The Effect of Welfare on Child Outcomes

Janet Currie

There is broad support for the idea that welfare should benefit poor children. Yet most research on welfare programs, as well as much of the debate about welfare reform, has focused on the way that parents respond to incentives created by welfare, rather than on its effects on children. Less work has been devoted to the fundamental question of whether any of the web of programs supporting poor families benefit children.

If it can be shown that they do, then there are many other questions to be addressed: First, are the benefits short or long term? Second, which types of programs or combinations of programs are most effective; for example, do cash or in-kind programs produce bigger benefits for children? Third, do welfare programs have different effects on different groups, and if so why? Fourth, how exactly do successful programs work? And finally, can efficacious programs pass the more stringent test of cost-effectiveness?

This review focuses on the eight large federal programs shown in Table 7-1: Aid to Families with Dependent Children (AFDC), which has been replaced with the new Temporary Aid for Needy Families program (TANF); the Earned Income Tax Credit (EITC); housing assistance; Food Stamps; the Supplemental Feeding Program for Women, Infants, and Children (WIC); school nutrition programs; Medicaid; and Head Start. The programs are evaluated with respect to their effects on the health and educational achievement of children. Where possible, documented effects on long-term outcomes are noted. The first section of this chapter is a brief discussion of how we know what we know about these programs. The evidence regarding the effects of cash programs and in-kind programs, respectively, is then reviewed in the next two sections.
### TABLE 7.1 Trends in Program Expenditures (billion 1995 $)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cash Transfers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFDC</td>
<td>23.8</td>
<td>21.8</td>
<td>21.8</td>
<td>22.0</td>
</tr>
<tr>
<td>Federal only</td>
<td>13.1</td>
<td>11.8</td>
<td>11.9</td>
<td>12.0</td>
</tr>
<tr>
<td>Earned Income Tax Credit Total</td>
<td>3.4</td>
<td>3.7</td>
<td>8.1</td>
<td>22.2</td>
</tr>
<tr>
<td>Refunded portion of credit</td>
<td>2.5</td>
<td>2.6</td>
<td>6.2</td>
<td>19.0</td>
</tr>
<tr>
<td><strong>In Kind Transfers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Housing assistance</td>
<td>7.0</td>
<td>10.0</td>
<td>18.2</td>
<td>23.7</td>
</tr>
<tr>
<td>Food Stamps</td>
<td>13.5</td>
<td>17.4</td>
<td>19.4</td>
<td>25.7</td>
</tr>
<tr>
<td>WIC</td>
<td>0.7</td>
<td>1.4</td>
<td>2.5</td>
<td>3.5</td>
</tr>
<tr>
<td>School nutrition</td>
<td>5.4</td>
<td>5.8</td>
<td>4.3</td>
<td>5.3</td>
</tr>
<tr>
<td>School lunch</td>
<td>0.4</td>
<td>0.5</td>
<td>0.7</td>
<td>1.2</td>
</tr>
<tr>
<td><strong>Medicaid</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>35.1</td>
<td>46.3</td>
<td>76.3</td>
<td>111.2</td>
</tr>
<tr>
<td>Federal only</td>
<td>20.1</td>
<td>27.0</td>
<td>47.8</td>
<td>86.6</td>
</tr>
<tr>
<td>To dependent children</td>
<td>6.3</td>
<td>6.0</td>
<td>10.7</td>
<td>17.8</td>
</tr>
<tr>
<td>To adults in families with dependent children</td>
<td>5.9</td>
<td>6.4</td>
<td>10.1</td>
<td>14.0</td>
</tr>
<tr>
<td>Head Start</td>
<td>4.1</td>
<td>1.3</td>
<td>1.9</td>
<td>3.5</td>
</tr>
</tbody>
</table>

---

The fundamental problem facing researchers and policy makers is that the children of welfare recipients may have bad outcomes for reasons that have nothing to do with the receipt of assistance per se. It is possible that a program could have substantial benefits for poor children and still leave many children disadvantaged relative to better-off peers.

Evidently, parents of children on welfare are worse off than other parents in observable ways; they are poorer, likely to have less education, and may also have health problems. Many datasets available to researchers contain at least crude measures of these observable variables so that observed differences between parents on welfare and other parents can be accounted for using standard regression models.

To take a simple example, suppose that children of high school dropouts have lower scores on standardized tests than children of college graduates. Then if mothers on welfare are more likely to be high school dropouts than college graduates, a simple comparison of the two group’s average scores might tell you more about the effects of maternal education than about the effects of welfare. A simple way to “control” for the effects of education in order to focus on the effects of welfare might involve drawing a sample of high school dropouts and comparing children of welfare mothers to other children within this group. Any differences between the welfare children and the others could then be attributed to welfare use and not to maternal education. Multiple regression techniques simply allow one to control for the effects of several observable variables at the same time.

The problem becomes much more difficult however if parents on welfare also differ from other parents in ways that are not observed. For example, they may lack motivation or be discouraged by previous misfortune. Failure to properly control for these differences could lead one to incorrectly infer that it was being on welfare that was associated with negative child outcomes, rather than these underlying conditions. Some underlying problem, such as maternal depression, might cause both welfare dependence and negative child outcomes.

There are basically two approaches to this issue of unobserved characteristics. First, one may design a social experiment, randomly assigning eligibles to a “treatment” group and a “control” group. Random assignment ensures that, on average, the two groups will have the same observed and unobserved characteristics. In principle, one can then assess the effect of the treatment simply by comparing mean outcomes for the two groups, just as one would do in a drug trial. The key advantage of an experimental evaluation is its transparency.

One disadvantage of social experiments is that they may be very expensive. But there are several disadvantages in addition to high cost (Heckman, 1990). These include differential attrition between treatments and controls (which causes the treatment group to become less and less like the comparison group over time); the fact that subjects assigned to the control group may not accept their fate passively (for example, subjects denied training in a government program might...
sign up for an alternative program; and the fact that it may be difficult to use the experiment to examine differential effects of the treatment on different groups.

Nonexperimental evaluations attempt to control statistically for unobserved variables associated both with participation in the program and with the outcome of interest. One method of doing this is to find a third set of variables, called "instruments," that are associated with participation in the program but not with the important unobserved variables. For example, a researcher interested in the effects of participation in Medicaid on child health might argue that the generosity of state AFDC benefits is associated with participation in Medicaid because of the link between AFDC recipiency and Medicaid eligibility, but that the level of AFDC benefits does not have any effect on child health other than through its effect on participation in Medicaid. If this assumption were true, then the level of AFDC benefits would qualify as an "instrumental variable."

This instrument would be used (along with other observable characteristics of the mother) to predict Medicaid participation, and predicted participation would be substituted for actual participation in the model explaining child health. The idea is that predicted participation will depend only on observable characteristics and differences in state AFDC benefit levels, and not on the unobserved characteristics of the mother. The procedure is analogous to an experiment in which AFDC benefit levels are varied across states, Medicaid participation responds, and only this source of variation in participation rates is used to identify the effects of Medicaid on health.

The difficulty with instrumental variables techniques is that the key assumptions may not be satisfied. Suppose that states with more generous AFDC benefits also have higher-income populations and that higher incomes are associated with better child health. Then unless one takes account of this relationship, one will tend to find a spurious positive relationship between participation in Medicaid and child health. Alternatively, suppose that states raise AFDC benefit levels in response to poor child health. Then one might observe a spurious negative relationship between predicted Medicaid participation and child health.

An alternative approach involves assuming that the relevant omitted characteristics are fixed within a family or for the same child over time. Suppose for example that the relevant unobserved variable is maternal attitudes towards education and that this remains fixed over some period of time. Suppose further that one sibling participated in Head Start and one did not. Then comparing the sibling who participated to the one that did not provides a measure of the effect of Head Start that is not affected by the fact that, on average, mothers of Head Start children may have more positive (or negative?) views of education than other similarly situated mothers. Of course, the problem with this approach is that the relevant variable may also not be fixed within households or over time.

The studies discussed below all rely on one of these methodological approaches. Their conclusions are only as valid as the assumptions underlying the chosen approach. It is in cases where the same result has been obtained using different assumptions and data sources that we can be most confident of the conclusions.

WHAT WE KNOW ABOUT CASH PROGRAMS

Aid to Families with Dependent Children

The term "welfare" has usually been identified with the Aid to Families with Dependent Children program. This oldest and largest of the federal welfare programs provided cash transfers to (predominantly female-headed) families with children. This is the program that recent welfare reforms (the Personal Responsibility and Work Opportunity Reconciliation Act of 1996 [PRWORA]) effectively ended, replacing it with the new Temporary Aid for Needy Families program. TANF differs from AFDC because it ends the "entitlement" of all needy families to welfare benefits, because it introduces time limits on welfare benefits and because it provides states with much more latitude in developing their own welfare programs. Nevertheless, since most of what we know about cash welfare programs comes from studies of AFDC, and because many states will respond to TANF by only gradually altering their AFDC programs, it is of interest to summarize this literature here.

Like TANF, AFDC was administered at the state level within federal guidelines. As a result, program characteristics varied widely from state to state. For example, as of January 1993, the maximum monthly AFDC grant for a one-parent family of four persons varied from $164 in Alabama to $923 in Alaska (U.S. House of Representatives, 1993). On average the federal government pays 54 percent of benefit costs, as shown in Table 7-1. The continuous erosion of real AFDC benefit levels over the past 15 years provides compelling evidence of the unpopularity of this program: the average monthly AFDC benefit declined from $483 (1993 dollars) in 1980 to $373 in 1993, even though the average family size remained constant at three persons (U.S. House of Representatives, 1994).

One of the problems involved in evaluating the effects of AFDC on children is that the benefits of a cash transfer program can be expected to be diffuse. Small increases in household expenditures on a wide range of items may produce overall benefits for children without affecting any one indicator a great deal. A second problem is that although income is often used as a shorthand summary of a household's socioeconomic status, it is in practice extremely difficult to separate the effects of income from the effects of other family background characteristics including neighborhoods (Mayer, 1996).

Most research about the effects of AFDC on children focuses on the fact that daughters of women who participate in AFDC are themselves more likely to participate (cf. Gottschalk, 1990; Murray, 1984). What is less clear is whether the relationship is causal or whether it merely reflects the fact that the children of the poor are more likely to be poor—older studies tended to conclude that the
relationship was not causal, but studies using more recent data have questioned
this conclusion. See Moffitt (1992) for a fuller discussion of this issue.

There has been comparatively little research linking maternal AFDC participation
to other child outcomes, but the empirical issues are the same. First, it is
necessary to control for some measure of income as well as for AFDC status
since otherwise the estimated effects of participation are likely to reflect the
relative poverty of AFDC mothers. Second, within the group of poor women,
one would like to control for the fact that women choose whether or not to go
onto AFDC. Blank and Ruggles (1996) show that only 60 percent of eligible
women actually take up welfare benefits. Those who do are likely to differ from
those who do not in many unobservables respects.

Hill and O’Neill (1994) find that, when instrumental variables methods are
used to take account of unobserved variables that might be correlated with AFDC
status, AFDC participation has no effect on children’s scores on a standardized
test of vocabulary, given income. Currie (1995a) confirms that their results hold
up even when sibling comparisons are used to account for unobserved maternal
background characteristics. Currie and Cole (1993) use data from the 1979 to
1988 waves of the National Longitudinal Survey of Youth (NLSY) to examine
the effect of AFDC participation during pregnancy on the utilization of prenatal
care and birthweight. They use both sibling comparisons and instrumental vari-
ables methods to take account of unobserved variables that might be correlated
with both participation in the AFDC program and outcomes, and find that AFDC
participation has no additional significant effect on birthweight given income.
Together, these studies suggest that income from AFDC has much the same
effect on children as family income from any other source.

The Earned Income Tax Credit:
A Comparison to the Negative Income Tax

The slack in the growth of AFDC payments over time has been taken up by
the growth in expenditures on the Earned Income Tax Credit, which doubled
between 1975 and 1990. The EITC was introduced in 1975 as a means of
granting tax relief to low-income tax payers. Because it is administered through
the tax system, the EITC is not always viewed as a welfare program. However,
unlike most tax credits, the EITC is “refundable,” that is, if the amount of the
credit exceeds the taxpayer’s federal income tax liability, then the difference is
refunded. Table 7-1 shows that, in fact, most EITC expenditures are outlays of
this kind rather than forgone tax dollars. The EITC differs from traditional cash
welfare programs primarily because the majority of recipients work and benefits
are available to all kinds of families. Thus, it creates fewer perverse incentives
than AFDC.

2 If it is difficult to identify the effects of cash transfers under AFDC, the
problems involved in identifying the effects of the EITC are even more formi-
dable. The fundamental problem is that the amount of the credit depends on the
parents’ earnings, and earnings are likely to reflect many unobserved factors
relevant to child well-being. However, the EITC is in many respects similar to
the negative income tax (NIT), an income guarantee program that was subjected
to exhaustive scrutiny through four large-scale social experiments, although it
was never implemented. The four experiments were conducted in New Jersey
and Pennsylvania; Seattle and Denver; Gary, Indiana; and rural areas of North
Carolina and Iowa. It is important to note that the North Carolina and Gary
samples were much poorer than the others.

The income guarantees paid out under the NIT program were large relative
to cash transfers that have been made under the EITC. The average payments in
the Seattle-Denver experiment, for example, ranged from $919 to $2,031 (1972
dollars), depending on the treatment group. By way of comparison, the poverty
line for a family of three persons was $3,099 in 1972. In 1992, the maximum
EITC was $1,384 and the poverty line $11,280. Since NIT participants were
randomly assigned to “treatment” and “control” groups, the NIT experiments
provide a unique opportunity to assess the effects of income transfers per se on
the well-being of children in poor families.

Despite the large transfers, findings about the effects of the NIT are inconsis-
tent across studies and experimental populations. In addition, econometric esti-
mates are sometimes at odds with those derived from simple comparisons of
treatments and controls. For example, Kohn and Wolin (1979) find that the
mean birthweight of infants born to the treatment group in the Gary experiment
was actually lower than the birthweight of the controls. Yet estimates from their
structural model suggest that the infants of treatments had higher birthweights in
9 out of 12 maternal age groups.

O’Connor et al. (1976) examine the effect of the NIT on child nutrition using
data from the rural experiment. Among subjects in North Carolina, they found
positive and significant treatment effects on nutrient intakes. However, the treat-
ment did not appear to have any significant effect in Iowa, a finding that the
authors attribute to the relative poverty of the North Carolina sample.
Maynard and Crawford (1976) found that elementary school children from NIT families in North Carolina showed statistically significant improvements in attendance, standardized tests, and grades. However, there were no effects for elementary school children in Iowa. Once again, this pattern of results is attributed to the fact that the children in North Carolina were more disadvantaged than those in Iowa. Maynard and Murnane (1979) found that in the Gary experiment the NIT treatment had positive effects on reading scores of young children but that these effects were statistically significant only among children whose families had been in the program for 3 or more years.

Finally, in an analysis of data from the New Jersey experiment, Mallar (1977) found that teenagers whose parents were enrolled in NIT were 20 percent to 90 percent more likely to complete high school depending on the NIT plan. However, Venti (1984) found only an 8 percent increase in the probability of completing high school for youth in the Seattle-Denver experiment. This lower estimate seems more probable in view of the relatively short duration of the experiments and the many long-term factors (such as achievement in early grades) that have been linked to educational attainment. These results may also be related to the fact that, in all four experiments, youths in treatment households were less likely to be employed than controls (Robins, 1985).

These studies suggest that the relatively large income transfers made to families under the NIT had a positive effect on the nutritional status and educational attainment of children in the poorest families. However, the magnitudes vary greatly from study to study. Perhaps unsurprisingly, studies of the effects of the NIT on consumption also show that families spent much of the subsidy on goods that may not have been directly related to the well-being of their children. For example, the NIT appears to have had a negative effect on the labor supply of married women, and positive effects on housing expenditures and purchases of consumer durables (Robins, 1985; Michael, 1977).²

### WHAT WE KNOW ABOUT IN-KIND PROGRAMS

A parallel ‘in-kind’ welfare system has grown up alongside the cash system. This system aims to directly provide for a child’s “basic needs” — decent housing, food, medical care, and quality early education. Table 7.1 shows that expenditures on virtually all of these programs have shown steady growth over time (the exception being the School Lunch Program). Table 7.2 indicates that in contrast to stagnant AFDC caseloads, caseloads for most in-kind programs have been increasing.

³No convincing evidence of a link between maternal employment and children’s well-being has been found. See Blau and Gossberg (1980) and Desai et al. (1989).

⁴The NIT may also have increased the probability of marital dissolution, although this finding remains controversial (e.g., Can and Wissoker, 1980; Hanum and Tuna, 1990).

<table>
<thead>
<tr>
<th>TABLE 7-2</th>
<th>Trends in Caseloads (millions)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cash Transfers</strong></td>
<td></td>
</tr>
<tr>
<td>AFDC</td>
<td></td>
</tr>
<tr>
<td>Total recipients</td>
<td>11.1</td>
</tr>
<tr>
<td>Child recipients</td>
<td>7.8</td>
</tr>
<tr>
<td>Earned Income Tax Credit</td>
<td></td>
</tr>
<tr>
<td>No. of families</td>
<td>6.2</td>
</tr>
<tr>
<td><strong>In-Kind Transfers</strong></td>
<td></td>
</tr>
<tr>
<td>Housing Assistance</td>
<td></td>
</tr>
<tr>
<td>No. of households</td>
<td>3.2</td>
</tr>
<tr>
<td>Food Stamps</td>
<td></td>
</tr>
<tr>
<td>Total recipients</td>
<td>16.3</td>
</tr>
<tr>
<td>WIC</td>
<td></td>
</tr>
<tr>
<td>No. of women</td>
<td>0.2</td>
</tr>
<tr>
<td>No. of infants</td>
<td>0.2</td>
</tr>
<tr>
<td>No. of children</td>
<td>0.3</td>
</tr>
<tr>
<td>School nutrition</td>
<td></td>
</tr>
<tr>
<td>School lunch</td>
<td></td>
</tr>
<tr>
<td>No. of meals</td>
<td>26.3</td>
</tr>
<tr>
<td>No. of meals</td>
<td>10.3</td>
</tr>
<tr>
<td>School Breakfast</td>
<td>2.5</td>
</tr>
<tr>
<td>No. of meals</td>
<td>2.0</td>
</tr>
<tr>
<td>Medicaid</td>
<td></td>
</tr>
<tr>
<td>Total recipients</td>
<td>22.0</td>
</tr>
<tr>
<td>Child recipients</td>
<td>9.6</td>
</tr>
<tr>
<td>Head Start</td>
<td></td>
</tr>
<tr>
<td>No. of meals</td>
<td>0.3</td>
</tr>
</tbody>
</table>

⁵These figures are for 1977.


Initial evaluation of these in-kind programs is more straightforward than the evaluation of cash transfer programs because we can ask whether the program has an impact on the specific child outcome it was designed to affect. For example, we can ask whether receipt of housing assistance is associated with improvements in housing or whether household participation in the Food Stamps program improves a child’s diet.

We might then wish to ask whether the program has additional effects on related child outcomes. For example, better nutrition could influence a child’s cognitive abilities. Also, subsidies to food and housing may influence child outcomes more generally by relaxing the family’s budget constraint (see Moffitt, 1989, and Citro and Michael, 1995, for discussions of the valuation of in-kind benefits).⁶ However, since the effects of income transfers are discussed above,

⁶The National Research Council (Citro and Michael, 1995) concludes that for simplicity’s sake, “near cash” benefits such as Food Stamps and housing assistance should be counted at their dollar...
the focus in this section is on any effects of participation in in-kind programs on
the specific outcomes that the programs were designed to affect. In practice, this
restriction eliminates very few studies from consideration.6

Housing Assistance

In contrast to AFDC and Food Stamps, housing assistance is not an entitle-
ment: when funds allocated to the program run out, people who are eligible must
be wait-listed. It is estimated that about half of federal expenditures on housing
assistance directly benefit children, while the elderly are the other large group of
beneficiaries.

Most expenditures are on rental assistance programs rather than on low-rent
public housing (which is what many people think of as ‘public housing’). And
since 1982, most new authorizations for rental housing assistance have been for
Section 8 programs (Pedone, 1988). The Section 8 existing housing program
provides rent subsidies to families who find an apartment of their own choosing,
as long as the rent is below the “Fair Market Rent” established by the Depart-
ment of Housing and Urban Development (HUD) and the unit meets minimum
quality standards. Rental assistance typically reduces a family’s rental pay-
ments to 30 percent of its income, after deductions for certain expenses are taken
into account.

Deficient housing is hazardous to children. For example, lead poisoning is
three times more common among poor children than among nonpoor children
and is directly related to housing conditions. The risk of accidental death is also
three times higher for poor children, and some of this increased risk may be due
to hazards in the home (Starfield, 1985). In 1989, 16 percent of poor households
(2.2 million households) lived in housing with severe or moderate physical prob-
lems compared to 7 percent of nonpoor households.1

It is not known whether, in general, housing assistance enables families in
deficient housing to move to adequate housing. A 1988 HUD study found that
more than half of public housing households lived in projects that needed moder-
ate to substantial rehabilitation just to meet HUD’s own standards. The estimated
cost of bringing these units up to standard would have exceeded $20 billion 1986
dollars (Lazere et al., 1991).

Section 8 programs require families to locate a landlord willing to participate
and to arrange with the landlord for inspections and repairs within a fixed period
of time. One case study of 56 single mothers in eastern Massachusetts in 1985
and 1986 found that after waiting an average of 2 years to receive a certificate, 24
women returned them unused because they were unable to find housing that met
program requirements within the allotted time (Mutzro, 1988). On the other
hand, there is some evidence that recipients of vouchers pay higher rent (Kennedy
and Finkel, 1987; Appar, 1990) and move to better neighborhoods (Johnson,
1986). The often dismal social conditions in many public housing projects must
be weighed against any improvements in the physical housing stock. However, it
is very difficult to identify the effects of neighborhoods and schools because any
relationship we observe between neighborhood characteristics and individual
outcomes could reflect the characteristics of the individual or of his or her family
that placed them in these neighborhoods in the first place.

The Gautreaux program sheds light on this issue. Under the program, resi-
dents in public housing projects can apply for Section 8 housing certificates and
move to private apartments. Some apartments are in predominantly white sub-
urbs, while others are in the inner city. Although the persons admitted to the
program are not a random sample of public housing residents,2 Rosenbaum
(Rosenbaum et al., 1986; Rosenbaum, 1992) asserts that the program assigns
apartments in an approximately random manner, since people get whatever is
available when they reach the top of the waiting list. He finds that 7 years after
their move, children who had moved to the suburbs were 15 percent less likely to
have dropped out of school, 16 percent more likely to be in a college-track
program, and 34 percent more likely to be employed than those who had moved
within the inner city. All of these differences are statistically significant at the 90
percent level of confidence.

These findings suggest that voucher programs can have a positive effect on
the life chances of children if they enable families to find housing in better
neighborhoods. On the other hand, they suggest that the disamenities associated
with large public housing projects may have significant negative effects. How-
ever, the study is marred by high rates of attrition from the sample. HUD is
currently conducting an experimental evaluation of a program similar to Gau-
treaux in four cities.3 An experimental evaluation that took care to minimize
attrition could shed light on the possible beneficial effects of housing vouchers,
and on the issue of the effects of neighborhoods more generally.

Despite their bad reputations, housing projects may be better than much of

1Applicants are screened to make sure that they have paid their rent regularly and that they have
adequate housekeeping abilities. The program does not serve families with more than four children
because few large housing units are available in the suburbs. In addition, the act of applying for an
apartment in an unknown location may indicate that a person is strongly motivated to improve his or
her circumstances.
2Personal communication, Lawrence Katz, Department of Economics, Harvard University, 1997.
THE EFFECT OF WELFARE ON CHILD OUTCOMES

188

JANET CURRIE

189

more likely to spend a given amount of income on food than men, and the female
head of household may have more control of Food Stamp coupons (which are
likely to be issued in her name) than she has over the household's cash income.11
Neither theory has been subjected to an empirical test.

Supplemental Feeding Program for Women, Infants, and Children

In addition to the Food Stamp Program, the federal government offers sev-
eral feeding programs that give food directly to needy children and their mothers.
The WIC program provides nutritional counseling and food supplements to preg-
nant and lactating mothers and their infants as well as to low-income children up
to age 5. All participants must be certified to be nutritionally "at risk." WIC is
funded by appropriation, and the size of each year's appropriation limits the
number of people that can be served. WIC is currently operated out of some
8,330 sites and serves approximately 60 percent of those eligible (Jones, 1992).
The law requires that the WIC program provide foods containing protein, iron,
calcium, vitamin A, and vitamin C. Food packages must be appropriately tai-
toled to meet the needs of each category of recipient.12 In fiscal year 1991, the
average monthly WIC package was valued at $31.67.

Many studies find that WIC has positive effects on the utilization of prena-
tal care and on measures of infant health including birthweight, the incidence
of low birthweight, gestational age, and infant mortality.13 Schramm (1985) and
Devaney et al. (1990) examine the effects of WIC on the Medicaid costs of
newborns. The results are of particular interest because they can be used to
compare the costs and benefits of the WIC program. Schramm found that in
1980, a dollar spent on WIC reduced Medicaid costs in Missouri by approxi-
mately 80 cents in the first 30 to 45 days after birth. Devaney et al. examined
Medicaid costs in the first 60 days after birth in five states and found that
reductions in Medicaid costs over this period more than offset the costs of
providing WIC.

Unfortunately, only two WIC studies, by Metcalf et al. (1985) and Caan et
al. (1987), have used random assignment to generate a comparison group. If
WIC participants are worse off than nonparticipants because places are scarce
and only the neediest are admitted into the program, then studies that compare

11Some circumstantial evidence pertinent to this hypothesis comes from the Washington State
Welfare Reform Demonstration Program. AFDC recipients in demonstration counties had the option
of choosing to receive their AFDC and Food Stamp benefits in the form of a single consolidated
check rather than continuing to receive Food Stamp coupons. Over 20 percent of these women opted
to continue receiving the coupons.
12The categories are children 0 to 3 months of age, children 4 to 12 months, women and children
with special dietary needs, children from 1 to 5, pregnant and nursing mothers, and postpartum
nursing mothers.
13See Devaney et al. (1990) for a review.

The housing available to poor families who do not have access to voucher pro-
grams. By combining data from the 1990 Census and the Current Population
Surveys, Currie and Yelowitz (1997) are able to examine the effects of residence
in public housing projects on housing quality as measured by the extent of
overcrowding and the density of the housing complex. They also examine the
effect on the probability that a child has been retained in grade, an important
index of educational attainment. They find significant positive effects on all
three outcomes.

The Food Stamp Program

Food Stamps are issued in the form of booklets of coupons that may be used
to purchase all foods except alcohol, tobacco, and hot foods "intended for imme-
diate consumption." In contrast to AFDC, Food Stamps are available to all
families who meet federally determined income-eligibility requirements, though
AFDC recipients are automatically eligible.

The value of a family's Food Stamp allocation is typically much less than
what the family spends on food. Hence, it is likely that the increase in the family's
food expenditures will be less than the value of the Food Stamps because
families can spend the same amount on food that they would have in the absence
of the program and use the "freeed-up" money for something else. In fact, eco-

 experimental studies of this issue have proved inconclusive. See Fraker (1980a, 1980b) and
Korenman and Miller (1992) for examples, and Currie (1995a) for a discussion.
WIC participants and nonparticipants will underestimate the effects of the program. Conversely, if WIC participants are more highly motivated or better informed than nonparticipants, then studies of this type may overestimate the program’s effects. Without knowing more about the selection mechanism underlying participation, it is difficult to assess the probable direction of this bias.

Still, given that the program is locally administered, the factors governing selection into the WIC program are likely to differ considerably over time and across sites. Hence, the fact that estimated effects are remarkably constant across states and over time suggests that the positive results are not entirely driven by selection. This conclusion is reinforced by a recent study by Brien and Swann (1997) who use both instrumental variables and sibling comparison methods to analyze data from the 1988 National Maternal and Infant Health Survey. They find significant effects of WIC on birth outcomes and on maternal behaviors (such as reductions in drinking while pregnant) among blacks, but they are unable to detect any effect among whites.

Studies of the effects of WIC on the nutrient intakes of children generally find positive effects (cf. Fraker, 1990a), but these studies are also plagued by possible selection bias. One way to control for bias is to follow the same child over time. The Centers for Disease Control reported the results of a study that followed child WIC participants in six states over a 2-year interval (United States Department of Health and Human Services, 1978). The study found that after three WIC visits the percentage of children who were anemic fell by more than half. In addition, the fraction of 6- to 23-month-old children below the tenth percentile of length-for-age fell from 21 percent to 15 percent after three WIC visits.

Hicks et al. (1982) focus on 21 pairs of siblings from rural Louisiana. Because of the design of the WIC program in that state, the younger child in each pair was eligible for supplementation beginning in the third trimester of pregnancy, while the older child became eligible for WIC only after the first year. The results show that the “early supplementation” group had significantly higher scores on a range of cognitive tests.

**School Nutrition Programs**

The federal government supports six other programs that provide meals or monthly food supplements to low-income children. The largest is the National School Lunch Program (NSLP). The NSLP is an entitlement that operates by reimbursing schools for each meal served. School lunches are provided free to children with family incomes less than 130 percent of the federal poverty line and are subsidized if the family income falls between 130 and 185 percent of the poverty line. In 1991, lunches were served to approximately 12.8 million students, and 10.3 million students received free lunches. The School Breakfast Program (SBP) serves fewer, typically needier, students.

The effects of school nutrition programs are controversial. Older studies found that participants had higher 24-hour nutrient intakes than nonparticipants and that SBP participants were more likely to eat breakfast than nonparticipants (Hanes et al., 1984). However, more recent studies show higher intakes of some nutrients, but also higher intakes of fat and cholesterol (Gordon et al., 1995). Surprisingly, there have been few attempts to evaluate the effects of school nutrition programs on cognitive outcomes. In one of the more compelling studies, Meyers et al. (1989) examined 1,092 third to sixth grade children in Lawrence, Massachusetts, before and after the SBP was introduced at their school in 1987. They found that the Breakfast Program participants showed greater improvements on the Comprehensive Test of Basic Skills, relative to their initial scores, than nonparticipant children. SBP participation also reduced tardiness.

**Medicaid**

Medicaid is the main system of public health insurance for poor women and children. It is a federal-state matching entitlement program, administered at the state level. Table 7-1 shows that expenditures on children account for a relatively small share of total Medicaid expenditures. The average expenditure on an AFDC child is $891 (1992 dollars) compared to $3,778 for an aged person (U.S. House of Representatives, 1994). Still, both expenditures and caseloads continue to grow as shown in Tables 7-1 and 7-2.

States were required to offer Medicaid coverage to AFDC recipients, and until recent extensions of coverage to other groups, there was a very close link between AFDC recipiency and Medicaid eligibility. However, evidence that many children and pregnant women were not receiving adequate preventive care led Congress to expand Medicaid coverage for pregnant women and children beginning in 1984. States are now required to cover all pregnant women and children under 6 with family incomes less than 133 percent of the federal poverty line, regardless of family structure.14 Beginning on July 1, 1991, states have been required to cover all children born after September 30, 1983, whose family incomes are less than 100 percent of the federal poverty line.

The recent 1997 Budget Reconciliation Act allocates $47 billion over the next 10 years to allow states to expand health insurance coverage to an even larger group of uninsured children, either through the Medicaid program or through separate state initiatives. States must contribute 70 percent of what the state would have contributed under the matching provisions of the Medicaid program—that is, states can get federal money to expand health insurance coverage at a very favorable match rate. These new provisions make it more pressing than ever to determine the effects of public health insurance on children.

Currie and Thomas (1995a) use panel data that follow the same child over time and show that, when children are covered by Medicaid, they are more likely

---

14 The coverage of pregnant women is limited to services related to the pregnancy.
to have had our doctor visits in the past 6 months. Moreover, the effect of being covered by Medicaid is larger than the effect of being covered by private health insurance, which probably reflects the fact that Medicaid has no copayments or deductibles. This effect is the same for black and white children. However, white children also receive more visits for illness when they are covered by Medicaid than when they are uninsured, and this is not true for African Americans. Thus, equivalent insurance coverage does not guarantee equal care.

Currie and Gruber (1996a) look at the effect of becoming eligible for Medicaid on the utilization of medical care and on child health. The effects of Medicaid eligibility are identified using the recent federally mandated expansions of the Medicaid program to pregnant women and children described above. They find that expansions of eligibility to pregnant women increased the fraction of women eligible from 12 to 43 percent. This increase was associated with an 8.5 percent decline in the infant mortality rate.

However, earlier extensions of Medicaid eligibility to very poor women who were already income-eligible for AFDC were much more cost-effective than later expansions to higher-income women. The reason is that higher-income women were less likely to become covered early in their pregnancies. Hence, they did not avail themselves of free preventive prenatal care available under the Medicaid program. There is evidence, however, that hospitals enrolled eligible women in Medicaid at delivery so that costly services received by unhealthy newborns were paid for by the program. These results suggest that outreach programs designed to improve take-up could increase the cost-effectiveness of the Medicaid extensions to pregnant women.

Currie and Gruber (1996b) use the same methodology to look at the effects of extending eligibility to additional groups of low-income children. They find that, although many newly eligible children did not take up coverage, becoming eligible for Medicaid reduced the probability that a child went without a doctor's visit in the past year and also improved the quality of care as measured by the fraction of these visits that took place in doctor’s offices rather than hospital outpatient clinics or emergency rooms. These changes were linked to significant reductions in child mortality from internal causes and had no effect on mortality from external causes (e.g., accidents). This is the pattern one might anticipate if the changes in mortality were linked to increases in the use of preventive care.

The complex relationship between formal take-up and benefits received is further explored by Currie (1995b) in a study that focuses on differences between children of immigrants and children of the native born. She shows that recent expansions of Medicaid eligibility had smaller effects on Medicaid coverage among immigrant children than increased utilization of basic services by at least as much among immigrants as among nonimmigrants.

The differences in patterns of take-up and utilization by race and natality are consistent with evidence from other countries that extensions of insurance coverage alone will not eliminate socioeconomic differences in health care utilization or health (Currie, 1995c). It is unlikely that lack of information alone can explain the differences, since black and immigrant parents are as likely as other parents to bring their children in for free preventive care when they become eligible for Medicaid. Similarly, purely cultural explanations that posit that some groups value medical care less than others are difficult to reconcile with this evidence.

Disparities in the availability of private health insurance, in the transaction costs associated with enrolling in the Medicaid program, or in access to providers willing to accept Medicaid payments may all be important determinants of group differences. Currie et al. (1995) examine the last of these three factors and show, using state-level data, that increases in Medicaid fee ratios for obstetricians/gynecologists are associated with significant declines in infant mortality, presumably because of increases in either effective physician supply or the quality of services provided.

The fraction of children with private health insurance fell over the period of the Medicaid expansions to such an extent that there was actually a small decrease in the fraction of children with any health insurance coverage. These trends lead one to suspect that public health insurance may have “crowded out” private insurance coverage. Cutler and Gruber (1996) estimate that as many as 50 percent of the people who became covered by the Medicaid expansions may previously have had private health insurance. While a switch from private to public insurance does not raise the fraction of children covered, the Currie and Thomas (1995a) results suggest that it may still improve the health of children by encouraging the utilization of preventive care.

Other analysts (see Dubay and Kenney, 1997) point out that private health insurance coverage was declining even among groups such as single men whom one would not expect to be affected by the expansions. If one asks what fraction of the total decline in private health insurance coverage is a result of substitution towards Medicaid, the answer is approximately 15 percent. Clearly, much research remains to be done on the causes and consequences of the decline in private health insurance coverage.

**Head Start**

Head Start is a federal-local matching grant program that aims to improve the skills of poor preschoolers so that they can begin schooling on a more equal footing with their more advantaged peers. Unlike Medicaid, it is not an entitlement program, and only about a third of eligible children are served (Stewart, 1992). Head Start has enjoyed widespread bipartisan support over a long period, although evidence regarding long-term effects is inconclusive. Experimental studies that focus primarily on inner-city African-American children typically find an initial positive effect on children’s cognitive achievement that fades out in 2 or 3 years.

Supporters of the program argue that a narrow focus on cognitive test scores
is inappropriate, given that Head Start is intended to affect a range of outcomes (see McKeczy et al., 1985). Evidence from the Perry Preschool Project, which found that program children were less likely to drop out of high school, engage in crime, or become pregnant as teenagers, is often cited. However, since the project included only 58 treatments and 65 controls, was funded at about twice the rate of a typical Head Start program, and did not involve a national sample, it is not clear that the findings generalize.

Currie and Thomas (1995b) examine sibling comparisons from a national sample and find that children who were in Head Start have higher test scores at the end of the program than either stay-at-home siblings or siblings who went to other preschools. The effects are of the same magnitude for both black and white children and indicate that Head Start closes one-third of the gap between these children and others. But consistent with the experimental studies, they find that the effects on black children fade out rapidly. These results suggest that the positive effects of Head Start may be undermined by subsequent deprivation among these children.

In contrast, the effects on the test scores of white children do not fade out. Moreover, white children 10 and over are significantly less likely to have repeated a grade if they attended Head Start and are thus less likely to have experienced the age/grade delay that often leads to high school noncompletion. Both black and white children who attended Head Start were more likely to be inculcated with high school standards, although there was no effect on height-for-age, a measure of long-term nutritional status.

In related work, Currie and Thomas (1996a) find that Head Start has large and lasting effects on the test scores of Latino students. A closer inspection of the data reveals that these positive effects are largest for Mexican-origin children and smallest for Puerto Rican children. However, due to sample size limitations it is not possible to use the data to estimate the effects of ethnicity and the effects of region. It is possible, for example, that the ethnic differences reflect differences in the ways programs are provided, such as where Puerto Rican children tend to be located, and California and Texas, where Mexican-origin children are concentrated, rather than any independent effect of ethnicity per se.

Currie and Thomas (1996b) ask whether differences in school quality can explain differences in the pattern of "fadeout" in test scores between whites and blacks. Specifically, the initial positive effects of the Head Start program may be undermined if Head Start children were subsequently exposed to inferior schools. And since we see fadeout for blacks but not for whites, it would have to be the case that black Head Start children are attending worse schools than other black children but that the same was not true among whites.

Currie and Thomas test this hypothesis using a sample of eighth graders from the National Educational Longitudinal Study of 1988 (NELS). Their work builds on earlier research by Lee and Loeb (1995) who showed, using these data, that the schools attended by Head Start children are of worse quality in some observable dimensions than the schools attended by other children. Even if family income and parent's education are controlled for, children who attended Head Start have lower test scores than other children. This result is to be expected if Head Start does not entirely compensate for early disadvantages.

However, among black children, the gap between Head Start children and other children is virtually eliminated when we compare children within the same school. That is, within schools, black Head Start children do no worse than other black children. But since they perform more poorly than other children on average, they must be attending schools in which all black children do badly. If a "quality" school is defined as one in which children do well, then these results suggest that black children who attend Head Start go on to attend schools of significantly worse quality than other black children. In contrast, among non-Hispanic white children there appears to be little difference in the schools attended by Head Start and other children.

WHAT WE NEED TO KNOW

The preceding discussion is summarized in Table 7.3. The table presents a matrix of programs and effects. Differences in the effects of programs across groups have been suppressed, although one theme that has emerged from the discussion so far is that they are important. The most striking feature of Table 7.3 is that there are many empty cells—we clearly need to learn a great deal more about the effects of welfare before we can make informed public policy. In some cases, research has been limited by lack of appropriate data. In others, existing information has not yet been fully exploited. This section highlights some unanswered research questions and discusses the extent to which data collection efforts could help.

Effects of Welfare on Long-Run Outcomes

Ultimately, what many people care about is whether investments in children today will produce productive, well-socialized adults tomorrow. However, Table 7.3 highlights the fact that little is known about the effects of welfare on long-term outcomes. Lack of data places major limitations on this type of research. Many important outcomes can only be examined 10 to 15 years after childhood participation in welfare programs. There are few existing datasets that combine information about childhood participation in welfare, other family background characteristics, and the outcomes of interest.

One exception is the National Longitudinal Survey's Child-Mother file (NLSMC). The NLSMC contains information about the children of a sample of approximately 6,300 women who were between the ages of 14 and 21 in 1978. Information about childhood participation in AFDC, the Food Stamp Program, Medicaid, Head Start, and WIC is available. By the time the 1994 wave is
released, there will be more than 800 children over 16. Of course, since these children will have been born to young mothers, they will not be a nationally representative sample of 16 year olds. Still, this sample is a valuable resource. If future waves of the survey continue to be funded, it will grow in size and in representativeness and allow us to address many questions about the relationship between welfare and long-term outcomes such as schooling attainment, teen parenthood, and crime.

A second exception is a special supplement to the Panel Study of Income Dynamic (PSID) that was fielded in 1995. This module contains retrospective information about early childhood education and criminal activity that can be linked to data about welfare participation from the original PSID file. The PSID is currently undertaking an even more ambitious data collection effort, the 1997 Child Development Supplement. The survey of 3,500 0 to 12-year-old children will have assessments of cognitive, behavioral, and health status. Data are being collected from the mother, a second caregiver, the absent parent (if relevant), teachers, school administrators, and the children themselves. The survey will also include time diaries for caregivers, children, and teachers, to examine inputs into child development. Finally, other inputs such as resources in the home and neighborhood will also be measured. Once again, this information can be linked to data about welfare participation from the main files, and follow-up on these children may help to identify long-term effects of participation. Fielding this type of supplement to existing data sources promises to be a cost-effective method of providing information on the link between the current outcomes of young adults and their participation in various programs as children.

An additional issue that can be addressed is whether there are links between the short-term outcomes that have been examined in previous research and longer-term outcomes. If it is found that particular short-term outcomes are reliable “markers” for longer-term outcomes, then future evaluations of welfare programs may not require as much costly long-term follow-up of the participants.

**Why Do Effects Appear to Vary with Race, Ethnicity, and Natality?**

The PSID and NLSCM datasets will both support analyses stratified by race, ethnicity, and natality. However, in many cases the sample sizes are very small. In order to properly document differences in outcomes, or even in utilization, it will be necessary to add questions to existing large-scale datasets. For example, the Census asks questions only about the use of cash welfare, even though expenditures on in-kind programs constitute the largest and fastest-growing share of the welfare bill.

A second problem is that large-scale, individual-level datasets typically lack information about neighborhoods and administrative procedures that could be used to test specific hypotheses about group differences. For example, one might believe that black children on Medicaid receive fewer visits for illness than white
children because the providers that serve them are overcrowded and it is more
difficult to get additional appointments. It would be very useful to know the
extent to which group differences are associated with the administration of wel-
fare programs, rather than with differences in parental tastes or circumstances.

It is unlikely that many detailed questions of this type will be added to large-
scale surveys, but it would be possible to match data from other sources to the
surveys if finer geographical information were made available to researchers.
While issues of confidentiality are important, the amount of information that
could be gained if it were routinely possible to match survey data to, say, zip-
code-level data from other sources can hardly be underestimated.

This type of matching is also greatly facilitated by the existence of a central
agency that collects program information (and is willing to give it to researchers).
There is a real danger that further devolution of responsibility for welfare to the
states will result in a loss of information about the administration of programs,
making it more difficult to identify program effects using state-level variation in
the programs.

How Do Programs Interact?

One glaring omission from this survey is that there has been no discussion of
multiple program participation. Many children are covered by more than one
program. For example, AFDC participants are covered by Medicaid and are
automatically eligible for Food Stamps. As of 1980, half of AFDC children
received free school lunches, 35 percent lived in public or subsidized rental
housing, and 19 percent participated in WIC. Conversely, half of all Food Stamp
recipients, 42 percent of Medicaid recipients, 38 percent of WIC recipients, and
24 percent of those in public housing also received AFDC. Moffitt (1992) esti-
mates that in 1984, 26.4 percent of nonelderly single-parent families received
AFDC, Medicaid, and Food Stamps, and 11 percent received at least one benefit
in addition to AFDC.

It is impossible to say how multiple program participation affects the child
outcomes discussed above since there has been little research on this topic.
Some programs may be duplicative, while others may interact to produce more
positive outcomes. For example, Currie and Thomas (1995b) found that chil-
dren in Head Start were more likely to be immunized than other children, even
though many Head Start children would have been eligible for free vaccinations
under the Medicaid program in any case. Head Start may help families to enroll
in Medicaid, may help them locate a Medicaid provider, or may bypass Medi-
care altogether by arranging for children to be immunized at the Head Start
center.

An analysis of multiple program participation would assist us in answering
the question of whether the current patchwork system of programs is an efficient
way to provide welfare. The proliferation of programs increases possibilities for

fraud, waste, and mismanagement. On the other hand, the evidence surveyed
here suggests that targeting specific benefits directly to individual children has
advantages in terms of ensuring that specific benefits are received. We need to
know more about the balance between these benefits and costs.

How Do Successful Programs Work?

Data limitations place severe restrictions on our ability to look inside the
"black box" of welfare programs. For example, we can show that expansions in
Medicaid eligibility have been related to reductions in child mortality rates at the
state level, but we do not know why. It could be due either to increased use of
preventive care or to more intensive palliative care for sick children. The two
possibilities have quite different implications for child well-being as well as for
efficiency and program costs. Better information about what goes on during doctor
visits and about objective measures of child health status (short of mortality sta-
tistics) could help us to address this question. It might be possible, for example, to add
questions about anemia, lead poisoning, and anthropometrics (e.g., height-for-age,
weight-for-height) to the next National Health Interview Survey.

Still, the most likely scenario is one in which we chip away at these questions
using an interactive, multidisciplinary approach: analysis of large-scale surveys
can be used to develop broad hypotheses, which can then be tested using case
studies. The case studies can then be used to develop more precise hypotheses
about the survey data and to suggest supplemental survey questions.

Cost-Effectiveness

Evidently, if a program has no effect at all on a desired outcome, then it
cannot be considered cost-effective. Many of the programs discussed above
have passed this initial test—they can be shown to have positive effects. The
question remains however, of whether they are cost-effective, that is, whether
the benefits outweigh the costs. The figures discussed above for WIC are quite
impressive in this regard. Cost-effectiveness studies exist for other small-scale
early intervention programs (not reviewed here) but have not generally been
conducted for large-scale federal programs. Although it is unlikely that there
will be agreement on all of the costs and benefits that should be included in such
an analysis, some rough calculations under varying assumptions would no doubt
be useful to policy makers.

CONCLUSIONS

This survey chapter discusses eight large federal welfare programs that af-
fect children. The available evidence is incomplete but suggests a consistent
story: programs that target services directly to children have the largest measured
effects, while unrestricted cash transfer programs have the smallest, perhaps because their benefits are more diffuse or because the amounts of money involved are typically quite small.

There are also striking and largely unexplained differences in the effects of some programs by race, ethnicity, and or natality. These differences could reflect nonlinearities in the effects of programs—that is, one might expect larger effects for poorer or for richer children, and children from some groups are more likely to be poor. Alternatively they may reflect differences in the programs available to children of different origins or unobserved differences between participants from different groups that have not been adequately accounted for.

This survey concludes with five questions for future research: (1) Do welfare programs have long-term effects on children? (2) Why do programs have differential effects by race, ethnicity, and natality? (3) How do programs interact? (4) How exactly do successful programs work? (5) Are programs cost-effective? These questions indicate that though we know much more than we did even 5 years ago about the effects of welfare on children, there is still much work to be done if we are to make informed decisions about public policy.

ACKNOWLEDGMENTS

The author is grateful to Lindsay Chase-Lansdale, Greg Duncan, Bentley MacLeod, and Robert Moffitt for helpful comments. Support from the Alfred P. Sloan Foundation, the National Science Foundation under grant SBR-9512670, and the National Institute of Child Health and Human Development under grant HD-31722-O1A2 is gratefully acknowledged. The author is solely responsible for the opinions expressed.

REFERENCES


JANET CURRIE

Koerner, S., and J. Miller

Lazere, E., P. Leonard, C. Dohlebar, and B. Zagas

Lee, V., and S. Leeb

Mallia, C.

Mayer, S.

Maynard, R., and D. Crawford

Maynard, R., and R. Murnane

McKey, R., L. Condell, H. Gann, B. Barrett, C. McConkey, and M. Planz

Metcalf, J., D. Rubin, M. Napoleone, and K. Nichols

Meyers, A., A. Sampson, M. Wettzman, and H. Kayne

Meyers, A., D. Rubin, M. Napoleone, and K. Nichols

Michael, R.

Mitchell J., and R. Sharman
1984 Access to private obstetrics/gynecology services under Medicaid. Medical Care 22(November):1026-1037.

Moffitt, R.


Moore, K., and S. Caldwell

Mulroy, E.


