Bos, and D. Polit, *New Chance: Final Report on a Comprehensive Program for Young Mothers in Poverty and Their Children* (New York: MDRC, 1997).

<sup>6</sup>On the most widely used scale of behavior problems, the Child Behavior Checklist, the published norms show that teachers’ average score for boys is higher than it is for girls. See T. Achenbach, *Manual for the Child Behavior Checklist 4-18 and 1991 Profile* (Burlington, VT: Child Behavior Checklist, 1991).

<sup>7</sup>A number of these intensive early interventions are discussed in a special issue of *Focus* 19, no. 1 (1997) on “Investing in Young Children.” The Abecedarian study is discussed in C. Ramey, F. Campbell, M. Burchinal, M. Skinner, D. Gardner, and S. Ramey, “Persistent Effects of Early Childhood Education on High-Risk Children and Their Mothers,” *Applied Developmental Science* 4 (2000): 2-14. This study shows four groups compared at age 12 and age 15. The New Hope study compares the Woodcock-Johnson reading scores for the two programs because they were the significant ones in New Hope. Abecedarian had four groups: control, intervention at preschool, intervention in school years, and intervention in both school and preschool. The preschool intervention was most important. At age 12, the Woodcock-Johnson score for the control group was 84; the two groups who received preschool intervention were 89 and 91 (hence the 5-to-7-point difference). At age 15, the control group was 88, and the two preschool intervention groups were 92 and 95.

# Respect, support, and accountability: Lessons in delivering the New Hope Project offer

By Julie Kerksick

---

Julie Kerksick is the Executive Director of the New Hope Project, in Milwaukee, Wisconsin. New Hope's mission is to create pathways to help individuals who can get out of poverty through work. New Hope pursues this mission in a variety of ways: through policy and advocacy work, through technical assistance, and, most recently, through running a transitional jobs and employment program as part of the Welfare to Work program.

---

The New Hope Project was not created as a welfare reform program, though its potential relevance to welfare reform is how and why most people became interested in us. New Hope was born out of people organizing to create better employment policies for all low-income workers and would-be workers. This bears saying in this context, because we came at our work with significant understanding of and experience with entry-level labor markets. We had spent ten years trying to connect individuals to jobs that would get them out of poverty.

We knew that most of the people who came through our doors could not immediately qualify for "good jobs," though we always tried to help them find the best possible job that they could get. We learned from our successes and even more from our failures that there were serious structural barriers in many of the jobs available to our constituents—jobs that would never lift a person beyond poverty, and not many automatic escalators from those jobs to better jobs. That was the basis for developing the New Hope offer, with its four components: access to work, including limited-term, guaranteed jobs for up to six months; earnings supplements; affordable health insurance; and affordable childcare.

When the evaluators from MDRC discussed their findings with the New Hope Board of Directors and staff back in 1999, they kept referring to a fifth component of the New Hope offer: effective case management. At first, I dismissed this "finding," because I took for granted that an antipoverty employment program would treat participants with respect, deal with them on an individual basis, and follow up even after they had begun working. I didn't like the evaluators making such a big deal out of how "nice" our staff were to the participants, as if that was the revolutionary characteristic of our work. I did not want the need for economic support to be pushed aside or conveniently ignored by policymakers thinking that you don't have to help people get more income, you just have to be warmly encouraging.

At the same time, as I studied the New Hope results—positive and inspiring in some ways, but more modest than I would have hoped in others—I began to understand that how we worked with people was as important as what we and they did together, or the economic supports we offered. This realization was strengthened over the next several years, as I looked at employment and welfare reform programs in Wisconsin, other parts of the United States, and the United Kingdom. I saw that programs needed to address both economic and noneconomic needs of individuals who do not have consistent or stable employment, and that I was taking too much for granted in assuming that program managers understood how to position and train staff to support participants effectively.

The New Hope Project staff understood that an individual's commitment to go to work was necessary, but not always sufficient, to guarantee success in the labor market. It was why New Hope was designed to make work pay. But we also knew that financial incentives would not make a difference if the individual was not ready to take ownership of the process and fact of being a worker, even if he or she hadn't necessarily come to grips with all that required. The motivation might have sprung from external pressures, such as changes in the rules governing eligibility, or it might come from internal motivators, such as a desire to improve the family's standard of living, or feel more in control of one's fate. It was crucial that staff begin by understanding what kinds of things were motivating the individual with whom they were working.

To the extent that New Hope was successful in helping participants make progress toward economic security through work, we did so by having both economic and noneconomic supports and staff who used all these supports to help participants take more control of their goals and undertake the tasks necessary to achieve them. To the extent that we were not successful with participants, some—though by no means all—of the explanation lies in our internal failures to adequately figure out how to work with people so that they would take greater ownership of their goals and of essential tasks.

I have learned a lot from the experience of running the New Hope demonstration and from studying other employment or welfare reform programs. I have seen high-quality programs and low-quality programs, and many in-between. I have thought a great deal about how to codify these lessons, to make them more transparent to myself and other managers. But a decade later, I am still strug-

gling to articulate what makes programs and frontline staff effective. Writing this article is part of helping me do this.

Here's what I think characterizes good case management:

## Respect

Good programs send messages to staff and participants at all levels that indicate respect. Respect for what? The list is long, but for starters, it includes the following:

- for the basic human dignity of both participants and staff;
- for the participant's own sense of what is important;
- for the participant's desire to do the right thing for himself or herself and the family;
- for the participant's time;
- for the participant's willingness to do unpleasant or hard things if they are seen as having the potential to lead to greater well-being;
- for the participant's experience;
- for the participant's ability to change;
- for the staff's experience and information;
- for the staff's time;
- for the staff's desire to do a good job;
- for the staff's ability to continually improve.

---

## Respect

We made appointments, but individuals might simply drop in, creating challenges for Project Representatives who were juggling already full calendars. But often participants needed something simple, or just wanted to drop off documents. We would begin by checking if the participant's Project Representative could see him or her, within a 15-minute window. If not, then we asked if someone other than the Project Representative could help. If the answer was yes, we looked for either the receptionist or another staff member to deal with the request. If none of this worked, then we offered to make appointments for the participants with their Project Representative. It didn't always go smoothly, but these steps generally set a responsive tone, while balancing respect for the value of both participants' and staff members' time.

We had a policy of returning all calls within 24 hours. When we were having a lot of trouble administering our child care subsidies, we identified an ombudsman who also had a response time of 24 hours or less. Even if she couldn't resolve all the issues within that time, the participant and her caregiver would have information about what was under way and what the expected timeframe was for resolution.

---

## Support

Support can mean economic and noneconomic support. Participants have come to the program because they need something. Their needs and wants may change over time, so it's important to begin establishing a relationship by asking them about the kinds of help that they want. (It's amazing how many programs do not take this into account. The programs assume they know what is needed, and jump right into whatever method they use for providing it, whether it be persuasion, command, exhortation, lecture, etc.) I think that Toby Herr's work and writings with the Erikson Institute (<http://www.erikson.edu/home.asp>) and Project Match (<http://www.pmatch.org/>) have done a great deal to codify these concepts.

Support takes many forms:

- help in defining individual goals, short- and long-term;
- help in finding work;
- help in understanding how systems work that may be affecting the participant's life, for example, the criminal justice system, government assistance, child support, or the school system;
- help in negotiating those systems;
- help in figuring out how to find good child care;
- economic assistance, such as cash assistance, subsidies for child care, transportation, housing, or costs associated with going to work;
- affordable health insurance;
- help in obtaining or recovering a driver's license;
- information about the kinds of experience, skills, or credentials that are required for specific jobs;
- information about how to get training or schooling, and help in finding ways to pay for it;
- help in figuring out the structures that will make it possible for a participant to work on a consistent basis, whether it be child care or the school circumstances of the participant's children;
- help in planning ahead;
- help in solving problems that arise to threaten whatever progress the participant is making;
- help in overcoming negative messages and actions—from the participant him- or herself, or from others.

I always try to combine what I call "concrete" help with less tangible but sometimes equally important moral support.

I remember one of New Hope's Project Representatives telling me how dissatisfied she was with the amount of time she had to put into processing monthly benefits such as child care or health insurance. She wanted to be coun-

seling participants. I reminded her that no matter how much participants appreciated our staff counseling and listening, they would not appreciate it if it meant that they did not get their child care payment processed on time.

It was the combination of economic and noneconomic help that made New Hope effective. And the ratio of one to the other varied by each individual's situation. Some people wanted nothing more from us than our very affordable health insurance. Even if we knew that they might benefit from attending one of our workshops to get a better job, we could not force that.

One of the most rewarding aspects of this work was undoubtedly linked to a sense that we really had made a difference to the economic well-being of the participant. Whether it was the positive effect of the extra money from the wage supplement that allowed a family to buy new shoes for the kids, or to be able to take care of all the family's medical needs, or just to treat themselves to a meal out or go to a movie, we could see the impact.

Equally rewarding, but more difficult to effect, were the times when we helped someone make a breakthrough in the way they were approaching a problem. We could often do this by asking questions, rather than telling them what *we* thought. "Have you thought about what you would like to do for your child's care when school is out for the summer?" would open up more possibilities than simply saying, "You need to deal with child care when school gets out."

The biggest breakthroughs, of course, had to do with helping people deal with real and perceived job barriers. Job cycling is often linked to real negatives, such as low pay or dangerous working conditions. But it's also very hard for us to look at our own behavior when we are not succeeding, to face the question "Have you thought about what you could do differently?" It's not a question of fault; it's a question of efficacy—the sense that you really can do something to change the situation.

Just asking a question didn't guarantee that the participant was open to looking at things differently. But as we often discussed, we needed to be there and ready to go when the participant was ready to make a move.

---

## Support

One staff member followed up the participant's responses to her opening question by talking about the variety of summer camps available. I remember the pleasure of that participant when she discovered she could send her daughter to a six-week dance day camp, using her child care subsidy. Then the challenge was to ensure that all our staff understood this as well, so that they could offer their participants the same kinds of creative options as had the Project Representative in this example.

---

To the characteristics of respect and support, we add a third and equally central element: mutual accountability.

## Accountability

Accountability is a word that is frequently used to describe the responsibilities that we expect participants to accept for themselves. And this is certainly one part of what we mean. But in the experience at New Hope, we applied it to the reciprocal nature of the relationship we tried to forge with participants. We expected participants to be accountable to us for doing what was required for their own progress, but we also expected ourselves to do what we said we would do.

The whole premise of the New Hope offer was that if participants did their part (work at least 30 hours per week and provide us with regular documentation of that work), then we would provide them with access to work supports as long as they were eligible.

In training our staff, we asked them always to think in terms of reciprocity and specific actions. It was important to end every meeting by summarizing what each of us would do: the participant and the Project Representative.

This list includes:

- Understanding that it is the participant who gets him- or herself a job, not the staff. We can create opportunities, find job openings, help get interviews, but it is the participant in the end who must persuade the employer to hire him or her. It is the participant who will perform in ways that either keep the job, or lose it. From the get-go, we want to encourage the participant to take as much initiative as possible, since that is a precondition of long-term independence.
- We are responsible for doing our part in a timely manner.
- We need to follow up with people, especially if they have not followed through on something they agreed to do. A lot of programs practice the-ball-is-in-your-court method of working with participants. There is a limit to what program staff should do, but it should never be set at the participant's first failure to follow through.
- Program staff should model the behavior they expect from participants.
- Accountability includes helping staff and participants understand why someone is not following through and using that information to break through to that individual's next level of achievement.

There isn't just one type of approach or personality or work style that captures effective case management or job



coaching. The need to develop greater staff capacities never ends.

I wish we had been able to apply all that I've written more consistently in the original New Hope demonstration. But for all the gaps, I know that we did enough of the right things to make a difference for a lot of New Hope's participants. I know that because of feedback from the evaluators' reports. Even more important, I know that because of encounters, years after we closed the program, with participants who credit New Hope staff with teaching them how to plan, or find resources, or be more adept at solving problems in their lives. From that, I take not only satisfaction, but also great encouragement and motivation to keep improving our understanding of staff roles and development. It's what makes our work worth pursuing.

---

### Accountability

*No One Is Unemployable*, by Debra L. Angel and Elisabeth E. Harney, is one of the best resources I know to help staff develop a good sense of effective and positive accountability. For their Web site, go to <[www.worknettraining.com](http://www.worknettraining.com)> .

---

## 2004 Luxembourg Income Study Summer Workshop

The Luxembourg Income Study has made comparable over 130 large microdata sets containing comprehensive measures of income and economic well-being for a set of 29 modern industrialized welfare states. The LIS databank currently covers countries including: Australia, Austria, Belgium, Canada, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Hungary, Ireland, Israel, Italy, Luxembourg, Mexico, the Netherlands, Norway, Poland, Romania, Russia, the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Taiwan, the United Kingdom, and the United States.

The LIS Summer Workshop is a one-week pre- and post-doctoral workshop designed to introduce young scholars in the social sciences to comparative research in income distribution and social policy using the LIS database. The 2004 Summer Workshop, our sixteenth such event, will be held in Luxembourg. Applications are available from the LIS homepage at: <http://www.lisproject.org/workshop.htm> and are due by April 15, 2004. Please note that space is limited.

The language of instruction will be English. The course of study will include a mix of lectures and assistance and direction using the LIS database to explore a research issue chosen by the participant. Workshop faculty will include the entire LIS staff (including Timothy Smeeding, Overall Project Director; Lee Rainwater, LIS Research Director; and John Coder, LIS Technical Director) and other experienced LIS users.

For more information about the workshop, please contact:

Caroline de Tombeur, LIS Administrative Assistant, 17, rue des Pommiers, L-2343 Luxembourg City, Luxembourg (e-mail: [caroline@lisproject.org](mailto:caroline@lisproject.org)) or Kati Foley, LIS Administrative Assistant, 426 Eggers Hall, Syracuse University, Syracuse, NY 13244-1020 USA (e-mail: [lisaa@maxwell.syr.edu](mailto:lisaa@maxwell.syr.edu))

For information about the LIS Project, see <http://www.lisproject.org/>.

# How do welfare sanctions work? New findings from Wisconsin and Illinois

---

*Although time limits may receive more attention in the media, many more families have been directly affected by sanctions, and sanctions have arguably played a greater role in reshaping welfare recipients' day-to-day experiences.*

– Dan Bloom and Don Winstead, *Sanctions and Welfare Reform*,  
Brookings Welfare Reform Policy Brief 12 (2002)

---

Financial sanctions have long been used to enforce work requirements in the welfare system, but under Temporary Assistance for Needy Families (TANF), they have gained greater importance and are more severe. Fewer families are now exempt from work requirements. Federal regulations permit states to penalize recipients for infractions of the rules by the loss of the entire cash grant, and even of Food Stamp or Medicaid benefits. Conciliation and mediation procedures, once required under federal law to resolve problems before sanctions were imposed, are now optional for states. Benefit reductions are larger than under previous welfare-to-work programs; sanctions are put in place more quickly and may last longer.

Yet since passage of the 1996 reforms, policymakers and researchers have paid less attention to sanctions than to the precipitous declines in TANF caseloads that ensued. In 2000, Jack Tweedie, of the National Conference of State Legislatures (NCSL), noted that there were not yet clear data about sanctions. Two years later, Bloom and Winstead noted that there was still “little systematic data on how, and how often, sanctions are imposed.”<sup>1</sup>

The dearth of substantial information about sanctions has meant that many questions are very hard to answer. How often are sanctions used? Who is at risk of being sanctioned? What are the effects of sanctions on behavior? on family well-being? Do some kinds of sanctions work better than others? How much do recipients really know about sanctions policy? How effective is the *threat* of sanctions as opposed to their imposition?

Broad policy conclusions are made even more difficult by the extraordinary variety of sanction policies. A summary of state sanctions policies as of April 2000 found great variation in severity and duration. In 15 states, a first sanction resulted in termination of the entire benefit. In 33 states, a first sanction reduced cash benefits by widely varying amounts—for example, 110 percent of the adult portion, one-third of the benefit, or 10 percent of the benefit. The duration of the sanction was equally variable: it might last until compliance, or for a set period of time, which itself was different among states—a month, two months, three months. The maximum possible sanction varied also: in 34 states, it was termination of the

benefit, but in 15 states, only a partial reduction. And in answer to the question whether a first sanction could escalate to a maximum sanction, 23 states said yes, 24 said no.<sup>2</sup>

Existing research concerning sanctions, mostly based upon cross-sectional studies of those leaving welfare, suggests that sanctioned families resemble long-time welfare recipients in a number of respects.<sup>3</sup> They are more disadvantaged than even the average welfare recipient— younger, less educated, less likely to live with a partner and more likely to have been in an abusive relationship in the past year. They are more likely to have grown up in a welfare-receiving family or to have health problems or children with health problems. As a group, they are more likely to have immediate practical disadvantages also— higher levels of financial strain, as evidenced by utility cutoffs, no car, or no telephone service.<sup>4</sup>

The two projects summarized here broke new ground in the study of sanctions. Both made use of longitudinal data. Chi-Fang Wu, Maria Cancian, and Daniel R. Meyer used administrative data from Wisconsin to examine the dynamic patterns of sanctioning (their severity, timing, and duration), the factors associated with being sanctioned, and the relationship between sanctions and subsequent welfare outcomes for sanctioned women. Bong Joo Lee, Kristen Shook Slack, and Dan A. Lewis used survey and administrative data from the Illinois Families Study (IFS) to examine whether and how welfare sanctions are associated with work activity, levels of earnings, welfare receipt, and material hardships among TANF recipients.<sup>5</sup>

---

*Tell policymakers that half of the welfare caseload has left and that you do not know why or what has happened to them and stand back. They will want to know why you do not have answers now and why you cannot produce them next week. . . . They will not be specifically concerned about sanctions and time limits unless you show them that this population is significantly different from all exits.*

– Jack Tweedie, NCSL, in 2000

---

# Sanction policies and outcomes in Wisconsin

By Chi-Fang Wu, Maria Cancian, and Daniel R. Meyer<sup>6</sup>

Under the Wisconsin Works (W-2) program, cash benefits are based on the welfare participant's hours of participation and her tier on a four-stage employment ladder. Those who fail to participate in assigned activities can be fully or partially sanctioned. Wisconsin is unique in its pay-for-performance policy that reduces the benefit by \$5.15 for every hour of work obligation unfulfilled without "good cause," which is defined to include domestic abuse, failure of child care arrangements, or other circumstances beyond the control of the participant. Good cause must be documented, and W-2 agencies and caseworkers generally have a high degree of discretion in determining whether or not to grant an exemption. W-2 also includes the concept of "strikes." Participants who fail to take part in an activity, without good cause, may receive a strike. Three strikes render the participant ineligible for benefits for life. (Strikes have, thus far, been rarely imposed.)

Implementing sanction policies presented a serious logistical challenge. The state had to establish systems to track the hours each participant was required to work and the actual hours she worked, and to notify her if she was being sanctioned. In some cases, participants or their employers were late or failed to turn in the required paperwork, and inaccurate or inappropriate sanctions were imposed.<sup>7</sup>

## The dynamics of sanctioning

The sample in this study included all women who received TANF cash benefits in Wisconsin during the first year W-2 was implemented, September 1, 1997, to August 31, 1998 (somewhat over 17,000 women). Researchers followed their history and outcomes for four years after they entered a W-2 employment tier that provided cash benefits (that is, either of the two lowest of the four W-2 tiers). Administrative data came from Unemployment Insurance (UI) files and from the CARES (Client Assistance for Re-employment and Economic Support) system. CARES provides information on the timing and severity of sanctions, on program participation, welfare status and welfare history, and demographic information. It documents good cause exemptions and the reasons for benefit reduction—whether sanctions or some other reason.

Researchers examined only cases in which sanctions were actually imposed, looking also at the severity of the sanction, and did not include the threat of sanction. They took into account a wide range of individual and family characteristics and the welfare history and current status of the participants. Because traditional regression ap-

proaches using cross-sectional data do not consider the length of time sanctioned nor whether the participant remains at risk of being sanctioned, they chose to employ event-history analysis to capture the timing of benefit receipt and sanctions, the predictors of being sanctioned, and their severity.<sup>8</sup> They were also able to explore how long participants were sanctioned, and how quickly they left the welfare system after sanctioning. Comparison of the results from a simple logistic analysis and from event-history analysis showed that measuring the levels of change and the timing, duration, and severity of sanctions is important to a full understanding of their effects.

## How frequent are sanctions?

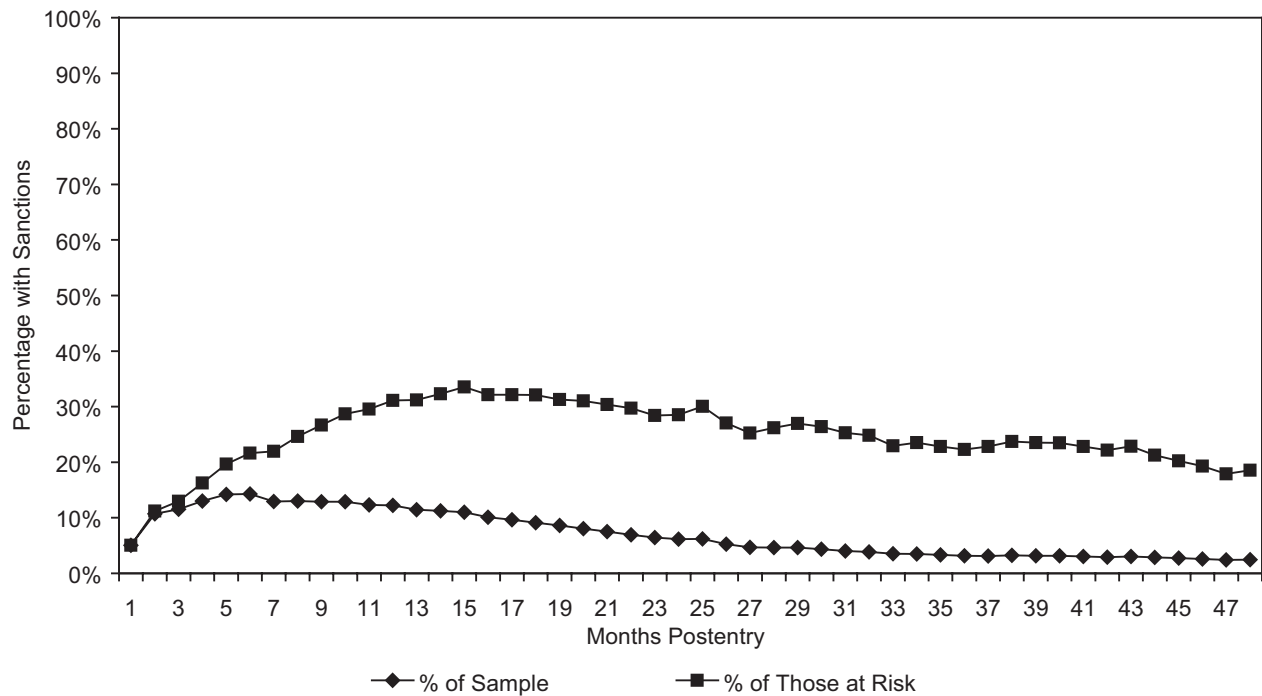
Calculating the frequency of sanctions is complicated by the very rapid decline in the numbers of women receiving cash assistance under W-2; at the end of the first year, fewer than half of the original participants were receiving cash benefits, and after four years, only 20 percent were. A simple calculation of the proportion of the original sample that was sanctioned showed that in the first month, only 5 percent were sanctioned; this percentage increased to 14 percent in the fifth and sixth months, and thereafter declined.

This estimate, however, ignores the fact that many women were no longer at risk of being sanctioned, because they no longer received cash benefits. Figure 1 shows what a difference emerges when the simple sanction rate is compared with the sanction rate among women still at risk of sanctioning (because they were still receiving benefits). The highest rate of sanctions was 34 percent of women, in the 15th month after entry, but even after four years, 19 percent of those in a cash benefit tier were being sanctioned. From the first year of data only, sanctioning emerges as a common experience, affecting 51 percent of the sample. The four-year time frame reveals that nearly two-thirds of recipients in this sample faced a sanction.

Sanction rates declined as the W-2 program matured. Among women entering in the second year of the program (September 1998 to August 1999) overall sanction rates were only 44 percent, and among third-year entrants they were only 37 percent.

## How severe are sanctions?

Researchers differentiated among low, high, and full sanctions. A low sanction is loss of less than 50 percent of the benefit, a high sanction is loss of 50–90 percent, and



**Figure 1. Frequency of Sanctions among W-2 Mothers in Lower Tiers.**

Overall sample: 17,119 W-2 participants who entered in lower (cash) tiers and received TANF benefits during the first year of implementation. This figure includes all spells for each woman.

Those at risk: W-2 participants who entered in lower (cash) tiers in first year of implementation and received cash benefits in each month after entry.

under a full sanction a woman is receiving either nothing or less than 10 percent of the benefit.

Most women in this Wisconsin sample received low sanctions; the percentage increased over the first 15 months and reached a maximum of 17 percent before leveling off at 11–13 percent. The rate of high sanctions also increased over time, to 12 percent, and then gradually declined. The proportion with full sanctions never exceeded 5 percent. The two most common patterns were no sanction (36 percent) or one or more partial sanctions, without ever a full sanction. But almost a quarter of women experienced both partial and full sanctions.

### What are the characteristics of those sanctioned?

As already shown, women who are sanctioned share many characteristics of disadvantage with long-term welfare recipients. It is important to distinguish the factors associated with their longer period of exposure to the risk of sanction from the factors that may simply be associated with their higher probability of remaining on cash assistance longer.

Event-history analysis makes it possible to estimate the relationships between a woman’s characteristics and the

time over which she was receiving benefits (and therefore at risk of sanctions), as well as between those same characteristics and the likelihood of being sanctioned. In a number of cases this approach yields substantially different results from the traditional logistic model. For example, the traditional model found that African Americans were more likely than whites to be sanctioned during their first spell on welfare, but found no discernible difference between Hispanics and whites in this respect. The event-history estimates suggest that part of the higher sanction rate for African Americans is associated with their slower rate of exit from welfare; they are simply at risk of sanction for longer. Hispanic participants left welfare more quickly than both African Americans and whites. When we take into account their shorter period of exposure, however, they were also more likely to be sanctioned than whites.

In other areas, too, event-history analysis revealed unexpected complexities. According to the simple analysis, the sanction rate was highest in Milwaukee County, where 77.3 percent of all participants and 92 percent of all black participants in the sample were living. Event-history analysis showed that Milwaukee County participants were actually less likely to be sanctioned than those in rural counties; they were merely on welfare longer. Likewise, the simple analysis suggested that those who were pregnant at entry or who came into the program as a



**Table 1**  
**The Likelihood of Being Sanctioned and the Severity of**  
**Sanctions in a Woman's First W-2 Spell, Discrete-Time Event-History Model**

Prototypical Groups <sup>a</sup>	Probability of Being Sanctioned <sup>b</sup>	Partial Sanctions	Full Sanctions
1. No work experience, long-term AFDC recipient, education less than HS, African American, lives in Milwaukee	73.7	69.9	11.4
2. Same as 1 except 8 quarters work experience	61.2	56.9	9.3
3. Same as 2 except no long-term AFDC reciprocity	52.8	49.0	7.1
4. Same as 3 except HS graduate	40.0	36.7	5.1
5. Same as 4 except white	30.2	27.5	3.6
6. Same as 5 except lives in urban area other than Milwaukee	46.2	44.0	2.9
7. Same as 6 except lives in rural area	31.6	30.4	1.0

**Note:** The sample includes all women who received TANF cash benefits in Wisconsin during the first year W-2 was implemented, September 1, 1997, to August 31, 1998.

<sup>a</sup>Table uses mean values for age of mother, language of mother, age of youngest child, number of children, household structure, entering cohorts, initial W-2 assignment, and unemployment rates.

<sup>b</sup>The predicted probability of being sanctioned in the discrete-time logistic model is within the first 12 months.

caretaker of a newborn (a status limited to the first 12 weeks after the child's birth) were less likely to be sanctioned, but event-history analysis showed they were actually *more* likely, once their shorter welfare experience was taken into account.<sup>9</sup>

Using the event-history approach, Table 1 estimates the probability that "prototypical" welfare participants will be sanctioned within the first 12 months that they are receiving welfare, and how severe those sanctions will be. For all these cases, researchers used the same mean values for a set of personal and demographic variables, varying only characteristics related to race, education and work experience, welfare history, and region of residence—factors already shown to be related to sanction status. The first case is that of a long-term welfare recipient, an African American woman with low education and no work experience, living in Milwaukee. Her estimated probability of being sanctioned is quite high, but as characteristics change, women's risk of sanction drops dramatically. For this set of prototypical recipients, however, it is always over 30 percent, whatever the combination of characteristics chosen.

### What happens after a sanction?

Just over 8,300 of the 17,000 women in this Wisconsin sample were sanctioned during their first W-2 spell. Sanction spells were quite short for most women. The probability of returning to full benefits after sanctioning was always higher than the probability of leaving W-2, regardless of the length of the sanction. Over 70 percent

returned to full-benefit status after a sanction—most of them in the next month.

Women who received a partial sanction (about 7,600 of the 8,300) most commonly were restored to full benefits in the next month, though each month about one in five left welfare completely. The risk that they would move from a partial to a full sanction was very low, less than 10 percent. Among the much smaller number of women who received a full sanction in the first welfare spell, 37 percent were restored to full benefit, and even among the very few who received two months of full sanctions, nearly 40 percent were restored to full benefits. Multiple sanctions were fairly common: 40 percent of women were sanctioned more than once and 14 percent four or more times.

### Do sanctions matter?

Exploration of this question is only just beginning. The Wisconsin research found that most commonly, women who were sanctioned returned rather quickly to full benefits, and that few participants spent any length of time in the sanctioned status—only 16 percent of those who received the more severe full sanction continued for longer than a month in that status. These findings can be interpreted as suggesting that sanctions are having their desired effect and inducing compliance with the requirements of the program. But an important minority of those sanctioned moved off welfare altogether, and it is not yet clear whether they and their children were getting by, or whether they faced serious economic distress.

# Sanctions policies and outcomes in Illinois

By Bong Joo Lee, Kristen Shook Slack, and Dan A. Lewis<sup>10</sup>

Illinois policy for moving TANF recipients into work combines somewhat generous incentives with gradual sanctions. To “make work pay,” there is a very liberal policy that disregards two-thirds of earned income indefinitely when determining TANF benefit levels. Illinois also “stops the clock” on time limits, putting in abeyance the 60-month lifetime limit if a recipient is working. For families that do not comply with TANF regulations, Illinois imposes a three-step sanction. At the first instance of noncompliance, cash benefits are reduced by half until the recipient cooperates. At the second instance, benefits are reduced by half for three months, even if the recipient begins cooperating within this three-month window. These two steps are known as “partial-grant” sanctions. After three instances of noncompliance, or three months under partial sanction without complying, the recipient loses all cash assistance for three months (this is a “full-grant” sanction).

To examine the effects of sanctions, researchers drew upon data from the Illinois Families Study, which is designed to assess the work, welfare use, and well-being of families affected by welfare reform. It uses a sample of nearly 1,900 TANF recipients, about half from Chicago, randomly selected from 1998 welfare caseloads. Data from three waves of interviews are linked to administrative data from the Illinois Human Services and Employment Security departments.<sup>11</sup>

This research focused on sanctions received by members of the IFS sample between January 1999 and March 2001. The study therefore included women receiving TANF at any point in that period (1,123 women). Over 80 percent were African American, and their average age was 31.5. They had had a first child at about age 19, and averaged 2.5 children, the youngest of whom was about 5. Nine percent were currently living with at least one other adult worker in the household; about the same percentage had worked informally in the past year. The recipients were predominantly long-term welfare users, and nearly a quarter reported either a chronic health condition or 10 or more depressive symptoms. About 60 percent had a high school diploma or GED, and 37 percent reported low job skills. Three-quarters lacked either a valid driver’s license or access to a car. More positively, they reported high levels of social support (a score of 10.11 on a scale from 4 to 12); 15 percent were receiving formal support and 46 percent informal support from the other parent of one or more children.

Researchers used the welfare agency’s reports of sanctions, grant amount changes, and actual case closings to measure the different kinds of grant loss. They distin-

guished grant reductions caused by sanctions from grant reductions resulting from, for example, increased earnings or marriage. They were also able to identify formal sanctions that were initiated but lifted before benefits were actually reduced. If sanctions have behavioral effects, one would expect that the *threat* of sanctions would predict employment and welfare behavior as well as, or better than, the actual imposition of sanctions. In that case, reconciled sanctions and imposed sanctions should affect work and welfare behaviors in the same way. If, instead, sanctions are primarily a punitive tool, one would expect greater effects on behavior and on family hardship from their actual imposition.

Researchers also considered the possibility that TANF recipients might change their behavior in response to their knowledge of sanction policies. Thus recipients were asked what they knew about five policies under welfare reform: time limits, earnings disregards, the stopped clock, the availability of transitional Medicaid, and the availability of food stamps after an individual left TANF. Although this measure of welfare knowledge did not include a question about sanctions per se, those who knew more about welfare policies in general seemed more likely to have greater understanding of the sanction policy, and to be more likely to be employed and less likely to experience hardships. Respondents, on average, had correct information on 2 or 3 of these 5 welfare policies.

About 20 percent of women in the sample had been sanctioned. For a further 8 percent, sanctions were initiated but lifted before the grant was actually reduced. About two-thirds had other types of grant reductions, and almost 80 percent had their TANF cases closed at least once during the study period, for a variety of reasons.

Those who experienced sanctions were more likely to be receiving cash benefits in the two quarters following the study’s sanction risk period; they were less likely to be working and had significantly lower levels of earnings than those who were not sanctioned in these subsequent months. They also reported that they had more difficulties obtaining food and greater levels of hardship.

Those whose sanctions were reconciled before any reduction in grant took place were similar in almost every way, except that they reported more difficulties in keeping up with the rent than those who were not under threat of sanction. It is not clear why the threat of sanction had this effect. Perhaps housing difficulties get in the way of complying with the welfare rules—or perhaps those with housing difficulties are more likely to have the sanction

removed before the actual grant loss because they can demonstrate “good cause” (e.g., they had never received notices of the impending sanction) or have more incentive to begin cooperating rather than face additional hardship in so crucial an area of their lives.

## Sanctions, work, and earnings

Even taking into account other factors that might affect women’s work and earnings—education, previous work, and welfare experiences—researchers found that those who were sanctioned were only about half as likely to be employed as those who were not. Neither the threat of sanction, nor case closing, had any effect on the likelihood that these women were working.<sup>12</sup>

Sanctions are by no means the only influence on employment. Many other characteristics were also associated with a significantly lower probability of working among this sample of welfare recipients. Older mothers, mothers with recent experience of domestic violence or of depression, with no car, and a history of informal jobs were less likely to be working. So too were mothers who had more social support or received formal child support payments. Women who had themselves received welfare as children were also less likely to be working.

The outcome for earnings was, as one would expect, very similar to that for work. Those who were sanctioned earned, on average, about \$1,325 less in the last two quarters of the study period than did those who were not.

## Sanctions and welfare receipt

Once other characteristics were controlled, sanctions appear to have had little effect on the likelihood that a woman was no longer receiving welfare. The biggest

---

*Margaret Simpson, 22, a mother of three in Cincinnati, lost her welfare check for seven months after she failed to show up for her state’s job readiness program.*

*“I wasn’t paying attention,” she said, “There was a letter with my check. Who pays attention to a letter with a check? You pay attention to the check.”*

*But eventually, when the check quit arriving, Simpson complied with the rules by helping her caseworker track down the father of her children to collect child support, working on her high school equivalency test and attending a job-preparation course. A new check is on the way.*

— Barbara Vobejda and Judith Havemann,  
“Sanctions Fuel Drop in Welfare Rolls,”  
*Washington Post*, March 23, 1998

---

effect came, instead, from marriage—women who had married were over three times more likely to be off welfare. So too were Hispanic women and women whose caseworkers took more time to explain the program rules. Women who were long-time recipients, who had greater social support, or who lived in public or subsidized (Section 8) housing were less likely to have left welfare.

## Sanctions and hardship

Not surprisingly, TANF grant loss for whatever reason was associated with increased levels of hardship, even when the benefit loss was due to higher earnings. Sanctions had a particularly strong effect on food hardships, which were three times more likely among the sanctioned than among those not sanctioned. Other types of grant loss were associated with increased rent and utility hardships, and with greater perceived hardship overall.

In summary, then, researchers found evidence that sanctions did not appear to promote work and reduce welfare dependency; in fact, they were associated with less work and lower earnings, and the sanctioned were more likely to be experiencing food hardship. As measured in this analysis, sanction threats, arguably the real force behind sanction policies, had no association with greater work effort or lower welfare dependency. Sanctions, it appears, are not primarily working to bring about behavioral changes among welfare recipients, but to penalize those who do not follow the rules.

One finding in particular is interesting. In general, the more people knew about the program rules, the more likely they were to be off welfare, employed, and to have higher earnings and lower food hardship. At this point one can only speculate about the reasons. Is being better informed a marker for unobserved characteristics, such as a woman’s motivation and efficacy? Does better understanding of the system help women navigate other supportive services such as job training and job search assistance, or food stamps?

Implicit in the strategy underlying the new welfare policies is that a potential or current recipient can know what the rules are, understand their meaning, and grasp the consequences of not following them. Yet recent research has demonstrated that many do not make these connections.<sup>13</sup> This analysis demonstrates the importance of paying attention to what recipients know and instituting policies to increase their understanding, so as to affect behavior and improve family well-being. ■

---

<sup>13</sup>J. Tweedie, “Sanctions and Exits: What Do States Know about Families Who Leave Welfare Because of Sanctions and Time Limits?” Working Paper 148, Joint Center for Poverty Research, 2000; D. Bloom and D. Winstead, *Sanctions and Welfare Reform*, Brookings

Welfare Reform Policy Brief 12, Brookings Institution, Washington, DC, January 2002.

<sup>2</sup>State Policy Documentation Project, “Summary of State Sanction Policies as of April 2000,” Center for Law and Social Policy and Center on Budget and Policy Priorities, Washington, DC, June 2000 (see the project Web site at <<http://www.spdp.org>>).

<sup>3</sup>L. Pavetti, M. Derr, and H. Hesketh, *Review of Sanction Policies and Research Studies: Final Literature Review*. Report submitted to the U.S. Department of Health and Human Services. Mathematica Policy Research, Washington, DC, 2003.

<sup>4</sup>Research based on official administrative records in the Illinois Families Study largely confirms the disadvantaged status of sanctioned recipients. D. Lewis, A. Stevens, L. Amsden, K. Hasson, K. Slack, and colleagues, *Preserving the Gains, Rethinking the Losses: Welfare in Illinois Five Years after Reform*. Third Annual Report of the Illinois Families Study. Institute for Policy Research, Evanston, IL, 2003. See also K. Shook, *Does the Loss of Welfare Income Increase the Risk of Child Welfare Involvement among Families Receiving Welfare?* Ann Arbor, MI: UMI Dissertation Services, 1999.

<sup>5</sup>The Illinois sanctions study is extensively discussed in B. Lee, K. Slack, and D. Lewis, “Are Welfare Sanctions Working as Intended? Employment, Continued Welfare Receipt, and Material Hardship among TANF-Recipient Families,” working paper, University of Wisconsin–Madison, September 2003. A report of the Wisconsin project is C.-F. Wu, M. Cancian, and D. Meyer, “How Do Welfare Sanctions Work?” working paper, University of Wisconsin–Madison, January 2004.

<sup>6</sup>Chi-Fang Wu is an IRP graduate fellow, Maria Cancian is Professor of Public Affairs and Social Work and Daniel R. Meyer is Professor of Social Work at the University of Wisconsin–Madison. Both are IRP affiliates.

<sup>7</sup>Wisconsin Legislative Audit Bureau, *Sanctioning of Wisconsin Works (W-2) Participants*, Madison, WI, 2002.

<sup>8</sup>For a detailed discussion of event history analysis, see P. Allison, *Event History Analysis: Regression for Longitudinal Event Data* (Beverly Hills: Sage, 1984) and *Survival Analysis Using the SAS System: A Practical Guide* (Cary, NC: SAS Institute, 1995); D. Cox and D. Oakes, *Analysis of Survival Data* (London: Chapman and Hall, 1984).

<sup>9</sup>Although program rules do not permit sanctioning of women in the Caretaker of Newborn tier, women who transition from that tier to one of the other tiers are especially likely to be sanctioned.

<sup>10</sup>Bong Joo Lee is an Associate Professor at Seoul National University, Korea, and a faculty associate at Chapin Hall Center for Children at the University of Chicago. Kristen Shook Slack is Assistant Professor of Social Work at the University of Wisconsin–Madison and an IRP affiliate, and Dan A. Lewis is Professor of Education and Social Policy at Northwestern University.

<sup>11</sup>D. Lewis, K. Shook, A. Stevens, P. Kleppner, J. Lewis, and S. Riger, *Work, Welfare, and Well-Being: An Independent Look at Welfare Reform in Illinois*. Project Description and First-Year Report of the Illinois Families Study. Institute for Policy Research, Evanston, IL, 2000.

<sup>12</sup>Other types of grant reductions had a *positive* effect on working, unsurprising because such reductions are quite likely to be tied to the earnings of TANF recipients.

<sup>13</sup>S. Anderson, “Ensuring the Stability of Welfare-to-Work Exits: The Importance of Recipient Knowledge about Work Incentives,” *Social Work* 47, no. 2 (2002): 162–70; Lewis and colleagues, *Preserving the Gains, Rethinking the Losses*.



# Benefit-cost analysis in the evaluation of child welfare programs

By E. Michael Foster and E. Wayne Holden

---

E. Michael Foster is Professor of Health Policy and Administration at Pennsylvania State University and an IRP affiliate. E. Wayne Holden is President of ORC Macro, Inc.

---

Broad evaluation of the financial consequences of social programs has been somewhat of an orphan stepchild of policy analysis, especially within states. A full benefit-cost analysis is expensive and complex to implement, requiring that evaluators reach far beyond the immediate framework of the program to consider the effects of choices made across a wide programmatic and social spectrum. It takes time to do properly, and public officials tend to need swift answers. But particularly now, as greater integration of social welfare programs is actively under consideration in many jurisdictions, benefit-cost analysis should be a central tool of social welfare program evaluation.

This article suggests how a benefit-cost analysis of child welfare programs might be set up, describing the structure of the analysis, identifying potential data sources, and noting difficulties.<sup>1</sup> An accompanying brief article (pp. 50–52) describes a benefit-cost analysis of an early childhood intervention program, the Chicago Child-Parent Centers, carried out as part of the evaluation by the Chicago Longitudinal Study under the direction of IRP affiliate Arthur Reynolds.

Spanning multiple programs at the local, state, and federal level, child welfare expenditures annually exceeded \$14 billion in the late 1990s. These expenditures had for some time been growing substantially; between 1986 and 1996, for example, federal foster care maintenance payments grew almost fivefold. In consequence, states began to experiment with different arrangements for organizing and delivering child welfare services, in the hope of controlling program costs and improving outcomes for children. Beginning in 1998, the federal government began allowing states to waive certain child welfare program requirements under Titles IV-B and IV-E of the Social Security Act. The waivers, which both reflected and further stimulated state initiatives, allowed states to change policies and service delivery. For example, states could now use IV-E funds for therapeutic, aftercare, or home-based services that previously had to be funded from

Medicaid or block grants, and could introduce service innovations such as managed care.

By February 2003, 26 Child Welfare Waiver Demonstrations had been implemented in 17 states. The federal Administration for Children and Families (ACF) that administers the waiver program summarized the tenor of these waivers as follows:

Collectively, the demonstration projects are aimed at reducing the number of children in foster care, the length of time in foster care, the use of more restrictive and costly placement settings, re-allegations of abuse and neglect, and re-entry into foster care. Some states have proposed discrete interventions focused on specific child welfare populations, while others are experimenting with flexible use of funds to produce system-wide reforms.<sup>2</sup>

Under the waivers, states are required to evaluate policies implemented or services provided. Such evaluations must assess the cost effectiveness of the project and its fiscal consequences for state and local jurisdictions. A primary reason for this mandate is to ensure that the program meets a key requirement of the waiver—that it be “cost neutral,” i.e., federal expenditures should be no higher in the new program than they would be in the absence of a waiver.

The fiscal analysis does, however, highlight two other possible consequences of a program innovation. First, the spillover of program effects onto other levels of government and onto the child’s family may result in “cost shifting,” where expenditures move from one budget or payment source to another; for example, the costs of mental health services may be shifted from Medicaid to Title IV-E. Cost shifting may or may not be associated with actual changes in the child’s condition or circumstances. If such changes do occur, then a second consequence, “cost offset,” may come into play; for example, improvements in the child’s condition may reduce the need for special services. Some offsets might actually be negative in the longer term, say if children returned to the community become involved in delinquency.

Cost neutrality, cost shifting, and cost offset naturally lead to a benefit-cost analysis more extensive than the “fiscal analysis” required under the terms of the federal waivers, which primarily emphasize budgeted government expenditures closely connected to the program un-

der analysis. A full benefit-cost analysis includes all resource uses. Some costs will be “opportunity costs,” which do not involve an explicit payment but may require the time of parents and caregivers. Others will be societal—the costs to victims of crime or the benefits of a lower crime rate.

### Steps in a benefit-cost analysis

The steps in a benefit-cost analysis can be succinctly conveyed (see box) but the choices prove to be much more complicated. Defining the program in this example is relatively simple: we establish it as the IV-E waiver. But the costs and benefits of a program will vary with the perspective. Costs important from one perspective will be irrelevant from another—social welfare agencies, for example, generally ignore costs that do not appear on their budgets. Some policy outcomes may be a cost from one perspective and a benefit from another; mental health services may return a child to the family (a benefit to the family and the mental health agency), but if the family thereby becomes eligible for cash assistance, the state’s Temporary Assistance for Needy Families (TANF) agency incurs a cost.

A benefit-cost analysis therefore needs to specify one or more perspectives. In the example we use, there are at least four points of view: (1) the child’s family or caregivers; (2) state agencies, both child welfare and other relevant agencies such as Medicaid or TANF providers; (3) other members of society, including taxpayers, victims of crime, and private or community services; and (4) society as a whole, a category which comprises the other three.

The time frame—the period during which the policy or program will be evaluated—differs from the analytic horizon—the period over which costs and benefits will be measured. In this example, the time frame is set by the waiver period. But services received during childhood may affect (indeed, are intended to affect) a child through the transition to adulthood and beyond. For example, intensive studies of early childhood interventions such as

the Chicago Child-Parent Centers have commonly continued into early adulthood, decades after the evaluation was implemented (see the accompanying article by Reynolds and colleagues). The analytic horizon is therefore likely to be limited mainly by practical concerns: available resources, or the ability of the evaluators to track participants over time or to project future costs and benefits.

Under the child welfare waivers, the outcomes are fairly generally specified as the improved health and safety of children, greater permanency of placement, better school performance, and reduced delinquency. Valuing these benefits and their costs in dollar terms involves multiple sources of data and a series of methodological choices. Reducing delinquency, for example, means fewer dollars spent in the juvenile justice system, but how should a dollar value be placed on crimes uncommitted or the benefits to the family of the child who remains out of the court system? The *net* benefits of a program also involve discounting (the conversion of future costs and benefits into today’s dollars) and sensitivity analyses (the recalculation of costs and benefits under alternative but plausible assumptions).

### Counting the costs and benefits of a Title IV-E waiver

Table 1 summarizes the potential sources of costs for program and service changes and for the consequences of those changes, categorizing them by the four primary perspectives we specified earlier. These categories represent the dollar costs and savings alone, and so only partly capture the effects of the waiver. For example, the personal or caregiver costs of abuse and neglect are identified as the costs of the medical services involved; these are clearly modest when compared to the emotional or social costs. Moreover, policy and program changes designed to reduce government expenditures in the long run may not do so in the shorter term. For example, the costs per participant of services under the first years of the Wisconsin welfare reforms (W-2) considerably outpaced Wisconsin’s per capita expenditure under the last years of AFDC.

Among the outcomes listed in Table 1, delinquency and school performance are only two of many outcomes that might be analyzed if research resources allowed. Children leaving foster care, for example, are at risk of behaviors and outcomes costly to themselves and society: substance abuse, homelessness, victimization, early pregnancy, and future welfare use.<sup>3</sup> Any improvements in these long-term outcomes that emerge from the waiver program are only partially captured through improved schooling and reduced delinquency.

As simply one example, we consider in detail the potential costs and benefits associated with one outcome from

#### Steps in a Benefit-Cost Analysis

1. Define the program, policy, or intervention being evaluated.
2. Specify the study perspectives.
3. Select the time frame and analytic horizon.
4. Identify relevant benefits and costs.
5. Measure those effects in dollar terms.
6. Produce a summary measure of the policy’s net benefits.

**Table 1**  
**Identifying Costs and Benefits**

Change/Outcome	Perspective			
	Parents and Caregivers	Government Agencies	Other Members of Society	Society as a Whole
Policy Changes	Costs of care (time and money)	Child welfare administrative costs Court costs	Taxes to pay for administrative and court costs	Resources used to administer programs and courts
Service Changes	Time and money costs	Medicaid expenditures Block grant expenditures	Taxes to pay for services and programs	Resources used to provide services
Accelerated Permanent Placement of Child	Costs of care (time and money)	TANF payments Medicaid expenditures Child welfare expenditures Adoption assistance	Community services Taxes to pay for services and programs	Resources used to provide services and care Resources used to administer programs Parental or caregiver time
Health and Safety	Costs to parents of services related to abuse and neglect	Child protection services, including court costs	Taxes to pay for services and programs	Resources used to provide services
Child's School Performance	Costs of services Time spent dealing with school-related problems	Expenditures on school and school services	Taxes to pay for services and programs Future taxes paid by child Effect on community	Parental or caregiver time Productivity and related societal benefits Resources used to provide services and schooling
Child's Delinquency	Time and money costs of delinquent child	Juvenile justice costs	Costs of victimization Taxes to pay for juvenile justice costs	Parental or caregiver time Costs to victims Juvenile justice costs

Table 1: acceleration of the child's permanent placement, either by reuniting the child with the parents or caregivers, or through adoption into another family.

*Reunification* has implications for many government programs, including other sources of child welfare funds, TANF, and Medicaid. Increases in these payments represent costs to taxpayers and, except for TANF cash assistance, they are also costs to society. State policy determines how other state funds, such as the Social Services Block Grants (SSBG), are used to fund child welfare programs such as family support, protective services for children, and special services for children at risk. Children returned home under the waiver might use any or all of these services.

TANF payments seem likely to increase under child welfare waivers: when children return to parents or kin, the family may thereby become eligible for cash assistance. But these payments, which constitute a cost to the state and to taxpayers, are a benefit to the family. Reunification may also produce emotional benefits for the family, but it consumes resources of time in caregiving and cash—for example, for child care or health services not covered under Medicaid. Other social groups may accrue costs, as children and their families make greater use of

the services of community or nonprofit agencies. From a broader societal perspective, however, the benefits and costs of reunification may offset each other.

*Adoption.* To the extent the waiver increases the rate of adoption without improving a child's need for special services, it may increase the cost of assistance the state offers to families willing to adopt such a child. But services to a child in foster care may improve a child's condition, increase the likelihood of adoption, and reduce the need for special services thereafter. This seems less likely if the waiver is aimed at children with particularly severe emotional and behavioral problems; in this respect, the intended target population of the waiver-based program enters into consideration of its benefits and costs.

### Sources of data

Because the effects of a waiver program may be complex and far-reaching, the data needed to understand them must be drawn from multiple sources. These will include (1) government payments under a variety of programs, (2) estimates of parental time and financial resources, and (3) estimates of costs and benefits to other members of society.

## Government payments

Administrative data on public costs such as those for TANF or adoption assistance are widely available, though the quality and accessibility of actual expenditure (as opposed to budgeted) data vary greatly. Tracking a family's Medicaid costs represents a special challenge because parents or caregivers may not accurately report their use of services or know the costs for particular services. Access to some data, such as child protective data, may require informed consent of the subjects.

One goal of a benefit-cost analysis is to identify those individuals for whom the costs of the waiver were particularly high or low, perhaps in order to target future waivers to specific groups for whom the net benefits are large. Aggregate program expenses for treatment or comparison groups will not provide these kinds of answers; instead, evaluators must have data concerning payments made on behalf of and services provided to individuals. Such records must often be gathered from programs at different levels of government, posing serious problems of confidentiality, access, accuracy, and reliability, as well as technical difficulties in linking very different types of data. Management information systems or billing data may not track participation in services or programs funded through block grants. Payments may not be linked to specific individuals or services. Estimating the actual costs of family preservation or a service provided to a particular child may therefore be next to impossible. One alternative may be to obtain estimated costs for similar services from other studies.

## Costs to parents

Self-reports appear to be the only feasible source of such information. Information on parents' time use can be converted to dollars using an estimated value of parental or caregiver time. There is some debate over how to value this time and over the appropriate measure of opportunity costs—the value of the activities forgone to participate in services or to monitor the child. One argument is that this time may not otherwise be spent productively—parents' leisure time may simply be reduced. The counter argument is that parents may miss work or have less time for productive activities such as housework or caring for other children. Furthermore, from an economist's perspective, lost leisure time has value as a good that parents choose to consume, and for which they pay through reduced earnings. Such "lost" time is in general valued at the individual's wage rate.

## Costs and benefits for other members of society and for society as a whole

Changes in health and safety, school performance, and delinquency all affect society more generally.

*Health and safety.* Medical services and child welfare programs (including court costs) may be estimated in

dollars using self-reports or administrative data. What of the extreme case where abuse results in the death of the child? Considerable controversy surrounds the valuing of human life but even conservative estimates may exceed several million dollars.<sup>4</sup> These costs are borne by the child's family, taxpayers, and society at large.

*School performance.* The costs of schooling and school services can be estimated from self-reports, review of school records, school budgets, and national cost estimates. There are, for example, estimates of the benefits of high school completion, including increased earnings and broader social benefits.<sup>5</sup>

*Delinquency.* Self-reports of crime and involvement with juvenile justice are merely a starting point; an accurate accounting will require review of court records. It may be difficult to estimate court costs or the costs of time spent in juvenile detention facilities in a specific community or state. One alternative, again, is to draw upon estimates from elsewhere—the Washington Institute of Public Policy, for example, provides estimates of juvenile justice facility costs—but these may be of limited applicability outside a particular area.<sup>6</sup> General estimates of the costs to victims, including medical costs, time lost from work, and pain or suffering, do exist and make it possible to value self-reported crimes in dollar terms.<sup>7</sup>

## Calculating and presenting net benefits

Calculating net benefits involves more than simply subtracting gross costs from gross benefits.<sup>8</sup> One must allow for future changes, for example, as children leave foster care and make the transition to adulthood. Because a dollar today is worth more than a dollar next year, even in the absence of inflation, future payments must be converted into their current value, but neither government agencies nor academics have been able to settle on a single discount rate. The best approach is to employ several within the range of annual values commonly used (2–10 percent), thus providing a plausible range of estimated net benefits.<sup>9</sup>

More important, however, is that net benefits are so presented as to reflect their true uncertainty. As with any calculation based on a sample of study participants, the net benefits for a given evaluation would be different were a different sample chosen for evaluation. For this reason, net benefits have a sampling error, as does the mean, or any sample statistic. This uncertainty can be captured by producing a confidence interval for the net benefits.

Other forms of uncertainty are introduced by the choices the evaluator makes—which measure or dollar figure to use from the range of costs for juvenile detention, for example. Sensitivity analyses, in which the evaluator calculates net benefits using a reasonable range of figures, will illustrate how the net benefit calculation changes



under alternative, plausible assumptions. Finally, it may simply be impossible to measure some benefits in dollar terms, but it is important to try to quantify them as best one can with nonmonetary measures of well-being or satisfaction. These findings are particularly important if the measured net benefits are negative. In that case, the program may still be worth undertaking; the policymaker needs to decide if the value of the unmeasured benefits exceeds the observed net costs of the program.

## Advantages and difficulties of benefit-cost analysis

### The advantages

The outcomes traditionally used to evaluate child welfare programs have generally focused on limited goals, such as length of placement or reunification, and have followed children for a relatively short time afterward. Child welfare professionals have increasingly come to realize that these criteria are too narrow and can lead to policy decisions that are bad for children—reunification, for example, may not promote long-term success. They are thus moving toward a broader array of outcomes that include measures of school performance and educational achievement, among others.<sup>10</sup> Some evaluators have suggested the children be followed for a minimum of three to five years after leaving the child welfare system.

As perspectives and outcomes expand, benefit-cost analysis provides a way of prioritizing a potentially very large body of information, focusing on those outcomes that have the greatest potential benefits or costs from a particular perspective. Benefit-cost analyses of waiver programs can document shifts in expenditures among different services and describe the distribution of expenditures across different types of individuals. For instance, total expenditures may remain the same but the waiver may reduce (or increase) the concentration of expenditures on a few high-cost children—information of real importance to policymakers. As noted earlier, waivers may dramatically change the types of services children receive. TANF expenditures will almost certainly rise if greater numbers of children are reunited with families. Medicaid expenditures may drop as Title IV-E funds are used to cover more services, or may increase as these funds are used to link individuals and families to new services. A benefit-cost analysis also emphasizes that a waiver mechanism is more than a matter of fiscal reform or the reorganization of service delivery, and that it may have far-reaching effects outside these narrow boundaries on other organizations, communities, and society as a whole.

### The difficulties

One problem associated with benefit-cost analysis is overreliance on net benefits as the criterion for determin-

ing a project's merits. Although a benefit-cost analysis can document the differential effects on taxpayers, parents, and so on, the net benefit figure, a measure of efficiency, simply subtracts all (discounted) costs from benefits without regard to who bears the costs or reaps the benefits. An obvious alternative criterion is equity.

A second, related pitfall is an exaggerated sense of precision attached to the net benefits. As we noted earlier, confidence intervals and sensitivity analyses should accompany any presentation of the benefits and costs, providing readers with a sense of the degree to which the study's findings might be a chance occurrence or might be sensitive to assumptions made in performing the analysis. At best, a benefit-cost analysis can only provide a range of plausible estimates.

Third, any benefit-cost calculation is limited to *measurable* costs and benefits. Possible nonmonetary costs and benefits should also be included in the report.

Looming above these pitfalls is the question of the cost of such a thorough analysis. Obtaining and manipulating large administrative databases, amassing the necessary budget data, finding and surveying families (many of them disadvantaged and unstable), and tracking outcomes over time are all very expensive undertakings. True, the quantity and quality of public administrative data are improving. This is in part because of the increased importance of longitudinal data to policymakers and program administrators who must track families over time to fulfill the requirements of laws such as the 1996 welfare reforms. Federal requirements to develop comprehensive databases on child outcomes have also stimulated better data gathering.<sup>11</sup>

Evaluators may reduce costs by relying on unconfirmed self-reports or national estimates, but will thereby reduce accuracy. One option is to lower the analytic horizon, for example by considering only benefits and costs until the child's 18th birthday. The evaluation will then likely underestimate the waiver's benefits, for the transition to adulthood is a crucial marker of the success or failure of a child welfare policy. If the net benefits observed during a shorter period are positive, then the results of the analysis may still be clear. If they are not, then the study findings and policy implications remain ambiguous.

Another option for reducing expenses is to limit the scope of the analysis to certain outcomes or behaviors and ignore others, as we earlier demonstrated in limiting consideration to school performance and delinquency. But how should we choose those outcomes? Importance to policymakers is one criterion; the potential dollar magnitude of the effects is another. Relevant benefits and costs are typically identified on the basis of the theory underlying the program or policy as well as prior research and practice. Limiting the list to a manageable size requires that the evaluators engage in a complex balancing act on

the basis of often insufficient information. How closely are particular data about policy and service changes linked to the desired outcomes? How large are the costs of collecting such information? Prior research may suggest outcomes that are likely to be most sensitive to the waiver program, but there may be very little prior research to draw upon.

Benefit-cost analysis has been an often neglected or underfunded feature in the evaluation of large federal demonstration programs. Yet it is consistent with the growing emphasis on broader outcome measures for child welfare programs and, indeed, for social programs in many other areas. The Title IV waiver demonstration programs provide a unique opportunity for the rigorous evaluation of alternative service organization and delivery mechanisms in the child welfare area, and a broad benefit-cost analysis should be seen as an essential feature of this kind of evaluation. ■

---

<sup>1</sup>This summary is based upon E. M. Foster and E. W. Holden, "Benefit-Cost Analyses of the Child Welfare Demonstration Projects: Understanding the Resource Implications of the IV-E Waivers," *Children and Youth Services Review* 24, nos. 6/7 (2002): 431–53, which is used here by permission of the publisher.

<sup>2</sup> Description from the ACF Web site, <<http://www.acf.hhs.gov/programs/cb/initiatives/cwwaiver.htm>>.

<sup>3</sup>M. Courtney, "The Costs of Child Protection in the Context of Welfare Reform," *Future of Children* 8, no. 1 (1998): 88–103.

<sup>4</sup>R. Zerbe and D. Dively, *Benefit-Cost Analysis in Theory and Practice* (New York: Harper Collins College Publishers, 1994).

<sup>5</sup>M. Cohen, "The Monetary Value of Saving a High-Risk Youth," *Journal of Quantitative Criminology* 14 (1998): 5–33.

<sup>6</sup>Washington Institute for Public Policy, *The Comparative Costs and Benefits of Programs to Reduce Crime: A Review of National Research Findings with Implications for Washington State* (Olympia: Washington Institute for Public Policy, 1999).

<sup>7</sup>M. Cohen, T. Miller, and S. Rossman, "The Costs and Consequences of Violent Behavior in the United States," in *Understanding and Preventing Violence: Consequences and Control of Violence*, ed. A. Reiss and J. Rother (Washington, DC: National Academy Press, 1994), 4:67–166.

<sup>8</sup>R. Plotnick, "Using Benefit-Cost Analysis to Assess Child Abuse Prevention and Intervention Programs," *Child Welfare* 78, no. 3 (1999): 381–407.

<sup>9</sup>The Office of Management and Budget uses 2 percent, the Congressional Budget Office uses 10 percent. These figures assume costs and benefits are already adjusted for inflation.

<sup>10</sup>One example is the Casey Outcomes and Decision-Making Project, whose publications include *Improving the Quality of Children's Services: A Working Paper on Outcome-Based Models of Service Delivery and Managed Care*, and *Assessing Outcomes in Child Welfare Services: Principles, Concepts, and a Framework of Core Indicators*, available on the Web site of the American Humane Society, <[http://www.americanhumane.org/site/PageServer?pagename=pc\\_best\\_practice\\_casey\\_homepage](http://www.americanhumane.org/site/PageServer?pagename=pc_best_practice_casey_homepage)>. "Measuring Success in Child Welfare," a Web site maintained by the Children and Family Research Center of the School of Social Work at the University of Illinois at Urbana-Champaign, summarizes the Casey and similar initiatives: <<http://cfrwww.social.uiuc.edu/pubs/Natl.Surv/backgrnd.htm>>.

<sup>11</sup>J. Waldfogel, "Child Welfare Research: How Adequate Are the Data?" *Children and Youth Services Review* 22, no. 9/10 (2000): 705–41.

# A cost-benefit analysis of the Chicago Child-Parent Centers

The study summarized here is the first cost-benefit analysis of a large-scale, publicly funded early childhood intervention for preschool children and their families, the Chicago Child-Parent Centers (CPC; see box). A cost-benefit analysis of the CPC was considered highly important because the length, intensity, and breadth of the services it provides to children are substantially greater than in most other intervention programs; it was also feasible, because the short- and long-term outcomes of the program have measurable economic benefits. The approach and the methodology used in the analysis are briefly described here, illustrating the procedures outlined in the accompanying article by E. Michael Foster and E. Wayne Holden.

## Establishing outcomes

Individual children participated in the CPC for anywhere from 2 to 6 years, setting the time frame for the evaluation. The analytic horizon, however, extends over two decades, for many of the most important consequences of early intervention programs are expected to occur in adolescence and over the transition to adulthood.

Research regarding the CPC and other intervention programs suggested that five main categories of outcomes would be important to a cost-benefit analysis: (1) reductions in expenditures for school remedial services; (2) reductions in criminal justice system expenditures; (3) reductions in child welfare system expenditures; (4) averted tangible costs to victims of crime and child maltreatment; and (5) increased earnings capacity of program participants and tax revenues. Because rates of public assistance were similar for parents in the program and comparison groups, the study did not estimate the benefits of the program for welfare participation.

## Estimating benefits and costs

The study distinguished three types of benefits, comparable to the perspectives described by Foster and Holden. These are, first, the benefits returned to children and parents, but not directly to others in society. These benefits are both long-range (e.g., the children would have greater earnings capacity in adulthood, because more would finish high school) and short-range (e.g., child care provided to the parents). Second are benefits to the general public. These include lowered or averted expenditures for remedial education and social welfare programs, reduced criminal justice expenditures because rates of crime and delinquency are lower, and increased tax revenues

as a result of the children's greater earnings capacity. Finally there are benefits to society at large—basically the sum of benefits to program participants and to the general public.

Program benefits and costs were calculated in dollars, converted to 1998 values to adjust for inflation. The present values of future costs and benefits were also computed in 1998 dollars and evaluated at a baseline age of 3, the beginning point of the program for all children. The annual discount rate was set at 3 percent, a rate recommended by the U.S. Public Health Service and the U.S. General Accounting Office; the study also undertook a sensitivity analysis using alternative rates.

## Costs

Taxpayer costs were broadly defined: they included all outlays for staff, family, and community support, administration of the program and related school district expenses, instructional materials, transportation and community services, and capital depreciation and interest.

The Chicago Child-Parent Center (CPC) program is, after Head Start, the nation's oldest, federally funded preschool program. Opened in 1967, CPC provides services for children aged 3 to 9 at 24 sites in high-poverty neighborhoods, including, among other things, structured educational activities, extensive outreach and parental involvement, and health and nutrition services. CPC is the subject of a quasi-experimental evaluation, the Chicago Longitudinal Study (CLS), directed by IRP affiliate Arthur Reynolds, which continues to investigate the life course development of over 1,500 children born in 1980. The vast majority of these children from low-income families were black; a small percentage were Hispanic. Almost a thousand of these children were enrolled in the CPC; also included in the study is a comparison group of about 500 low-income children who did not participate in the CPC but attended full-day kindergarten in randomly selected schools in the same Chicago neighborhoods.

A full discussion is A. Reynolds, J. Temple, D. Robertson, and E. Mann, "Age 21 Cost-Benefit Analysis of the Title 1 Chicago Child-Parent Centers," *Educational Evaluation and Policy Analysis* 24, no. 4 (2002): 267–303. Note: "cost-benefit" and "benefit-cost" are equivalent terms.

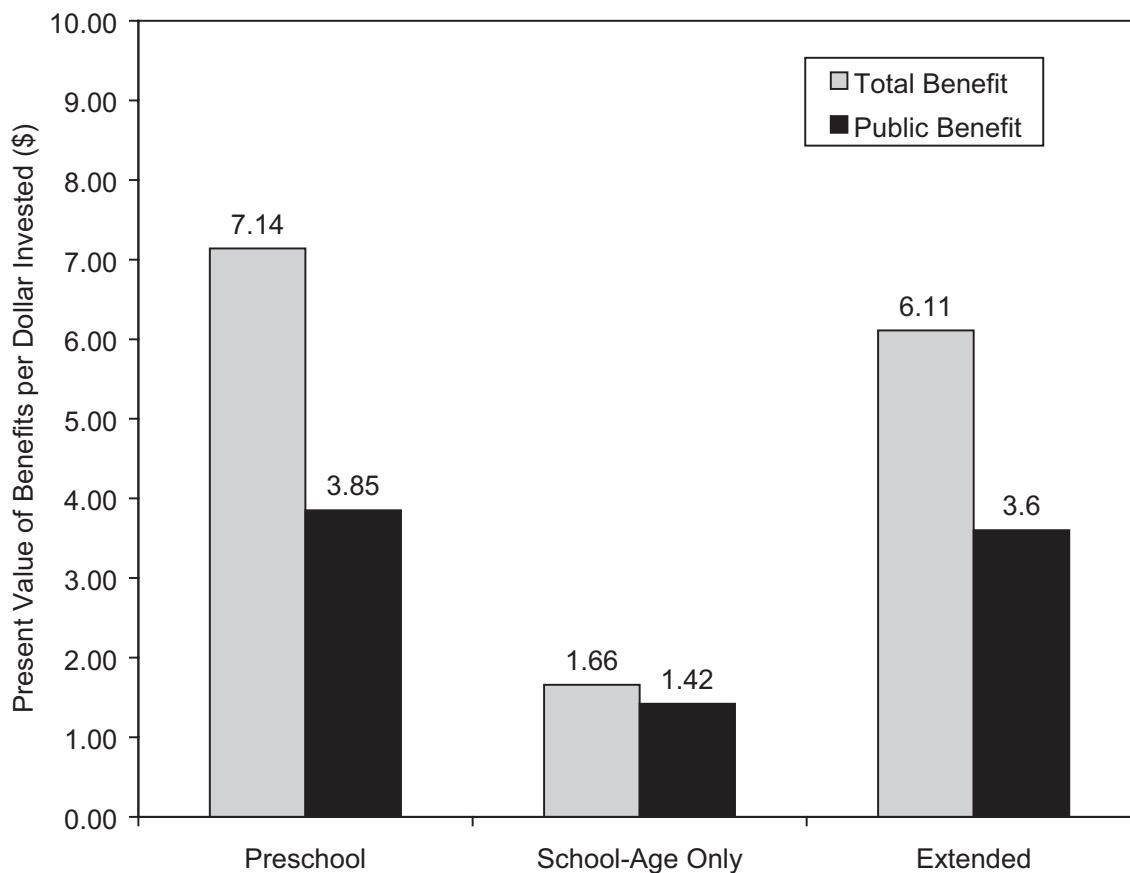


Figure 1. Costs and benefits estimated by the Chicago Longitudinal Study (see box, p. 50) for the Chicago Child-Parent Centers program.

The study did not include the cost for the half- or full-day CPC kindergarten program because the comparison group also participated in a kindergarten program, and costs for these children approximated or exceeded the costs for the CPC kindergarten. Direct costs for the program were supplemented with other costs, including the opportunity costs of parents' participation, based on 10 hours of participation per month at the minimum wage of \$3.35 an hour.

### Benefits

The study estimated the savings on school remedial services by using average per-pupil expenditures of the Chicago School District for grade retention and special education placement from kindergarten to grade 12.

Increases in lifetime earnings for ages 18–65 were estimated from the differences in high school completion rates between the program and comparison groups. The calculations were based on multiple data sources—school records and youth surveys, U.S. Census and Department of Labor data for black, full-time workers aged 25–29—and four categories of educational attainment from “less than high school” to “college graduation or more.” The

projections also included estimates of fringe benefits received and taxes paid, all requiring choices among different projected tax rates, real income growth, and benefit packages.

For the criminal justice and child welfare systems, benefits are estimated as savings resulting from reduced crime and abuse rates. Because the strongest predictor of adult crime is juvenile crime, the study projected decreases in expenditures for the adult criminal justice system from juvenile arrests; this required manual and computer searches of court records in Chicago and other cities. Criminal justice expenditures were based on administrative expenses and weighted national averages of the costs of residential or community treatment and incarceration. The study defined crime-victim savings as direct expenditures incurred as a result of delinquent or criminal behavior, exclusive of pain and suffering. Dollar amounts were based on national estimates of tangible losses to victims of violent and property crime.

Savings for the child welfare system were based on court and child protective system data for Chicago. Our measure was the number of referrals (“substantiated reports”) to the juvenile court by the Illinois Department of Child



and Family Services. About 70 percent of children that are the subject of substantiated reports receive in-home services, and the remainder are placed in foster care. To estimate the savings arising from fewer substantiated cases of abuse and neglect, the study used the tangible losses estimated by the National Institute of Justice: medical and mental health care, police and fire services, and lost productivity of victims and their families (work or school missed, and some legal expenses).

A direct benefit to families was the child care provided by enrollment of children in half-day preschool five days a week; the value of parents' free time was estimated at the minimum wage of \$3.35 an hour in 1986 (adjusted for inflation).

For both the preschool program only and the extended (4–6 year) program, the study estimated that just over 40 percent of the social benefits of the program went to program participants; a similar proportion constituted savings to government, and about 15 percent reflected savings to potential victims of crime averted by children's participation in the program.

### **The benefit:cost ratio for the CPC program**

Whether researchers considered the preschool program only, the school-age program only, or the extended program that included both, the estimated ratio of benefits to costs was impressively high (Figure 1). For every dollar invested in the preschool program, the most intensive and comprehensive component, about \$4 were returned to the general public through government savings on school remedial services, criminal justice and child welfare system costs, and averted crime-victim expenditures. About \$7 were returned to society at large through increased economic well-being and reduced public expenditures on educational and social welfare services. The returns were higher for preschool boys than for preschool girls (\$9.06 vs. \$4.67 per dollar invested). The largest share of benefits was attributable to the link of preschool participation with higher rates of school completion and lower rates of juvenile arrest.

The study's estimates are conservative. For example, for victims of crime only tangible savings (arising from property losses, hospitalization, and lost productivity) were measured. Had the study included averted intangible costs such as pain and suffering, the benefits per preschool participant would have risen by over \$20,000 (intangible costs to victims of crime are generally estimated at about three times tangible costs).

In their discussion of the net benefits, the study authors note some limitations. First, the estimates of increased earnings capacity and of cost savings in the criminal justice system were, of necessity, projections. Future eco-

nomie and social conditions are difficult to predict with confidence and could affect the assumptions underlying these projections.

Second, the benefits of some outcomes of the program were not measured—indeed, almost impossible to measure. These include well-documented nonmarket benefits of higher school attainment, such as health status and future family health status, fertility decisions, and even the educational attainment of the children of these children.

Third, causal inference is more difficult in a quasi-experimental than in a rigorously experimental design. Without further study, results should not be generally applied outside the context of urban minority populations and programs with a successful history of implementation. Nevertheless, the results reflect many years of different analyses in which the mechanisms by which the programs effects were achieved have been clarified, minimizing the likelihood that effects may be due to the study sample selected. Encouragingly, the patterns of effect sizes and economic benefits observed in this study are consistent with those of other programs implemented in different contexts and time periods. Our cost-benefit findings, for example, are similar to those reported in studies of the High/Scope Perry Preschool Program and the Prenatal/Early Infancy Project. Findings also compare favorably with the federally funded WIC program.<sup>1</sup>

---

<sup>1</sup>These and other programs were discussed in a special issue of *Focus*, Vol. 19:1, Summer/Fall 1997, devoted to the subject of early childhood interventions.

## Recent IRP Special Reports

Robert, S. A. *Wisconsin's family care long-term care pilot program: Care managers' perspectives on progress and challenges*. 2003. 38 pp. SR87.

SR 87 is part of an ongoing study of Wisconsin's Family Care pilot long-term care program, focusing on the perspectives of care managers working in the Family Care program in four counties implementing Care Management Organizations.

Bartfeld, J., and David, C. *Food insecurity in Wisconsin, 1996–2000*. 2003. 31 pp. SR86

SR 86 includes comparisons of the rate of food insecurity among households that differ in income, household composition, age, race, health and disability status, employment status, and location. It also compares Wisconsin and the country as a whole.

Cancian, M., Haveman, R., Meyer, D. R., and Wolfe, B. *The employment, earnings, and income of single mothers in Wisconsin who left cash assistance: Comparisons among three cohorts*. 2003. 53 pp. SR85.

SR 85 compares women leaving welfare in 1999 with those who left welfare in 1995 and 1997, and finds substantially higher rates of exit in the later periods. Later leavers are somewhat more likely to work, but their earnings are lower. On average, substantial earnings growth is outweighed by declines in benefits, so that women's measured net income is lower after they leave welfare. The income drop is more pronounced for those who left in 1995 than for those who left in 1997 and 1999.

Bartfeld, J. *Forgiveness of state-owed child support arrears*. 2003. 43 pp. SR84.

Public concern about the magnitude of child support arrearages, the growing awareness of the complex circumstances behind such arrears, and the likely effects on all concerned, especially the children, has led to increasing interest in potential remedies. This report considers one general policy approach: the forgiveness of arrears owed to the state. It provides an overview of the magnitude of arrears, the contributing factors, and the problems stemming from high arrears. In addition, it discusses current arrears-forgiveness programs around the country and highlights the many unanswered questions that remain about the impacts of such programs.

## Recent IRP Discussion Papers

Stormer, A., and Harrison, G. G. "Does household food security affect cognitive and social development of kindergartners?" 2003. 46 pp. DP 1276-03.

Fitzgerald, J. M., and Ribar, D. C. "Transitions in welfare participation and female headship." 2003. 30 pp. DP 1275-03.

Magnuson, K. "The effect of increases in welfare mothers' education on their young children's academic and behavioral outcomes: Evidence from the National Evaluation of Welfare-to-Work Strategies Child Outcomes Study." 2003. 61 pp. DP 1274-03.

Bruce, M. A. "Contextual complexity and violent delinquency among black and white males." 2003. 38 pp. DP 1273-03.

Wolfe, B., Scrivner, S., with Snyder, A. "The devil may be in the details: How the characteristics of SCHIP programs affect take-up." 2003. 39 pp. DP 1272-03.

Scrivner, S. and Wolfe, B. "Universal preschool: Much to gain but who will pay?" 2003. 80 pp. DP 1271-03.

Chernick, H., and Reschovsky, A. "State fiscal responses to welfare reform during recessions: Lessons for the future." 2003. 25 pp. DP 1270-03.

Holzer, H. J., Lane, J. I., and Vilhuber, L. "Escaping low earnings: The role of employer characteristics and changes." 2003. 32 pp. DP 1269-03.

Holzer, H. J., Raphael, S., and Stoll, M. A. "Employer demand for ex-offenders: Recent evidence from Los Angeles." 2003. 35 pp. DP 1268-03.

Holzer, H. J., Raphael, S., and Stoll, M. A. "Employers in the boom: How did the hiring of unskilled workers change during the 1990s?" 2003. 46 pp. DP 1267-03.

Flinn, C. J. "Minimum wage effects on labor market outcomes under search with bargaining." 2003. 56 pp. DP 1266-03.

Moffitt, R. "The role of non-financial factors in exit and entry in the TANF program." 2003. 66 pp. DP 1265-03.

Moffitt, R. "The role of randomized field trials in social science research: A perspective from evaluations of reforms of social welfare programs." 2003. 58 pp. DP 1264-03.

Fram, M. S. "Managing to parent: social support, social capital, and parenting practices among welfare-participating mothers with young children." 2003. 33 pp. DP 1263-03.

Bollinger, C. and Hagstrom, P. "Food Stamp Program participation of refugees and immigrants: Measurement error correction for immigrant status." 2003. 44 pp. DP 1262-03.

Wilson, F. D. "Labor specialization, ethnicity, and metropolitan labor markets." 2003. 47 pp. DP 1261-03.

Moffitt, R. "Milton Friedman, the Negative Income Tax, and the evolution of U.S. welfare policy." 2003. 31 pp. DP 1260-03.

Swann, C. A. "The dynamics of prenatal WIC participation." 2003. 30 pp. DP 1259-03.

Plotnick, R. D., Garfinkel, I., McLanahan, S. S., and Ku, I. "The impact of child support enforcement policy on nonmarital childbearing." 2002. 31 pp. DP 1258-02.

Oliver, P. and Yocom, J. "Racial disparities in criminal justice: Madison and Dane County in context." 2002. 48 pp. DP 1257-02.

Chernick, H. and Reimers, C. "Welfare reform and New York City's low-income population." 2002. 63 pp. DP 1256-02.

Bitler, M., Currie, J., and Scholz, J. K. "WIC eligibility and participation." 2002. 65 pp. DP 1255-02.

Holzer, H. J., Raphael, S., and Stoll, M. A. "Perceived criminality, criminal background checks, and the racial hiring practices of employers." 2002. 43 pp. DP 1254-02.

Bartfeld, J. "Single mothers and emergency food assistance in the welfare reform era." 2002. 34 pp. DP 1253-02.

Bhattacharya, J., Currie, J., and Haider, S. "Food insecurity or poverty? Measuring need-related dietary adequacy." 2002. 33 pp. DP 1252-02.

Hu, Y.-W. and Wolfe, B. "Health inequality between black and white women." 2002. 46 pp. DP 1251-02.

Edelhoch, M., Liu, Q., Martin, L., and Wiseman, M. "The impact of South Carolina's Family Independence Program on movers, stayers, and those in between." 2002. 49 pp. DP 1250-02.

Dunifon, R. and Kowaleski-Jones, L. "Associations between participation in the National School Lunch Program, food insecurity, and child well-being." 2002. 22 pp. DP 1249-02.

Quillian, L. "The decline of male employment in low-income black neighborhoods, 1950-1990: Space and industrial restructuring in an urban employment crisis." 2002. 40 pp. DP 1248-02.

Holzer, H. J. and Offner, P. "Trends in employment outcomes of young black men, 1979-2000." 2002. 35 pp. DP 1247-02.

Chatterji, P., Bonuck, K., Dhawan, S., and Deb, N. "WIC participation and the initiation and duration of breastfeeding." 2002. 43 pp. DP 1246-02.

# IRP Publications Order Form

- Focus** (1 copy free of charge; multiple copies \$3.00 each; formatted text of issues may be downloaded from the IRP Web site).

**INDIVIDUAL PUBLICATIONS:** (Please fill in number or title and author)

**Discussion Papers** (may also be downloaded free from the IRP Web site)

(USA & Canada \$3.50; Overseas \$8.50) \_\_\_\_\_

**Reprints**

(USA & Canada \$4.00; Overseas \$9.00) \_\_\_\_\_

**Special Reports** (prices vary; may be also downloaded free from the IRP Web site) \_\_\_\_\_

Send to: Publications  
Institute for Research on Poverty  
1180 Observatory Drive  
Madison, WI 53706

Phone: (608) 262-6358  
Fax: (608) 265-3119

Name: \_\_\_\_\_

Address: \_\_\_\_\_

City

State

Zip

- Please indicate here if this is a change of address.

---

---

## An Appeal for Support of *Focus*

For over 25 years, the IRP newsletter *Focus* has been used to inform individuals, organizations, libraries, and members of the media about important research into the causes and effects of poverty and about policies to ameliorate it. *Focus* has always been distributed free of charge, and we plan to continue doing so. Each issue is also posted in full on the IRP World Wide Web site.

The increasing costs of producing and mailing the printed copies, in times of straitened funding, and the breadth of public access to the Internet, have led us to consider whether or not to continue to produce printed copies, or to move entirely to electronic production. We encourage those who wish to see *Focus* continue as a printed publication to send contributions to the U.W. Foundation/IRP Fund, using the form below. We are grateful for your support.

---

---

### Institute for Research on Poverty

**I wish to support the continued printing of the IRP newsletter, *Focus*.**

**Enclosed is my gift of \$\_\_\_\_\_.**

**Name** \_\_\_\_\_

**Address** \_\_\_\_\_

**City** \_\_\_\_\_ **State** \_\_\_\_\_ **Zip** \_\_\_\_\_

*Please make check payable to UW Foundation/IRP Fund*

*Address: Institute for Research on Poverty, 1180 Observatory Drive, Madison, WI 53706*



# ***Focus***

**Institute for Research on Poverty  
University of Wisconsin–Madison  
1180 Observatory Drive  
3412 Social Science Building  
Madison, Wisconsin 53706**

---

**Read *Focus*  
on the Web**

<http://www.ssc.wisc.edu/irp/>

---