The Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) has enjoyed extraordinary bipartisan support stemming from the widespread belief that research studies have proven that WIC "works." Although some studies suggest real dietary and health improvements, the greatest benefits only apply to WIC's prenatal program (just a small part of the total program). Even here, weaknesses in the research render the findings highly uncertain. The three most significant weaknesses are (1)selection bias, (2) simultaneity bias, and (3) lack of generalizeability. The resulting uncertainty places WIC's possible impacts on infant mortality, prematurity, and birthweight on a range from zero *to* substantial. *For infants, children, and postpartum and breastfeeding mothers, the only impacts seem to be small to modest effects on anemia and nutrient intake. This paper does not argue that WIC's weaknesses justify abandoning or even cutting the program. On the contrary, there should be a sustained effort to make the program more effective. This effort should start with a policy debate about WIC's role and impacts, coupled with a grant of greater flexibility to state and local WIC agencies to open the program to innovation and experimentation. To increase WIC's positive impacts, we propose a series of possible reforms, each to be thoroughly evaluated.*

EVALUATING WIC

DOUGLAS J. BESHAROV
PETER GERMANIS
University of Maryland

1. INTRODUCTION

"WIC works, perhaps better than any other government program in existence," Agriculture Secretary Dan Glickman (1997) declared. Former Health and Human Services Secretary Louis Sullivan (1995) made a similar claim: "The WIC Program results in significant Medicaid savings that far outweigh the program's costs by a ratio of 3 to 1. . . . That is clearly an overwhelming return on a small national investment." Such statements testify to the extraordinary bipartisan support enjoyed by the "Special Supplemental Nutrition Program for Women, Infants, and Children," known as WIC to reflect the initials of its three target populations.

WIC was established in 1972 as a 2-year pilot program, partially in response to the 1969 White House Conference on Food, Nutrition, and Health. The conference report concluded that nutritional deficiencies among low-income

EVALUATION REVIEW, Vol. 24 No. 2, April 2000 123-190

women and children threatened their health and led to higher medical costs. To avoid preventable physical or medical conditions, WIC seeks to improve the diets and, therefore the health, of low-income pregnant, breastfeeding, and postpartum women, and their infants and children (up to age 5).

WIC is a $5 billion program, serving about 7.4 million women and children. Although WIC is a program of the United States Department of Agriculture (USDA), most of its grantees are state health departments. These state agencies, in turn, fund WIC services through local health-related agencies such as health departments, hospitals, public health clinics, and community health centers.

WIC's popularity stems from the widespread belief that research studies have proven that WIC works. But although some studies suggest real improvements in the diets and the health of recipients, the large benefits cited above relate only to research conducted on WIC's prenatal program (which comprises only about 11% of program participants). Even here, the evidence suggests that WIC's benefits are, at best, modest.

Some will argue that this exaggeration is for a good cause: $5 billion a year in additional food, nutritional education, and counseling services for low-income infants and children, pregnant women, and breastfeeding or postpartum mothers. But overstating WIC's effectiveness undermines support for the research and programmatic flexibility needed to increase its beneficial impact. For example, over the past decade, additions to WIC's funding have had the effect of expanding the program into the lower middle class, when they probably could have been much more effectively used to improve or intensify services for more needful families (a point that we return to later in this article). WIC's rigid spending rules, for example, effectively prevent local programs from spending more than about 30 minutes for nutritional education every 6 months with clients and preclude enriching food packages with, for example, iron supplements.

This need to improve the WIC program is extremely important. Even if WIC were as effective as its advocates claim, much more needs to be done to improve diet-related health outcomes for low-income Americans. In 1998, America's infant mortality rate was 0.7% of live births, or about 28,000 babies. The low birthweight rate was about 7.6% of live births (up 12% since 1986), or about 300,000 babies (Martin et al. 1999). Surely, we should strive for a WIC program that is as effective as possible.

Moreover, WIC was developed almost 30 years ago, when hunger was the major nutrition-related problem facing disadvantaged Americans. Since then, hunger has been superseded by overweightness as our most serious nutrition-related health problem. It is time to consider updating WIC's mission by adding a specific focus on preventing overweightness.

Parts 2 and 3 of this article describe the WIC program; Parts 4, 5, and 6 review and assess the research on its impact; Part 7 recommends the state-based experimentation along the lines of the policy reforms listed above; and Part 8 calls for a series of randomized experiments to evaluate the program and any changes made to it. The essential descriptive points made in this article are found in the appendix, titled "WIC at a Glance."

2. PROGRAM BENEFITS

WIC has three parts: (a) vouchers to purchase specific high-nutrition food packages to supplement diets, (b) nutritional and health counseling, and (c) referrals to health care and social-service providers.

FOOD PACKAGES

After applicants are approved for participation, they receive WIC vouchers that are valid for 1 to 3 months, after which the vouchers must be reissued. Like food stamps, vouchers can be redeemed only at participating food stores. But unlike food stamps, which enable recipients to purchase a wide array of foods, WIC vouchers can be exchanged only for specified foods in the WIC "food package." (They cannot even be used for vitamins or nonfood nutritional supplements.)

The amount of WIC benefits is not related to household size or income, as are food stamps and many other social programs; rather, it is related to the age and pregnancy status of recipients. Thus, recipients receive one of seven basic food packages, depending on their recipient category and nutritional need.[1]

WIC's food "packages" are meant to supplement the diets of recipients rather than to meet their entire food needs. High in protein, calcium, iron, and vitamins A and C, they are designed to provide the nutrients often lacking in the diets of WIC's target populations. Packages typically include iron-fortified infant cereal and formula, juice, milk, cheese, eggs, peanut butter, and beans. Many WIC agencies tailor food packages to meet the nutritional deficiencies of individual WIC clients. For example, if it is known that a WIC participant has high cholesterol, the standard food package may be modified accordingly, replacing such high-cholesterol foods as eggs and regular peanut butter with such low-cholesterol foods as reduced-fat peanut butter and skim milk.

TABLE 1: Monthly WIC Food Package Costs—1997 (in dollars)

Pregnant women	38.48
Breastfeeding	38.48
Postpartum	30.82
For all women	36.51
Infants	21.88/79.93 before rebate
Children	34.46
Entire WIC population	31.67/46.28 before rebate

SOURCE: U.S. Department of Agriculture, Food and Nutrition Service, Office of Analysis and Evaluation, "Fiscal Year 1997 WIC Food Package Costs," April 7, 1999, p. 2.

In 1997, the average WIC food package was worth about $46 per month. The actual cost to the government, however, was only $32 because of manufacturers' rebates on infant formula. (States are required to pursue cost-containment strategies, such as obtaining rebates from infant formula manufacturers, a practice that saved $1.3 billion in 1997. Essentially, the rebate savings must be used to expand program participation, as described below.) This is the average across all WIC recipients, however. The total value of the two WIC food packages provided to a postpartum mother and her newborn can exceed $100 a month. Table 1 shows the average monthly cost of WIC food packages for the program's various target groups. (It also shows the actual value of the infant food package, both before and after taking into account the rebates.)

NUTRITIONAL AND HEALTH COUNSELING

Each time WIC recipients are certified, they must be offered at least two nutrition education sessions. These are voluntary sessions; vouchers cannot be conditioned on attendance (although it appears that some participants do not realize this). Designed to teach about the importance of good nutrition and its relationship to good health, these voluntary sessions may be either one-on-one or in groups. Participants are also instructed on how to deal with their own particular "nutritional risks" and those of their children. Pregnant women are encouraged to breastfeed, for example, unless not medically advisable. (These sessions generally last only about 15 minutes.)

States are limited in the amount that they can spend on "nutrition services and administration." In 1997, it was only about $11 per participant per month. The amount is based on a formula that essentially limits expenditures to an inflation-adjusted per-participant cost. (The initial figure was established in

1987 based on the average nationwide cost per participant.) States may vary the amount spent on individual recipients, but the total they spend must fall within the USDA limit. (Local WIC agencies are required to spend at least one sixth of their administrative funds on nutrition education.)

There is evidence that this cap is too low, or at least that, under the cap, agencies are increasingly unable to provide the level of services that they once did. To assess the impact of the WIC's rapid growth and new regulatory and legislative requirements imposed during the 1988 to 1993 period, researchers at Macro International, Inc., and the Urban Institute (1995) used a mail survey of all state WIC agencies and a nationally representative sample of local agencies, along with case studies in 22 local agencies. (Among the new mandates were measures encouraging cost containment; providing drug abuse education; encouraging breastfeeding and immunizations; expanding access to the homeless, incarcerated women, and working mothers; and requiring public assistance agencies, including WIC agencies, to provide voter registration services.)

Macro International, Inc., and the Urban Institute (1995) concluded that both the expansion and the new federal requirements led to funding constraints on "nutrition services and administration" that, in turn, led to scheduling less time per participant; increasing the use of group nutrition education (as opposed to individual sessions); and issuing standard food packages (as opposed to tailoring them to the individual needs of participants).

No wonder, then, that the limited evidence available suggests that participants still have major gaps in their knowledge of appropriate dietary practices. For example, a recent study of six WIC offices in three states, conducted by Mary Kay Fox and her colleagues (1999) at Abt Associates, found that although many participants have "reasonably high levels of nutrition knowledge," a substantial number still engage in a number of inappropriate feeding practices (p. xi). More than 40% of the mothers "offered their babies something other than breastmilk, formula, or plain water before the age of 4 months" (p. xii). (In some sites, more than two thirds reported doing so.) Less than half of prenatal and postpartum participants said that they learned something from WIC (p. 5-6).

That such practices continued, despite the parents' participation in WIC, should not be surprising. Counseling sessions are voluntary and, as Peter Rossi (1988, 55) concludes in *Feeding the Poor: Assessing Federal Food Aid*, "certainly inadequate for all but superficial instruction." In fact, it appears that many parents choose not to participate. According to the Fox et al. (1999) study, although most WIC participants were offered two nutrition education contacts each certification period, the number of contacts is actually much

lower. For example, only 5% to 59% (depending on the site) of postpartum WIC participants received two contacts, because many participants failed to show up for scheduled nutrition education activities (p. 2-4). The researchers (p. 2-5) also noted that in some sites, women "were observed openly refusing to attend nutrition education classes."

REFERRALS FOR OTHER SERVICES

WIC agencies are also expected to refer clients to other health care and social service providers. For example, WIC counselors might send pregnant women to prenatal-care or smoking-education classes, as well as to drug- and alcohol-addiction and family-planning services.

Although WIC is generally administered by a health organization, this connection to health care providers does not always translate into easy access to health services. According to one nationally representative survey of local clinics, only 52% of WIC agencies (serving just 39% of all WIC recipients) reported that health care services were available on-site, most commonly through a public health clinic (Macro International, Inc., and The Urban Institute 1995). Moreover, even when health services were available, the range of services often was limited. (The most common services available at WIC clinics were breastfeeding promotion, family planning, and immunizations. Less likely to be available were obstetrical/gynecological care and Medicaid screening.) As a result, WIC staff do not always have all of the health information or services that WIC families need, so that the families must be referred to various health care providers located elsewhere. This means that some WIC families have to make visits to multiple providers, leading some of them to become frustrated and stop seeking assistance.

3. PROGRAM COVERAGE

Because of the relatively high income cutoff (almost twice the poverty line), loose interpretations of nutritional risk, and the fact that the middle class is having relatively fewer children than in the past compared to lower income Americans, a surprisingly large proportion of Americans receive WIC benefits: nearly half of all infants, one quarter of all children ages 1 to 4, and the same proportion of pregnant women. (We estimate that another 10% of American children are eligible for benefits but do not receive them because WIC is not fully funded; that is, appropriations are not sufficient to provide services for all those who are statutorily eligible.)

ELIGIBILITY

Under federal rules, eligibility for WIC is based on low income plus nutritional risk. Income eligibility is set at family incomes up to 185% of the poverty income guidelines ($30,443 for a family of four as of July 1, 1998). (Recipients of TANF [temporary assistance for needy families], food stamps, and Medicaid benefits are automatically deemed income eligible.)[2]

In addition to income, however, WIC recipients must also be at nutritional risk. (Eligibility for other major federal food programs is usually based solely on income.) Nutritional risk is a broad concept including medical conditions such as anemia and low weight or overweightness; a mother's age, history of pregnancy complications, or poor outcomes in prior pregnancies; and inadequate diet.

WIC is not an entitlement program, another difference from other federal nutrition programs. Not all the people who fall within WIC's eligibility rules receive services, because the program is not fully funded to that level of benefits. The number of women and children served in a given year is established by the amount Congress appropriates, with states free to add supplemental funding. (Federal regulations require that openings in the program be filled by giving priority to pregnant and breastfeeding women and infants who have a nutritionally related medical condition, followed by other target groups and those with less severe nutritional risks.)

Up until now, at least, the process of determining nutritional risk appears to have been inexact, at best. According to the Institute of Medicine's (1996) Committee on Scientific Evaluation of WIC Nutrition Risk Criteria, some states have used "generous" cut-off points and "loosely defined risk criteria." The committee concluded that there were "serious gaps on the evidence" for some of the risk criteria, with unreliable tools used to measure them. Moreover, some of the criteria seem to have been loosely applied: "Street-level bureaucrats" were able to qualify someone who is marginally at risk or not at risk at all (Rossi 1988, 98). Although the extent of these problems is not known, effective April 1, 1999, the USDA revised WIC's eligibility guidelines to tighten the process for determining nutritional risk (U.S. Department of Agriculture 1998).

It is too early to tell whether the changes will affect frontline decision making. However, our own analysis, described below, suggests that in many if not most places, nutritional risk is assumed if the family meets WIC's income criteria. Moreover, it appears that even income determinations may not be as accurate as one would like. A 1988 USDA study of WIC participants estimated that 5.7% were not eligible because their income was too high (see U.S. General Accounting Office 1999, 23). We speculate that the

TABLE 2: WIC Expenditures and Recipients—1977 to 1997 (in thousands)

Year	*Expenditures (1996 dollars)*	*Women*	*Infants*	*Children*	*Total Recipients*
1977	668	165	213	471	848
1978	926	240	308	633	1,181
1979	1,161	312	389	782	1,483
1980	1,413	411	507	995	1,913
1981	1,530	446	585	1,088	2,119
1982	1,546	478	623	1,088	2,189
1983	1,775	542	730	1,265	2,537
1984	2,093	657	825	1,563	3,045
1985	2,174	665	874	1,600	3,138
1986	2,244	712	945	1,655	3,312
1987	2,296	751	1,019	1,660	3,429
1988	2,397	815	1,095	1,683	3,593
1989	2,450	952	1,260	1,907	4,118
1990	2,572	1,035	1,413	2,069	4,517
1991	2,646	1,120	1,559	2,214	4,893
1992	2,875	1,222	1,684	2,505	5,411
1993	3,045	1,365	1,742	2,813	5,920
1994	3,349	1,499	1,786	3,192	6,477
1995	3,555	1,577	1,817	3,500	6,894
1996	3,688	1,648	1,827	3,712	7,188
1997	3,768	1,710	1,863	3,835	7,409

SOURCE: 1998 Green Book (U.S. House Committee on Ways and Means 1998), and for 1997, USDA, Food and Nutrition Service.

rapid growth in program participation since that time may have made this a more pervasive problem today.

PARTICIPATION

Always a popular program, WIC has grown rapidly. (In 1994, WIC's name, the "Special Supplemental Food Program for Women, Infants, and Children," was changed, replacing the word *food* with *nutrition*.) When it was permanently authorized in 1974, WIC served only 88,000 women and children, at a cost of $10.4 million. Participation increased to nearly 2 million in 1980, at a cost of $725 million, and to 4.5 million in 1990, at a cost of $2.1 billion. By 1997, participation reached 7.4 million, at a cost of $3.7 billion. If one also includes the rebates that infant formula manufacturers are pressured into giving, the program provides another $1.3 billion worth of food, for a total value of about $5 billion (U.S. Department of Agriculture 1999, 1).

TABLE 3: WIC Participation—1991 to 1997 (as a percentage of eligible persons[a])

	Pregnant Women	*Postpartum/ Breastfeeding*	*Infants*	*Children*	*Total*
1991	61	47	91	44	56
1992	52	—	96	44	56
1993	52	79	98	48	60
1994	59	101	111	57	70
1995	58	105	109	64	75
1996	62	117	114	69	81
1997	69	122	122	75	87

SOURCE: U.S. Department of Agriculture, Food and Nutrition (or Consumer) Service, "Special Supplemental Nutrition Program for Women, Infants, and Children (WIC): Eligibility and Coverage Estimates," selected issues.
NOTE: According to the U.S. Department of Agriculture (USDA), the high estimated participation rates for some groups (including more than 100% for the postpartum/breastfeeding mothers and for infants) are due to differences between the way the number of income-eligibles is estimated and the certification practices applied in local WIC agencies. In addition, some imprecision is present in any survey-based estimate. USDA concludes, "These data do strongly suggest that the program has likely achieved virtually full coverage of persons in this category at the national level."
a. Because pregnant women are unlikely to participate in WIC for a full 40 weeks, their participation rate is expected to be less than 100%. For example, if all eligible pregnant women enrolled in WIC for 6 months, their participation rate would be 65%.

In 1997, slightly more than half of all participants were children between ages 1 and 4, one quarter were infants, and nearly one quarter were women, about half of whom were pregnant (about 11% were pregnant, 5% were breastfeeding, and 7% were postpartum) (U.S. Department of Agriculture 1999, 1). Spending by participant category was as follows: children, $2.175 billion; infants, $670 million ($2.450 billion if one includes the fair market value of the foods purchased with the rebate); and women, $1.030 billion.[3] (Table 2 shows the historical growth in WIC expenditures and recipients.)

Because eligibility is based on the vague concept of nutritional risk, it is difficult to gauge the actual size of the eligible population and, hence, the percentage of the eligible population that is participating. The best estimates come from the USDA Food and Nutrition Service (1997), which are based on health survey data and approximate the percentage of those income-eligible who also have at least one nutritional risk. The USDA applies an estimate that "about 4 out of 5 income eligible persons are also at nutritional risk and thus fully eligible for the WIC Program."[4]

Based on the USDA's estimates of the number of income-eligible families who are at nutritional risk, it appears that in 1997, about 87% of eligible individuals were participating. As Table 3 shows, participation rates for pregnant women and older children were lower, 69% and 75%, respectively. (The

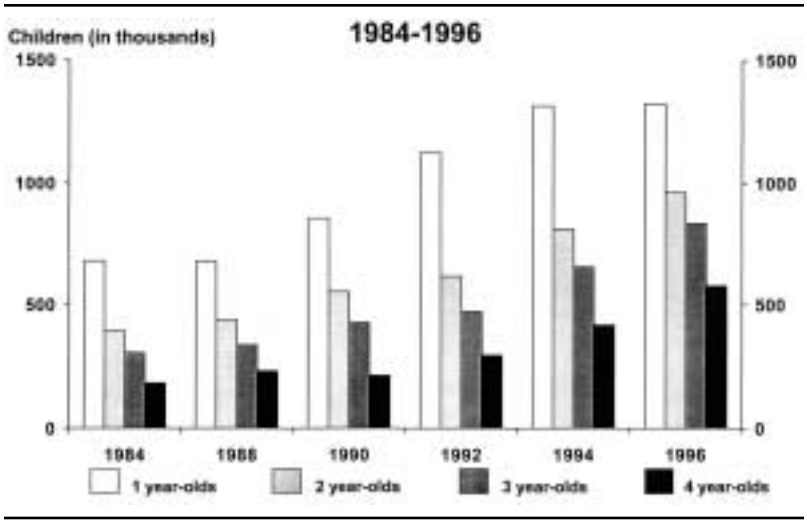

Figure 1: WIC Enrollment of Children Ages 1 to 4

figure for pregnant women, however, is artificially depressed because it is estimated by using as the base all eligible pregnant women, from the very first week of pregnancy. Hence, even if all women participate for 6 months of their pregnancy, the participation rate would be only 65%.) Most striking is that 122% of the estimated number of eligible postpartum and breastfeeding mothers and 122% of the estimated number of eligible infants are participating. (Note that this is a percentage of the estimated eligible persons. It is possible that not all eligible persons are participating and that there are more ineligibles than the 122% participation rate suggests.) We analyze this inconsistency in the next section.

There is a significant drop-off in WIC participation among families with older children. Theoretically, a mother who starts in a WIC program when she is pregnant should continue until her youngest child reaches age 5 (assuming continued income eligibility and nutritional risk). However, participation drops off rapidly after the first year. For example, in 1996, nearly 2 million infants participated in the program, but only about 1.4 million 1-year-olds did. With each successive age group, participation fell, so that only 600,000 4-year-olds participated in the program (Randall, Bartlett, and Kennedy 1998, 36). (See Figure 1.)

Although part of the drop-off may be due to the difference in eligibility criteria for infants and children, as well as to the fact that family incomes tend to be higher in families as children grow older, we suspect that a major contributor is the smaller value of the food package once the mother is no longer

eligible to receive benefits for herself. For example, the value of a food package for a postpartum mother and her infant was worth about $110 per month, compared to just $34 for a child (U.S. Department of Agriculture 1999, 2). After a while, many mothers may simply decide that the small amount of food is not worth the time or trouble of continued participation. (In a survey of WIC recipients, the supplemental foods WIC provides were listed as the most attractive program attribute among mothers in the prenatal and postpartum components [Fox et al. 1999]. The study reported, "This was the only program characteristic that was consistently included in the top three positive aspects of the WIC Program" [Fox et al. 1999, p. xiv].)

ELIGIBILITY CREEP?

Eligibility for WIC is set at 185% of the poverty line plus nutritional risk. However, as we saw, nutritional risk is an amorphous concept, and especially as funding has increased, most agencies seem to have assumed that all income-eligible applicants are at nutritional risk. (Like so many other areas of WIC implementation, there is little systematic evidence of such practices. However, this is the most likely explanation for estimated participation levels that otherwise imply that all income eligible people are at nutritional risk, even as medical evidence indicates otherwise.) Hence, in most places, eligibility, essentially, has become solely a matter of income. Moreover, as program funding has increased, according to some local WIC staff, even income testing seems to have become less rigorous, with many participants having incomes over eligibility limits. Remember that 122% of the estimated number of eligible postpartum and breastfeeding mothers and of eligible infants are participating in WIC.

USDA estimates seem to confirm this "eligibility creep." For example, for various WIC target groups, the number of participants now exceeds the number that is income eligible for the program. Between 1991 and 1997, the number of infants estimated to be eligible for WIC fell from 1.7 million to 1.5 million. During the same period, however, the number of infants in WIC rose from 1.6 million to 1.9 million, so that by the end of the period, according to the USDA's own data, there were 336,000 more infants in WIC than were estimated to be eligible (22% over the estimate).

There are many valid reasons why there could be more WIC participants than estimated to be eligible: Medicaid policies that allow persons with incomes above 185% of poverty to enroll in Medicaid and thus be eligible for WIC, problems with the census data used to estimate income eligibility, and differences in the family unit and income measures used by WIC agencies

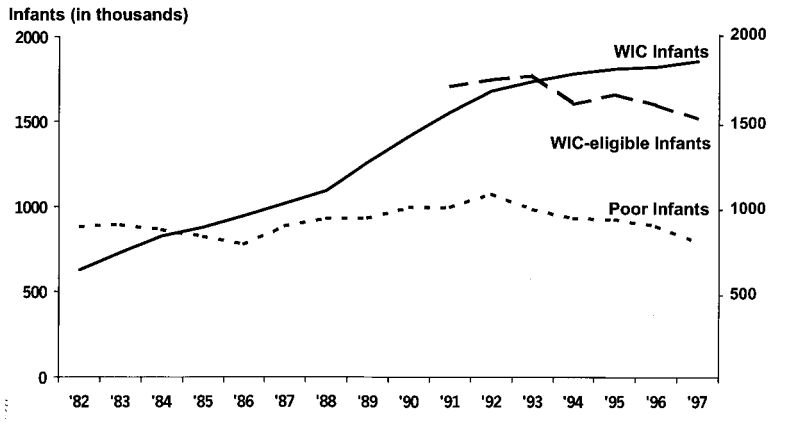

Figure 2: WIC Coverage of Infants

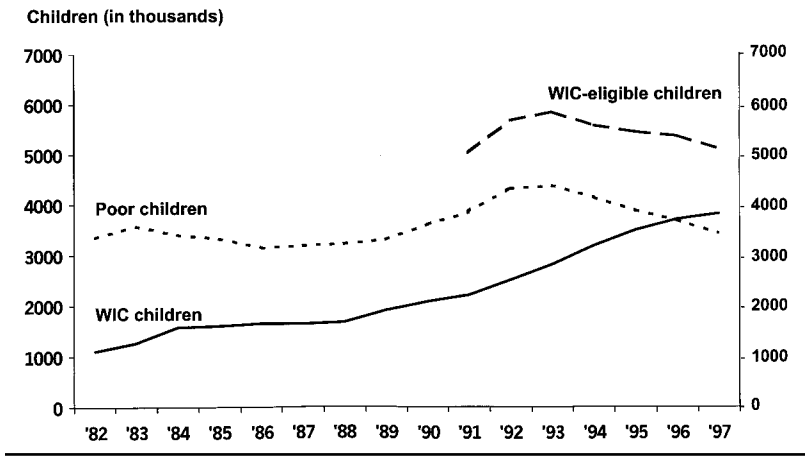

Figure 3: WIC Coverage of Children

and those used by USDA to estimate eligibility (see Lewis and Ellwood 1998). But the larger point is that at the very least, all infants falling within the family income limits are now in the program. That means that either all infants under 185% of the poverty measure are at nutritional risk or that the nutritional-risk criteria is not being applied. Like many other observers, we think that it is clearly the latter.

Similar trends apply to pregnant, breastfeeding, and postpartum women, and the same process is well on its way for children ages 1 to 4, with about 75% of eligibles participating. (As noted in Table 3, the 69% participation estimate for pregnant women is artificially low because it is unlikely that such women could participate a full 40 weeks.)

Figures 2 and 3 portray this expansion in the number of WIC infants and children, respectively, together with the number of poor in both groups. The graphs suggest that the progression of expansion was from "poor" to "less poor" infants and children, because the lines cross in 1984 and 1996, respectively. However, the graphs are also consistent with increased participation at all eligible-income levels.

Data from the Survey of Income and Program Participation (SIPP) suggest that WIC's expansion has been primarily from low-income poor to those with relatively higher incomes. Richard Bavier (personal conversation with Peter Germanis, 1999) of the U.S. Office of Management and Budget used SIPP data to examine the distribution of WIC participants by income level in 1988 and 1996. He found that the percentage of WIC participants in families with annual incomes above $25,000 (measured in constant 1996 dollars) rose from 21% in 1988 to 29.4% in 1996, a 40% increase. As Bavier notes, these data are composed of differences in the reporting of WIC receipt in the two years, with considerably higher reporting in 1996 than in 1988. (It is unknown whether these differences bias the findings.) Nevertheless, his findings are consistent with informal reports from the field (and common sense). Thus, it appears that once the program reached those with the lowest incomes, eligibility was expanded by bringing in participants from higher income families.

There are a number of explanations for this "eligibility creep." Outsiders do not understand the program and assume that because the statutory criteria require that recipients be at nutritional risk, all WIC families must need government aid. Insiders have been hesitant to raise the question, partly because they believe in the program and want its benefits to go to as many people as possible and partly because any suggestions that the program should be better targeted would suggest that the basic program requires improvement; and of course, that would be inconsistent with the inflated benefit-cost estimates that have become part of the political landscape.

Feeding this process has been the unalloyed success of the infant formula rebate program, which has provided billions of dollars to WIC with little legislative oversight. In 1988, the rebates provided WIC with $32 million in additional funds, permitting the program to add 63,000 participants. In 1991, the rebates had grown to $650 million, fueling an increase in participation of nearly 1.1 million. By 1997, the rebates totaled $1.3 billion, adding 1.9

million participants to WIC, roughly one quarter of the program's entire caseload and one third of its appropriated funding.

Coming to the program outside the normal appropriations process, these billions of dollars have been automatically applied under WIC's eligibility and funding rules, without considering whether the additional funds should be used to adjust program benefits or services. As a result, states have been forced to use these savings to expand participation, generally to those with higher incomes (and lower nutritional risk), rather than to improve the program. There are legitimate reasons for some limits in a program as large and diverse as WIC, but forcing states to add more and more families to the program is not one of them. In so doing, we have lost an opportunity to improve WIC's effectiveness, by directing enhanced benefits and services to the most needful.

More expansions of WIC are looming. All persons eligible for Medicaid are automatically deemed income eligible for WIC. Recent legislation has allowed states to expand Medicaid to children in families with incomes above the 185% of the poverty standard (which is also the WIC income limit). In response, 22 states have increased income eligibility levels for Medicaid beyond 185% of poverty (generally, to somewhere between 200% to 300% of poverty for at least one of WIC's target groups) (Krause 1999). Hence in many states, WIC eligibility is now more than 185% of the poverty standard, effectively increasing the size and relative income of the WIC eligible population. Other states will likely follow suit.

4. PAST RESEARCH

All of us would like WIC to be as successful as claimed. And it makes common sense that providing food packages and nutritional counseling to the poor ought to improve their diets. The plain and almost undisputed fact, however, is the following: Beyond modest reductions in anemia and modest increases in the intake of selected nutrients, there is little research evidence about the effectiveness of almost 90% of the funds expended under the WIC program. The purported 3:1 savings calculation comes from research on the roughly 11% of the program that serves pregnant women and is based on WIC as it existed more than a decade ago. Moreover, even that research is fraught with methodological uncertainty.

In the following pages, we review the research results for WIC, target group by target group. There is, however, much more research on WIC than reported here.[5] As Leighton Ku (1999, 108) notes, however, "Many of the

most interesting papers are not rigorous evaluations, but reports of field studies of descriptive analyses." Hence, like Rossi (1988) in *Feeding the Poor*, we have limited our discussion to those studies that roughly meet accepted standards of scientific rigor. As we will explain, however, even these studies are plagued by serious statistical concerns, largely around questions of selection bias, simultaneity bias, and generalizability.

POSTPARTUM AND BREASTFEEDING MOTHERS

About 12% of WIC participants are postpartum and breastfeeding mothers (accounting for about 12% of WIC spending). One would assume that nutritional supplements would aid such women or their next child, especially if the period between births is relatively short. There is, however, scant research on the impact of nutrition supplements on postpartum and breastfeeding mothers.

The only study that we could find on WIC's effects on postpartum mothers was conducted by Bette Caan of the School of Public Health, U.C. Berkeley, and her colleagues (Caan et al. 1987). Using California administrative data from the 1980s, they evaluated the effect of WIC's postpartum benefits for women in California who (a) gave birth to a child in the early 1980s and (b) had a subsequent birth within 3 years. They found relatively small effects: For infants born to mothers who received WIC postpartum benefits for 5 to 7 months between their two pregnancies, the mean birthweight was about 4% higher than for infants born to women who received postpartum benefits for less than 3 months or not at all. (Women in both groups received WIC prenatal benefits during both pregnancies.)

The validity of even this small positive impact, however, is undermined by questions of selection bias: The better functioning mothers presumably sought the WIC benefits for a longer period. Another problem is that about 20% of the original sample was dropped due to missing data (which could introduce attrition bias if, for example, the dropped cases have systematically different birth outcomes from those that remained).

Some observers have suggested that the free infant formula that WIC provides may discourage breastfeeding, ordinarily considered better for newborns. This is a concern often expressed by local WIC staffers. Because although they encourage women to breastfeed, many believe that the free infant formula it provides (with a market value of about $80 a month) gives a powerful signal in the other direction, especially as more of these mothers have jobs, making it more difficult for them to maintain a regular breastfeeding schedule. The research is not enlightening.

The National WIC Evaluation, conducted in the early 1980s by a group of researchers at the Albert Einstein College of Medicine in New York City and the Research Triangle Institute in North Carolina, compared the rates of breastfeeding of a nationally representative sample of WIC participants at hospital discharge to a similarly low-income group of pregnant women not enrolled in WIC (Rush, Sloan, et al. 1988). After attempting to control for socioeconomic differences between the two groups, it found no statistically significant difference in the rates of breastfeeding between the two groups.

A more recent study, also using nationally representative data and multivariate statistical techniques, compared the rates of breastfeeding initiation and the duration of breastfeeding between WIC participants and eligible nonparticipants (Schwartz et al. 1992). Both prenatal WIC participants and eligible nonparticipants had comparable rates of breastfeeding, approximately 35%. However, prenatal WIC participants who received advice about breastfeeding had much higher rates than those who did not (44% compared to 25%), although there was no difference in the duration of breastfeeding.

It is not clear how to interpret this finding. On one hand, given this impact of counseling on the likelihood of breastfeeding, one might expect that it would also extend the average period of breastfeeding for those mothers who would have done so without the counseling. On the other hand, the addition of some mothers who otherwise would not have breastfed may have reduced the period of breastfeeding, if their commitment to breastfeeding were more tenuous.

In response to continuing concerns that WIC discourages breastfeeding, in 1993, the USDA introduced an enhanced food package for breastfeeding mothers who forego the infant formula food package. No rigorous research has been conducted on the reconfigured program. Thus, as Urban Institute researchers concluded in 1994, the most that can be said is, “It is not clear whether WIC promotes or hinders breastfeeding” (Ku, Cohen, and Pindus 1994, 5).

INFANTS

About 25% of WIC participants are infants (accounting for about 17% of spending), and again, there is little research evidence of effectiveness. The National WIC Evaluation, conducted in the early 1980s and described above, is also the most comprehensive examination of WIC’s impact on infants (Rush, Leighton, et al. 1988). It found that although infants participating in WIC did not have higher caloric intakes than those in the comparison group, they had higher mean intakes of certain nutrients, especially iron and vitamin

C. WIC infants also had lower mean intakes of calcium, reflecting the fact that they were more likely to consume infant formula than whole milk. In addition, on a given day, significantly fewer infants were likely to fall below 77% of the recommended daily allowance (RDA) for iron, vitamin A, and vitamin C (Rush, Leighton, et al. 1988, 487-90). They were also more likely to have a regular source of medical care and to be immunized than similar low-income children not on WIC (Rush, Leighton, et al. 1988, 504-07).

However, as Rossi (1988) concludes, the overall study was

> plagued by a number of technical problems concerning the suitability of the comparison group used as well as severe response rate problems for both the WIC and comparison samples. Accordingly, the findings derived cannot be regarded as anything more than suggestive. (P. 54)

CHILDREN

Slightly more than half (52%) of all WIC participants are children ages 1 to 4, making this the largest target group served by the program. (This part of the program accounts for about 56% of expenditures.) But here, too, the body of research on effectiveness is disappointingly sparse, with a few studies concluding that WIC reduced iron deficiency among poor children and modestly increased the intake of selected nutrients.

Again, the National WIC Evaluation, conducted in the 1980s, found that children participating in WIC also did not have higher levels of caloric intake than the comparison group, but they did have higher mean intakes of certain nutrients, especially iron and vitamin C. On a given day, fewer children were likely to fall below 77% of RDA for iron and vitamin C (Rush, Leighton, et al. 1988). The strongest positive dietary effects were for the least advantaged children, including those who were very poor and from female-headed households. (No significant differences were found between past WIC participants and the comparison group, which suggests that WIC does not make long-term dietary improvements, whatever its short-term effects may be.) WIC children were also more likely to be immunized and have a regular source of medical care than similar low-income children not on WIC. However, as noted above for infants, these findings should be viewed as nothing more than suggestive because of technical problems related to the evaluation.

A 1998 study by Donald Rose of the USDA's Economic Research Service, Jean-Pierre Habicht of Cornell University, and Barbara Devaney of Mathematica Policy Research (MPR) used the 1989-1991 Continuing Survey of Food Intake by Individuals (CSFII) to compare preschoolers ages 1 to 4 years old in households with incomes below 130% of the poverty level

participating in WIC to a comparable group of preschoolers not in WIC (Rose, Habicht, and Devaney 1998). After attempting to control for differences in socioeconomic characteristics and possible selection bias, the study reported that WIC had statistically significant positive effects for 10 of the 15 nutrients examined, with particularly significant increases in the intake of iron and zinc (often lacking in the diets of WIC children). (The increases in iron and zinc were 16.6% and 10.6% of RDA, respectively.) Although the diets of most children already exceeded the RDAs for many nutrients, these findings suggest significant dietary improvements for particular groups.

This research, unfortunately, suffers from the generic weaknesses of comparison group studies. Moreover, the considerable differences in the socioeconomic characteristics of WIC participants and the comparison group of eligible nonparticipants make it particularly difficult to judge how successfully the study controlled for these differences. (For example, 47% of those receiving WIC were White compared to 66% of those eligible, but not participating, in WIC. There were other important differences in ethnicity, age of the youngest child, and geographic location.)

These studies, however, were conducted before the rapid expansion of WIC participation, among children, that occurred during the 1990s. A new study by USDA economists Victor Oliveira and Craig Gunderson, using data from the 1994-1996 CSFII, compared the nutrient intake of preschoolers ages 1 to 4 with incomes below 200% of the poverty level participating in WIC to a comparable group of preschoolers not in WIC (Oliveira and Gunderson 2000). After attempting to control for observable differences in socioeconomic characteristics, the study reported that WIC has statistically significant positive effects for five of the eight nutrients examined: iron, vitamin C, vitamin A, vitamin B-6, and folate. However, this analysis did not control for possible selection bias.

Recognizing that the two groups may differ in unobservable ways, such as the nutritional motivation of the parents, they attempted to control for possible selection bias by limiting their analysis to households in which a woman or infant was participating in WIC. From this group, they compared the nutrient intake of WIC children to non-WIC children. Because all households had at least one WIC participant, the theory is that this would control for such unobservable household characteristics as "nutritional awareness and motivation." (The authors point out that the traditional selection bias methods to control for selection bias are based on modeling the participation process but that the CSFII did not provide sufficient information to do so. Thus, they resorted to an "indirect" method.) Using this sample, they reported that WIC has a significant positive effect on children's intake for three of the eight

nutrients examined: iron, folate, and vitamin B-6. They suggested, however, that these findings understate WIC's impact, because the statistical model could not control for nutritional risk (due to data limitations) and it seems likely that the non-WIC children in the comparison group may be less disadvantaged, particularly in terms of their nutritional risk.

Here again, however, the generic weakness of comparison group studies and differences in the socioeconomic characteristics of WIC participants and the comparison group of eligible nonparticipants raise uncertainty about how well the study controlled for these differences. (For example, 43% of the WIC sample consisted of 2-year-olds, compared to 22% of those in the comparison group. There were other important differences in ethnicity, children's age, and geographic location.)

Ray Yip at the Centers for Disease Control (CDC) and his colleagues took a different approach, using aggregate time series data rather than comparison groups. With data from the CDC's Pediatric Nutrition Surveillance System (PNSS), they examined the incidence of anemia among children ages 6 months to 60 months who were enrolled in public-health programs (primarily WIC) (Yip et al. 1987). (The analysis was limited to six states that consistently participated in the PNSS during the time period examined.) Anemia among children ages 6 months to 60 months declined steadily from 7.8% in 1975 to 2.9% in 1985. Other studies report similar decreases (Yip 1989; Sherry, Bister, and Yip 1997). The trends were examined separately for children seen at preenrollment screening visits as well as for those seen at follow-up visits. (The rate of anemia fell by more than 50% for both groups during the study period.)

The decline in anemia for both groups could be due to several factors. First, because the data are limited to children covered by programs such as WIC, this decline could also have reflected a change in the composition of the children covered by the programs. For example, as WIC's funding and enrollment have expanded, the children in the program have probably become less disadvantaged. If the newly eligible children had lower rates of anemia than the earlier WIC recipients, they would have lowered the average rate of anemia, even if the WIC benefit itself had no effect. (This effect is likely, because the priority system used by the program targets those with the greatest nutritional risks and WIC expansion would generally lead to the enrollment of those with less severe nutritional problems.)

An example of this relationship was illustrated by Sherry, Bister, and Yip (1997) in their examination of the decline in the prevalence of anemia in Vermont. Although the prevalence of anemia dropped steadily from 1984 on, there was a slight increase in 1991 and 1992. The researchers noted that

funding constraints led to some restrictions on eligibility in those years. They explained, "This increased the proportion of high-risk children, thus increasing the likelihood of anemia" (p. 930).

Second, this reduction may be part of a more general downward trend in iron-deficiency anemia, caused by such factors as an increase in breastfeeding, the substitution of iron-fortified formula for unfortified formula and cow's milk, and general nutritional education efforts. WIC may also have contributed indirectly to this trend through its required iron fortification of many foods. As Barbara Devaney and her colleagues (Devaney, Ellwood, and Love 1997), who conducted some of the major WIC studies for MPR, observe,

> WIC may also have had a role in reducing the prevalence of iron-deficiency anemia over time among all infants and children, including those who do not receive WIC. Because WIC vouchers constitute a large share of the market for infant formula and children's cereal, manufacturers may have changed the iron content of their products to meet WIC's eligibility requirements that include iron fortification. (P. 96)

"Thus," Devaney (1998, pp. 120-21) adds in another publication, "much of these foods that are on the shelves of supermarkets are iron-fortified and affect the diets of nonparticipants as well as program participants." If this is the real explanation, a more direct way of achieving the beneficial effects on anemia claimed for WIC would be simply to require manufacturers to fortify infant formula, cereal, and bread.

But this is speculation. No study has successfully isolated WIC's impact from the changing socioeconomic characteristics of recipients nor from the apparent secular decline in anemia. (In Tennessee, where information on the socioeconomic characteristics of participants was available, the prevalence of anemia declined for several different socioeconomic groups, suggesting that the overall decline was due to more than just a changing composition of those covered by WIC [and other public health programs].) Moreover, the existing studies are limited to a small number of states and, therefore, are not necessarily representative of the nation as a whole.

More important, the practical significance of these modest findings is unknown. If there had been a socially significant reduction in anemia, one would expect, for example, to see it reflected in a reduction in the behaviors associated with anemia. But as Devaney, Ellwood, and Love (1997, 95) comment, "Little is known about the long-term effects of WIC on improving behavioral and cognitive development, outcomes that would presumably result from better iron nutrition status." This sentiment is echoed by James Ohls (1998), a senior researcher at MPR, who states,

Critics of the program correctly note that there is essentially no information available about the effects of the program on children in the one- to four-year-old range, the group currently "at the policy margin" in terms of how more (or less) funding would be used.

PREGNANT WOMEN

The purported 3:1 savings calculation, cited by so many advocates and politicians, comes solely from research on one of the smallest parts of WIC, the program for pregnant women (11% of participants and 14% of expenditures). Giving impetus to this claim was a 1992 report from the U.S. General Accounting Office (GAO) that reviewed 17 "methodologically strong" evaluations of this part of WIC. The GAO estimated that prenatal WIC participation resulted in a 25% reduction in low birthweight births (less than 2,500 grams) and a 44% reduction in very low birthweight births (less than 1,500 grams).

Only one study examined the impact of WIC using a random assignment methodology. James Metcoff and his colleagues (1985) randomly assigned 471 women attending the prenatal clinics at the Oklahoma Memorial Hospital to WIC or to a control group. WIC participation was associated with an increase in mean birthweight of 91 grams (2.9%). However, despite the researchers' reliance on random assignment, WIC mothers were heavier than mothers in the control group at the time of entry into the study. When the researchers controlled for this difference, the birthweight finding was no longer statistically significant, except among mothers who smoked. (As discussed below, this suggests targeting WIC to women who are at greatest risk.) In any event, despite its rigorous methodology, the findings are not generalizable to the United States or the current WIC program, because the sample is drawn from a population receiving prenatal care in Oklahoma in the early 1980s.

One of the most influential studies the GAO reviewed was conducted by Devaney and her colleagues (Devaney, Bilheimer, and Schore 1991). In five states, they matched 1987-1988 Medicaid and WIC records and then compared the outcomes of Medicaid recipients receiving WIC services with those not receiving WIC. After applying statistical controls for identifiable demographic and parental-care characteristics that could also affect birth outcomes and Medicaid costs, the researchers found that WIC participants had fewer premature births and that their newborns had 1.6% to 3.8% higher birthweights (depending on the state). (WIC participants had higher birth weights, ranging from 51 grams in one state to 117 grams in another, although more significant gains were achieved for preterm births, ranging

from 138 to 159 grams.) Moreover, their estimates suggest a striking 28% decrease in the number of newborns with *low* birthweights (less than 2,500 grams) and an even larger 59% decrease in the number of newborns with *very low* birthweights (less than 1,500 grams). They estimated that WIC reduced the incidence of low birthweight newborns from an estimated 13.6% without the WIC program to an estimated 9.8% with WIC. (These estimates were derived by averaging the findings across all five states included in the study.)

These apparent effects on birthweights are what have led to the estimates that WIC saves money. Low birthweight, and especially very low birthweight, is linked to higher neonatal mortality and various childhood illnesses and disabilities (Paneth 1995). According to one estimate, 35% of the amount spent on health care for infants is due to the added costs of providing medical care to low birthweight infants, estimated to be about $15,000 more for each such birth in 1988 (Lewit et al. 1995, 40). If WIC reduces the incidence of low birthweight, it would reduce the need for subsequent Medicaid and other services and, thus, save money.

Based on the 17 studies it reviewed, the GAO estimated that each dollar spent on WIC for pregnant women saved $3.50 (over an 18-year period) in Medicaid and disability payments paid by federal, state, and local governments, and in medical care costs borne by private health care providers. (The estimated savings to the federal government were just $1.14, with $1.04 in savings to state and local governments and $1.32 to private payers, such as hospitals and insurance companies.) Here are the GAO's (1992) exact words:

> [Every dollar] invested in WIC benefits returns an estimated $3.50 over 18 years in discounted present value, and $2.89 within the infants' first year to federal, state, and local governments and to private payers. Because GAO did not quantify all program benefits and estimate all potential cost savings at current eligibility and participation levels, savings may be greater than these estimates. (P. 4)

Because the 17 studies varied in quality and in the effects they found, the GAO's statement was actually a composite and weighted estimate of WIC's effects.[6] Moreover, GAO did not measure the savings associated with WIC by examining reductions in Medicaid and other program costs directly, as did some evaluations (see Devaney, Billheimer, and Schore 1991; Schramm 1985). Instead, GAO estimated the savings in health care, disability, and special education costs associated with averting low and very low birthweight births. It then applied this savings estimate to the estimate of the number of infants who were born at normal birthweight rather than low birthweight. This erroneously assumed that averting a low birthweight birth is the same as producing a normal birthweight birth. Many of the "averted" low birthweight

births may simply have moved from just below the 2,500 gram cutoff to just above it, suggesting that the actual savings would be much lower.

A deeper look at this body of research reveals more reasons why the GAO's conclusions should have been considerably more tentative. All but one of the studies the GAO reviewed were based on nonexperimental comparisons of WIC participants to nonparticipants, with statistical attempts made to control for other factors that influence birth outcomes. As we will see, serious methodological problems with all of these studies undercut their findings.

5. RESEARCH WEAKNESSES

As mentioned above, this review does not include all research on WIC. Excluded were those studies that failed to meet basic tests for validity or that provided insufficient information for which to make a judgement. But even those included in this review have weaknesses that require their findings to be considered highly uncertain. The three most significant weaknesses are (a) selection bias, (b) simultaneity bias, and (c) lack of generalizability. (Because the major WIC research tends to concern its effect on pregnant women and their newborns, this section focuses on it as well.)

SELECTION BIAS

From a scientific standpoint, the preferred approach for measuring the effectiveness of WIC, or any social intervention, is a randomized experiment in which individuals eligible for WIC are randomly assigned to a treatment group (that receives WIC) and a control group (that does not). Properly implemented, random assignment should result in comparable treatment and control groups, so that any difference in subsequent outcomes can be attributed to the program rather than to some personal characteristic or external force that is systematically different in each group.

Unfortunately for research, but fortunately for recipients, high percentages of eligible families now participate in WIC. On the mistaken belief that it would be necessary to deny benefits to currently eligible individuals—necessary for most randomized experiments—most WIC studies rely on statistical comparisons between those who received WIC benefits and those who did not. However, because participation is voluntary, some unobserved factor, such as parental motivation, may affect both participation and outcomes.

The problem is this: If the pregnant women who voluntarily enroll in WIC are more concerned about their babies than those who do not, then their better birth outcomes may reflect this higher level of concern more than the program's nutritional counseling or food supplements. Conversely, if WIC recipients are at greater nutritional risk than nonparticipants, comparisons to nonparticipants may understate the positive effects of WIC. Gordon and Nelson (1995) of MPR describe the two possibilities:

> Some pregnant women may not participate in the WIC Program because they lack access to or knowledge of publicly funded programs that provide health care or other services, which may independently affect birth outcomes. Thus, the estimated improvement in birth outcomes may overstate the effect of the Program since, relative to nonparticipants, WIC participants may have better outcomes even in the absence of the WIC Program. Conversely, if the WIC Program is successful at reaching high-risk, low-income pregnant women, WIC participants may be more likely than nonparticipants to have poor birth outcomes in the absence of the program. In this case, the estimated improvement in outcomes would understate the true effect of prenatal WIC participation. In either case, estimates are contaminated by selection bias. (P. 124)

Careful regression analysis often can reduce selection bias, but even in the best circumstances, it cannot do so completely. But the results of a regression analysis are of limited credibility when the differences between participants and nonparticipants are large and there is little data available about their social, economic, and personality characteristics with which to explain differences in outcomes, as is the case in many WIC studies.

For example, no WIC study seems to have known the mother's marital status or age at first birth, possibly crucial determinants of her motivation and functioning. Sometimes, even more basic information is unavailable. In their 1991 study of WIC's impact on low birthweight, for example, Devaney and her colleagues (1991) relied on data from birth records and several selection bias models. Although they did not describe the statistical models they used to correct for selection bias, they reported that these models "yielded very unrealistic results that were extremely sensitive to both minor changes in model specification and the estimation procedure employed" (p. 58). (Unpublished data on Medicaid savings show widely divergent findings, depending on the model used, although none were statistically significant.) As a result, they did not publish their selection bias adjusted findings.

This absence of realistic results should not be surprising. Devaney and her colleagues did not even have data on household income and, instead, had to rely on the mother's educational attainment as a proxy. As a result, concludes Rossi (1988, 57), "It may not have been possible to control for some differences in economic status between WIC recipients and nonrecipients."

This "omitted variables" problem is why many researchers say that they cannot determine whether selection bias understates, or exaggerates, WIC's effectiveness. In our experience, we find it much more plausible that the more motivated and higher functioning mothers enroll themselves into the program, so that selection bias would tend to increase estimated effects. A decade ago, Wayne F. Schramm (1986) wrote,

> As with other studies of this type, WIC mothers were self-selected in that they were motivated to apply for and receive WIC benefits. Other Medicaid mothers may not have been aware of WIC because Medicaid referrals were not as good in their counties. Still others probably had the same opportunity but were not motivated to apply for WIC. These mothers may have been less interested in health and nutrition and these factors may have affected their infant's birth weights and Medicaid costs. (P. 614)

If there were no other evidence, however, we would agree with those who say that it is impossible to gauge the direction or the amount of selection bias and that it could be quite small. (See Rossi 1988, 58, stating, "It appears likely that estimates uncorrected for selection biases over-estimate WIC effectiveness but the degree of overestimation cannot be reliably determined.") However, recent studies persuade us that the bias in the past studies is quite large and substantially exaggerates WIC's estimated impact.

The first is a study by Gordon and Nelson (1995), which used data from the 1988 National Maternal and Infant Health Survey (NMIHS), a nationally representative source of information on the characteristics and experiences of women who experienced a live birth in 1988. They compared the birth outcomes of WIC recipients to income-eligible nonparticipants, but again, the two groups differed on a number demographic and socioeconomic characteristics, with WIC serving a generally more "disadvantaged" population. (WIC recipients were more likely to be teenagers [26.6% vs. 17.4%], Black or Hispanic [51.8% vs. 37.8%], and welfare recipients [37% vs. 21.9%]. They were less likely to have a high school diploma [66.5% vs. 75%], be married [48% vs. 64.3%], and have worked within the past year [54% vs. 61.9%].) Thus, statistical controls were needed to adjust for identifiable demographic, socioeconomic, and behavioral characteristics that could also affect birth outcomes.

After controlling for the observable characteristics of the two groups, Gordon and Nelson (1995) reported that WIC participants had infants with mean birthweights about 2% (68 grams) higher than infants in the comparison group. They also estimated that WIC reduced the incidence of low birthweight newborns (less than 2,500 grams) by 27%, from a mean of 10.8% without WIC to 7.9% with WIC. An even larger, 45%, reduction was

estimated for very low birthweight births (less than 1,500 grams), from a mean of 2.2% without WIC to 1.2% with WIC.

To identify the impact of selection bias, Gordon (1993) tested their results using three selection bias models: the Heckman two-step or "Heckit" procedure, an instrumental variables estimator, and a maximum likelihood method.[7] These tests are designed to measure the impact of nonrandom selection into the program through the use of "instrumental variables," that is, variables that explain program participation independent of factors that affect birth outcomes. The instruments they used to correct for selection bias were the following: state WIC program food expenditures per capita (intended to measure the variation in the availability of WIC across states, where more WIC food expenditures per capita are hypothesized to lead to greater WIC participation); a variable capturing whether the mother's family had any earned income (intended to serve as a proxy for eligibility for public assistance, where greater family earned income is hypothesized to lead to a lower likelihood of WIC participation); and a variable indicating whether the mother had previously participated in WIC (intended to capture knowledge of program availability).

If sample sizes are large enough and the models are well specified, that is, capable of explaining why individuals participate in the program, then each of these methods should produce similar estimates of WIC's impact.[8] These instruments are used, along with other observable characteristics of the mother, to predict WIC participation. The predicted participation is then substituted for actual participation in the statistical models to measure the impact of WIC on birth outcomes. As Rossi (1988, 58) warns, however, "The success of this approach is heavily dependent on having good measures that relate to participation." Unfortunately, this appears to have been a problem in the Gordon (1993) analysis, as her efforts resulted in widely divergent and inconsistent findings.

Unadjusted for selection bias, the estimated impact of WIC on mean birthweight was 70 grams (see Table 4). But the three selection-bias tests (with slight variations in the instruments used) produced estimates ranging from a negative 307 grams to a positive 47 grams. (Gordon [1993] first estimated the impacts by using state WIC program food expenditures per capita and a variable capturing earned income as instruments. She then added a variable indicating previous WIC participation as well. Gordon did not have faith in these estimates, so the findings can only be found in an unpublished memorandum.) The Heckit procedure resulted in estimates that WIC reduced birthweight by 371 to 89 grams. The instrumental-variables approach led to estimates that WIC lowered birthweight by 303 to 69 grams (although the latter estimate was not statistically significant). Finally, the maximum likelihood

TABLE 4: The Effect of Prenatal WIC Participation on Birthweight (grams/percentage change)

Study	*Regression Adjusted Impacts*	*Selection Bias Correction Impacts*	*Simultaneity Bias Correction Impacts*	*Selection + Simultaneity Bias Correction Impacts*
Devaney et al. (1991)	51*** to 117***	Very unstable and unrealistic estimates	11 to 60**	—
Gordon and Nelson (1995); Gordon (1993)	70*** (2.1%)	Estimation Method Heckit Model 1 –371*** Model 2 –89* Instrumental variables Model 1 –303*** Model 2 –69 Maximum likelihood Model 1 –81*** to 47*** Model 2 57***	Alternative definitions of WIC participation First 8 months 68*** First 7 months 53** First 6 months 10 Gestational cohorts 28 weeks –4 32 weeks 28 36 weeks 27 40 weeks 39 Controlling for gestational age 25	Estimation method Heckit Gestational age 28+ weeks –273* Gestational age 32+ weeks –242* Gestational age 36+ weeks –295** Gestational age 40+ weeks –414* Instrumental variables Gestational age 28+ weeks –314* Gestational age 32+ weeks –242* Gestational age 36+ weeks –314** Gestational age 40+ weeks –307 Maximum likelihood Gestational age 28+ weeks –29 Gestational age 32+ weeks 23 Gestational age 36+ weeks 24 Gestational age 40+ week 9
Brien and Swann (1997)	White Model 1 54 (1.6%) Model 2 91*** (2.7%) Black Model 1 148*** (4.8%) Model 2 117*** (3.8%)	White 2SLS Model 1 178 (5.3%) Model 2 228 (6.8%) Fixed effects Model 1 (0.3%)	—	—

(continued)

TABLE 4 Continued

Study	*Regression Adjusted Impacts*	*Selection Bias Correction Impacts*	*Simultaneity Bias Correction Impacts*	*Selection + Simultaneity Bias Correction Impacts*
		Black 2SLS Model 1 421*** (13.6%) Model 2 408*** (13.2%) Fixed effects Model 1 (4.9%***)		
Brien and Swann (1999)	White Model 1 34 (1.0%) Model 2 68** (2.0%) Model 3 82** (2.4%) Black Model 1 173*** (5.6%) Model 2 186*** (6.0%) Model 3 146*** (4.7%)	White 2SLS Model 1 –153 (–4.5%) Model 2 –65 (–1.9%) Model 3 –248 (–7.3%) Model 4 –24 (–0.7%) Fixed effects Model 1 (–0.3%***) Black 2SLS Model 1 294*** (9.5%) Model 2 195 (6.3%) Model 3 393*** (12.7%) Model 4 266 (8.6%) Fixed effects Model 1 (4.8%***)	—	—

NOTE: 2SLS = Two Stage Least Squares. Percentage reductions are calculated by authors.
*, **, and *** coefficient is significantly different from zero at the 90, 95, and 99 percentage level in a two-tailed test.

estimates were characterized as "unstable" and ranged from a reduction of 81 grams to an increase of 47 grams.

Gordon (1993, 3) rejected these large negative estimates and, instead, adopted "as a premise that it is simply not possible that participation in WIC reduces newborn birthweight or related outcomes." She did not believe that the estimates were reliable or that they should be used to draw any conclusions about the effects of WIC, even the direction of selection bias. Gordon may well be correct. It is, after all, her analysis, and in other places, she and her colleagues have provided a more fulsome justification for rejecting the findings (Fraker, Gordon, and Devaney 1995). However, given the parallels between her rejected findings and the findings of others, including, for example, those of Brien and Swann (1997, 1999) discussed below, we find Gordon's results a strong sign that failure to control for selection bias overstates WIC's impact. (See Mark Lopez 1999, 9, explaining that because Gordon's selection-adjusted results "are so consistent in sign (they mostly suggest negative program effects), and are statistically significant, the selection-unadjusted results . . . are likely biased up.")

Those who have seen them believe that a series of ongoing analyses by Michael Brien and Christopher Swann (1997, 1999), conducted while they were both at the University of Virginia, provide the best window into the uncertain effects of selection bias. Like Gordon and Nelson (1995), they used data from the live birth sample of the NMIHS. Unlike Gordon and Nelson, however, they added the infant death sample of the NMIHS and further restricted their sample to non-Hispanic Whites and non-Hispanic Blacks. They selected WIC participants (the treatment group) and eligible nonparticipants (the comparison group) if they had family incomes below 250% of the poverty guidelines.

Brien and Swann (1997, 1999) estimated WIC's impact on birthweight using a basic regression model with a variety of demographic, socioeconomic, and behavioral variables. Originally, they analyzed the impact of WIC for Whites and Blacks separately, so that they could focus on "homogeneous" groups. We asked them, however, to pool their sample of Whites and Blacks, so that we could compare estimates of the program's impact for the population as a whole using their approach to the estimates derived from the WIC evaluations described above. (There were too few Hispanics in the sample for a special analysis.) Their (M. Brien, personal conversation with Peter Germanis, 1999) pooled analysis (unadjusted for selection bias and not presented in their paper) suggested that WIC increased birthweight by 2.3%. There were striking differences by subgroup. The results suggested that WIC increased birthweight by 4.7% to 6% for Blacks and by 1% to 2.4% for Whites, although the lower bound estimates were not statistically significant.

Brien and Swann (1997, 1999) then attempted to correct for selection bias by using a variety of state-level instruments reflecting the generosity and availability of the WIC program in each state, which in turn may affect the chances of program participation. (These variables include state eligibility policies related to the self-declaration of income and use of income allowances or exemptions, the imposition of brand restrictions on food that can be purchased, the use of AFDC eligibility for WIC eligibility, and the first trimester hemoglobin cutoff used for nutritional risk. Other variables used include the number of WIC clinics per 1,000 poor people, the number of WIC clinics per 1,000 square miles in the state, and the generosity of each state's welfare benefits.) They used Two Stage Least Squares (2SLS) to correct for selection bias, an approach equivalent to the instrumental variables method used by Gordon and Nelson (1995). The result was to make WIC's estimated overall impact no longer statistically significant. For Blacks, however, estimated birthweight impacts increased to between 9.5% to 12.7% in two of the models they used (although the findings were positive but not statistically significant in two other specifications). The impact for Whites, on the other hand, became negative and statistically insignificant in all four specifications.

Brien and Swann (1997, 1999) hypothesize that their instruments might work better for Blacks than for Whites. That is possible, but in the context of research on the impact of programs similar to WIC and what we know about the national WIC program, we believe it is more likely that the program is simply more effective for some subgroups than for others.

Thus, at least two sets of analyses, one by Brien and Swann (1997, 1999) and one by Gordon (1993), made elaborate attempts to control for selection bias. In both cases, WIC's estimated positive birthweight effects disappeared for the population as a whole, suggesting that failure to control for selection bias overstates WIC's impact on birthweight (and probably other related outcomes). Once again, it bears noting that Brien and Swann's findings also suggest that particular subgroups within the WIC population benefit from the program (perhaps substantially).

SIMULTANEITY BIAS

The longer a woman is pregnant, the more likely it is that her baby will be healthy. At the same time, however, the longer she is pregnant, the more likely it is that she will enroll in WIC, because she has more opportunities to learn about the program and more time to enroll. Thus, independent of any program effect, the positive birth outcomes for women who enroll in WIC late in their pregnancies are more likely due to the length of their pregnancy

than to the effects of WIC, because the program has little time to have any substantial impact, and yet, the baby is more likely to be born healthy. Failure to account for these "simultaneous" effects exaggerates WIC's estimated impact. As Gordon and Nelson (1995) explain,

> While WIC participation may increase gestational age, causality may also operate in the opposite direction—in that a longer pregnancy may increase the likelihood of WIC participation. . . . Thus, the results of the basic model are likely to overstate the effects of WIC on birthweight and gestational age, and to overstate the reductions in low birthweight births, preterm births, and neonatal and infant mortality associated with WIC participation, because a portion of the estimated positive effects reflects reverse causality. (P. 104)

Yet, many WIC evaluations have ignored the problem of simultaneity bias and, therefore, have presented estimates likely to overstate the impact of the program. Rossi (1988) comments,

> WIC effects, although generally positive for birthweight and lowering the incidence of low and very low birthweight, are not likely to be as large as calculated in earlier studies since most of those studies did not control for the simultaneity problem. (P. 59)

Those few studies that have explored simultaneity bias show that it, too, substantially overstates WIC's estimated impacts. For example, Devaney, Bilheimer, and Schore (1991) attempted to control for gestational age by adding it as an independent variable to the birthweight and Medicaid cost regression equations. Doing so provides estimates of WIC's effects on infants of a given gestational age and excludes any benefits associated with WIC's impact on increasing gestational age (to the extent that there are such impacts). As a result, estimated birthweight impacts were more than halved: falling from gains of 51 to 117 grams to gains of 11 to 60 grams. (Estimates are provided by Barbara Devaney, the principal investigator for the five-state WIC study.) Estimated Medicaid savings declined to about 40% to 45% of the original estimates, resulting in a benefit-cost ratio of –$0.87 to $2.29 (down from $1.77 to $3.13) (Devaney, Bilheimer, and Schore 1992, 591-92). (The benefit-cost ratios are based on the authors' calculations using data reported in the study.)

Devaney and her colleagues (1991) also tried a second approach: redefining WIC participation by excluding very late WIC entrants. (Women who enrolled after 36 weeks and who redeemed less than $55 in food benefits were reclassified as nonparticipants.) This resulted in much smaller reductions from the unadjusted estimates, with estimated birthweight impacts ranging from 46 grams to 112 grams (a reduction of about 5% from the

unadjusted estimates in most states) and a benefit-cost ratio of $1.38 to $3 (pp. 54, 79). (Benefit-cost ratios were calculated using the data on Medicaid savings and WIC costs.) The change in estimated effects may have been small because this approach did not entirely eliminate the bias, because women whose pregnancies reach 36 weeks still have a longer period in which they can enroll in WIC (and thus be counted as WIC participants) than women whose pregnancies end earlier.

Gordon and Nelson (1995) applied more extensive corrections for simultaneity bias. First, they controlled for gestational age by including it as an independent variable in their regression equations. The estimated impact of WIC on average birthweight for infants of a given gestational age was 25 grams, compared to the unadjusted estimate of 68 grams, although the result was not statistically significant. Second, they defined WIC participation in several ways, based on when a participant enrolled. For example, WIC participants were defined as those who enrolled in the program during the first 6, 7, and 8 months of participation. (Devaney, Bilheimer, and Schore [1991] used a 36-week definition, thereby allowing more late entrants to be counted than under some of the narrower definitions used by Gordon and Nelson [1995].) The estimated effect of WIC declines with the number of months used to define participation. For example, using an 8-month definition suggests an impact of 53 grams on mean birthweight, whereas a 6-month definition leads to an impact of just 10 grams (which is not statistically significant). Third, they created cohorts of women whose pregnancies reached a minimum number of weeks (28, 32, 36, and 40). (This differs from the second approach in that for each cohort, all women had pregnancies lasting at least the specified number of weeks. For example, for the 28-week cohort, all women who enrolled during the first 6 months are counted as WIC participants, and all those who did not are in the comparison group.) For the 40-week cohort, only pregnancies lasting at least 9 months are counted. Thus, the sample size (and the sample's representativeness of the broader WIC population) diminishes as the time frame is extended. However, for purposes of adjusting for simultaneity bias, the shorter cohorts are more important. There were no statistically significant birthweight impacts using this approach.

Thus, correcting for simultaneity bias reduced WIC's estimated impacts, but again, different approaches yielded different estimates. This uncertainty led Gordon and Nelson (1995) to present their resulting impact estimates of birth outcomes within ranges, rather than as precise estimates. For example, they estimated that WIC increased birthweight from 25 to 68 grams (approximately 1% to 2% of mean birthweight), reduced the incidence of low birthweight by 1 to 3 percentage points (from a mean of 10.8%), and lowered the percentage of preterm births by 2.4 to 3.6 percentage points (from a mean of

TABLE 5: The Effect of Prenatal WIC Participation on the Incidence of Low Birthweight

Study	*Regression Adjusted Impacts*	*Selection Bias Correction Impacts*	*Selection + Simultaneity Bias Correction Impacts*
Devaney et al. (1991)	–22%** to –32%**	—	—
Gordon and Nelson (1995); Gordon (1993)	–27%***	Univariate Probit Model Model 1 79% Model 2 –5%	Gestational cohorts 28-week –11% 32-week –13% 36-week –21%** 40-week –35%**
Brien and Swann (1997)	White Model 1 0.2 (3%) Model 2 –0.1 (–2%) Black Model 1 –5.4*** (–39%) Model 2 –3.8*** (–28%)	White 2SLS Model 1 –3.5 (–54%) Model 2 –5.6 (–86%) Fixed effects Model 1 –0.8 (–12%) Black 2SLS Model 1 2.1 (15%) Model 2 3.9 (28%) Fixed effects Model 1 –4.2*** (–30%)	—
Brien and Swann (1999)	White Model 1 0.2 (3%) Model 2 –0.2 (–3%) Model 3 –0.4 (–6%) Black Model 1 –5.4*** (–40%)	White 2SLS Model 1 3.0 (47%) Model 2 1.0 (16%) Model 3 11.6** (181%) Model 4 9.1 (142%)	—

(continued)

TABLE 5 Continued

Study	*Regression Adjusted Impacts*	*Selection Bias Correction Impacts*	*Selection + Simultaneity Bias Correction Impacts*
	Model 2 –5.6*** (–41%)	Fixed effects	
	Model 3 –3.9*** (–29%)	Model 1 –0.9 (–14%)	
		Black	
		2SLS	
		Model 1 –7.9 (–58%)	
		Model 2 –7.4 (–54%)	
		Model 3 –7.5 (–55%)	
		Model 4 –6.1 (–45%)	
		Fixed effects	
		Model 1 –4.5*** (–33%)	

NOTE: 2SLS = Two Stage Least Squares.
*, **, and *** coefficient is significantly different from zero at the 90, 95, 99 percentage level in a two-tailed test.

14.2%) (Gordon and Nelson 1995, xvii). In fact, they found that WIC had no impacts on selected birth outcomes for some of their specifications.

Two caveats are needed to interpret these corrections for simultaneity bias. First, by effectively defining away any effect that it has on lengthening pregnancies, these corrections can understate WIC's impact. As Rossi (1988, 59) cautions, "It should be emphasized that controlling for gestational age may not be entirely appropriate because doing so may obscure WIC's positive effects on increasing gestational age."

Second, these estimates are corrected only for simultaneity bias, and, hence, still overestimate the effects of WIC, as they do not also correct for selection bias. When Gordon and Nelson (1995) attempted to control for both, they found large (statistically significant) negative effects using the instrumental variables and Heckit approaches and very small (statistically insignificant) effects using the maximum likelihood approach. However, as shown in Tables 4 and 5, studies rarely correct for both. If selection bias is exaggerating WIC's estimated impact, then the range of true impacts is even lower.

GENERALIZABILITY

Another limitation of this research is that it may not be applicable to the national program, especially as it exists today. Many of the studies examine the WIC program in only one state or only a few states. Even if their findings are valid, it is an open question whether those states' experiences are representative of the nation as a whole. In fact, those studies that look across states often find impacts that vary widely from state to state. Consider, for example, the five-state study conducted by Devaney and her colleagues. The estimated infant mortality rates associated with prenatal WIC participation were an almost incredible 45% lower in three of five states studied (Devaney and Schirm 1993). However, there was no statistically significant impact in Minnesota, and in South Carolina, the estimated reduction was so large (from 36.9 to 8.7 infant deaths per 1,000 live births, a reduction of 76%), that one of the authors of the study "expressed the opinion that the calculated effects . . . were likely overestimated" (see Rossi 1988, 167). Moreover, Gordon and Nelson (1995, xvii), using a similar methodology but a national data set, could find no impact on infant mortality.

In addition, all of the studies are based on data collected more than a decade ago. (Some of the research on WIC's nonprenatal components is based on data that are almost two decades old.) Since then, there have been important changes in the program and in the size, composition, and characteristics of

the WIC population, all of which could affect the applicability of earlier research. Rossi (1988) points out,

> When programs expand, the characteristics of participants change, a shift that might affect effectiveness estimates. As WIC expands to cover a larger proportion of poor pregnant women, the women newly brought into the program may not have the same level of concern for the health of their fetuses or may not have as serious nutritional deficiencies. (P. 90)

(As we described in Part 3, it appears that as WIC has grown, application of the nutritional-risk criteria has become less rigorous.)

Moreover, many of the prenatal findings are restricted to those mothers receiving Medicaid, which in 1988, represented only 46% of the caseload. Because these were presumably the most disadvantaged families in WIC, the findings may not be applicable to the non-Medicaid, and hence somewhat higher income, families on WIC. Indeed, one study reported that WIC increased birthweight by an average of 106 grams for Medicaid recipients compared to only 62 grams for those not receiving Medicaid (Gordon and Nelson 1995, 19). Thus, as Devaney, Bilheimer, and Schore (1992) explain in relation to their own study, the changes in Medicaid and WIC over time make it difficult to generalize from earlier studies to today's WIC program:

> Since the analysis period of the WIC/Medicaid study (1987), major changes have occurred in the Medicaid and WIC programs and in the environments in which these programs operate. Thus, the long-term stability of the study results is an important issue. Higher Medicaid income-eligibility ceilings for pregnant women, in conjunction with increased coordination between the Medicaid and WIC programs, means that a higher-income group of women is likely to participate in the WIC programs. If prenatal WIC participation is more beneficial for lower-income women, then the benefits of prenatal WIC participation observed in 1987 may be greater than what would be observed under the current Medicaid income-eligibility standard for pregnant women of 133% of the federal poverty level. On the other hand, aggressive outreach, streamlined eligibility procedures, and the growing problem of substance abuse may bring a higher-risk group of pregnant women into both the Medicaid and WIC programs. The net effect of these changes is uncertain. (Pp. 591-92)

Finally, the world has changed. At the bottom of the income scale, mothers are now more disorganized and subject to greater drug problems, so helping them may be even more difficult than it was to help their earlier counterparts, as Devaney, Belheimer, and Schore (1990) observe:

> In particular, the increase in alcohol and drug abuse among pregnant women—especially cocaine and crack—has become a major public health policy problem. While much of

> the information on this issue has been anecdotal, recent studies suggest that the number of newborns exposed to drugs is increasing dramatically and that low-birthweight rates are thus rising. (P. 63)

ACCOUNTING FOR RESEARCH WEAKNESSES

As we have seen, past research is plagued by problems of selection bias, simultaneity bias, and generalizability. There is no objective way to adjust for generalizability problems, but the impact of selection and simultaneity bias can be illustrated. Tables 4 and 5 summarize the findings of the key birth-outcome studies. For the tables, we have not second-guessed the researchers involved. If they assert that steps were taken to deal with selection and simultaneity bias, we accept their estimates for the purpose of presentation. As the tables demonstrate, both selection and simultaneity bias are potent factors that must be considered in assessing any research on WIC. It is on this basis that we reach the judgements described next.

6. DOES WIC WORK?

Studies of WIC's impact are almost entirely nonexperimental in nature and, so, are based on statistical comparisons made between those who received WIC benefits and those who did not. As a result, many are subject to severe problems of selection and simultaneity bias. Nevertheless, many have been conducted by first-class researchers, and even their corrected impacts are usually positive, if often modest. (Compare with Rossi 1988, 90, "The research in question is of high quality, and results from the several independently conducted studies are largely consistent.") How, then, should this body of research be assessed?

MODEST IMPACTS, AT BEST?

Careful researchers often present their findings about WIC's impact as ranges. At the risk of great oversimplification, we have done the same for all the major studies described in this paper. Doing so places WIC's possible impacts on infant mortality, prematurity, and birthweight on a range from "zero" to "substantial." For infants, children, and postpartum and breastfeeding mothers, the only impacts seem to be small to modest effects on anemia

and nutrient intake. And as we saw, the impacts on anemia could, at least in part, be explained by changes in the nutrient content of many foods.

In addition, there is also a glimmer of evidence that WIC's beneficial effects are concentrated among the most needful recipients. Several studies suggest that WIC's impacts are greatest among the most needful. Rush, Leighton et al. (1988, 484-511) reported that the strongest dietary effects were for the least advantaged children, including those who were very poor and from female-headed households. Gordon and Nelson (1995) also found that impacts on birthweight were greater for women receiving Medicaid than for those not receiving Medicaid. Women who received Medicaid had lower average incomes than those who did not. Finally, Brien and Swann (1999) concluded that WIC had larger impacts on birthweight for Blacks than for Whites. A comparison of their demographic characteristics suggests that the Black women were poorer and more likely to be on welfare. Although these studies do not provide definitive proof that WIC is more effective for the most needful, they are at least suggestive.

Thus, based on the research, about the most that can be said is that WIC probably has "modest" impacts on some of its target populations, depending on a host of demographic and community factors, and the characteristics of the particular local program.

- WIC probably makes at least a small improvement in the diets and behaviors of some pregnant women, especially the most disadvantaged, and that in turn may improve the birth outcomes for some number of infants.
- WIC probably increases the nutritional intake of some infants, especially those who would not have been breastfed, but the health consequences of the increases are not clear. Moreover, WIC may reduce breastfeeding, which can have negative health consequences.
- Overall, WIC probably makes little significant difference in the diets of 1- to 4-year-olds, but these average impacts may miss important effects for subgroups, especially for those children whose intake of nutrients might otherwise be considered inadequate.
- As the summary in the appendix suggests, existing WIC research, at least when read in the most favorable light, provides some (and perhaps substantial) support for the proposition that WIC has socially and policy significant impacts on particular subgroups of participants. The research has not clearly established the makeup or identity of these subgroups, but the subgroups seem to be composed of the most needful families (in some substantial measure, the equivalent of the poorest of the poor).

These conclusions are based on data that are almost all from more than a decade ago and from research involving a limited number of state programs. So

that even this constrained conclusion is not necessarily applicable to the program as it exists today.

Others have similarly concluded that WIC's overall impacts are likely modest, even in regard to what many see as WIC's most successful component, that for pregnant women. Brien and Swann (1999, 9), for example, conclude, "The existing literature finds mostly consistent evidence that the WIC program has a small positive effect on birth outcomes." Goldenberg and Rouse (1998) characterize WIC in the same way:

> The Special Supplementation [sic] Program for Women, Infants, and Children, which provides a calorically enriched diet to low-income pregnant women, has been in operation in the United States for more than 20 years. Studies of this and other caloric-supplementation programs in developed countries suggest that they result in small increases in birth weight. (P. 315)

Our judgment is that WIC probably does make "modest" improvements in the health and well-being of at least some groups of disadvantaged children and their mothers. And as described above, its effectiveness probably depends on a host of demographic and community factors, and the characteristics of the particular local program.

What does this reassessment suggest about WIC's cost-effectiveness? First, we have found no responsible authority that claims that 90% of the WIC program for infants, children, and postpartum and breastfeeding mothers is cost effective. And if WIC's impacts on pregnant women and their newborns are only "modest," then it is unlikely that even that component of the program could pass any real test of cost effectiveness.

Furthermore, when Devaney and her colleagues (1992, 591-92) just took simultaneity bias (and not selection bias) into account, estimated Medicaid savings fell 60% from the original estimates, resulting in a benefit-cost ratio of –$0.87 to $2.29 (down from $1.77 to $3.13). Likewise, Brien and Swann (1999) found that controlling for selection bias (but not simultaneity bias) resulted in smaller benefit-cost effects than those estimated by GAO. They concluded, "The program is not cost-effective for whites" (p. 26). (Their results for Whites suggest that WIC saves $0.64 for each dollar spent.) However, for Blacks, their findings suggested a savings of $1.57 to $3.50 per dollar spent on WIC, but even these findings were based on estimated reductions in the incidence of low birthweight that were not statistically significant.

Taken together, then, these studies strongly suggest that WIC's overall cost-effectiveness, even for the prenatal component, is considerably less than the 3:1 ratio estimated by GAO and cited by advocates, and is quite possibly negative. WIC may be cost-effective for specific subgroups, however.

A NEW GAO ESTIMATE?

As described above, a 1992 GAO report estimated that prenatal participation in WIC reduced low birthweight births by 25% and very low birthweight births by 44%, so that each dollar spent on WIC for pregnant women saved $3.50 (over an 18-year period) in medical and disability costs. This GAO report legitimated the claims of advocates that WIC worked. After all, if the GAO said so, then WIC really must be a success.

WIC partisans use these GAO estimates on pregnant women and their newborns as if they were undebatable, and as if they applied to the entire WIC program. Here is a 1998 quotation from the prepared remarks of Shirley R. Watkins (1998), Under Secretary for Food and Nutrition and Consumer Services, who should have known better:

> We can still say that for every $1 in WIC, $3 is saved in Medicare [sic] costs. These are numbers that people can relate to and understand in very real terms, and we should use every opportunity to tell people about how successful this program is, not just for mothers and babies, but for everyone.

A more subtle effect of the GAO's seal of approval has been to freeze the program in time. After all, goes the thinking, if the program works, why spend money studying its impact, and why engage in the politically difficult task of programmatic improvement.

The GAO should conduct a new cost-benefit assessment of WIC. When the Bush Administration's USDA reviewed a prepublication version of the GAO report, it raised concerns about GAO's methodology and suggested that its estimates may have exaggerated WIC's impact. Here is the GAO's (1992) summary of the USDA's comments, many of which echo points made in this article:

> USDA expressed concern that we overestimated cost savings attributable to WIC because certain of our model's assumptions might not hold. They were concerned that our model did not capture all of the program's positive and negative effects and depended upon evaluations that themselves may not have accurately separated program effects from all measured and unmeasured differences between WIC and non-WIC populations. They also said that the calculated reduction in low birthweight rates due to WIC based on the evaluations we used might not represent the true national effect of WIC, and that the evaluations used to develop the effect size were dated in that characteristics of the WIC population had changed. (P. 43)

Unfortunately, these concerns seem to have been discounted as being ideologically based and were lost in the din of support for WIC.

Whatever the wisdom of the GAO's initial judgement about WIC's cost effectiveness, our review of the research strongly suggests that it is time for the GAO to reconsider its estimates and, in particular, that it makes the appropriate adjustments for selection bias, simultaneity bias, and lack of generalizability—or that it concludes, as it did in regard to Head Start research, that no reliable estimate can be made (U.S. General Accounting Office 1997).

This would be consistent with GAO's evaluation criteria for other programmatic research. For example, GAO's 1997 assessment of Head Start research rejected the findings of many studies because of methodological problems, particularly selection bias and lack of generalizability.

> One of the most serious of the methodological problems was noncomparability of comparison groups. The most reliable way to determine program impact is to compare a group of Head Start participants with an equivalent group of nonparticipants. The preferred method for establishing that the groups are equivalent at the outset is to randomly assign participants to either the Head Start group or the comparison group. . . . In most cases, researchers matched participants on one or more demographic variables, usually by including some variable related to socioeconomic level. In other cases, researchers did not match treatment and comparison groups but tried to compensate statistically for any inequality between the groups. Neither of these methods compensates for lack of random assignment to group. (P. 10)

GAO's conclusion was that as a result of these, and other, weaknesses, the research provided little information on the program's impact. Applying the same rigorous standards to the WIC research might lead to a different conclusion than the agency reached in 1992.

In addition, the GAO should give consideration to altering the manner in which it calculated the savings estimate. The GAO used an unsupported assumption that all of the added costs of low birthweight births can be avoided for each such birth averted. If WIC's real effect is to raise birthweights from just below the 2,500 gram cutoff point for being classified as "low" birthweight to just above it, the savings were probably exaggerated, because such births would generally result in higher costs than normal birthweight births.

Researchers at RAND (1998) describe how sensitive savings are to changes in birthweight:

> Interventions that simply shift very low birthweight infants into higher weight categories can save substantial amounts. An increase of 250 grams saves an average of $12,000-$16,000 in the first year. An increase of 500 grams generates savings of $28,000.
>
> However, not all increases in birthweight produce cost savings. At the lowest birthweights, added weight can increase treatment costs because the infant is likely to live longer and require a lengthier hospitalization. Thus, an increase in birthweight from less

> than 750 grams to between 750 and 999 grams would actually increase treatment cost by $29,300; an increase to the 1,000 to 1,249 gram range would increase costs by $13,500.

If possible, the GAO should use more fine-grained estimation methods that recognize these complexities. It might also state its estimates as a range, as do many other researchers in this field.

Finally, the GAO should recognize the diversity of WIC's caseload and examine the possible variation in benefit-cost ratios for different subgroups. A recent report by Lynn Karoly and her colleagues (1998) at RAND estimated the economic impact of a nurse home visitors program. Their findings suggest that the program produced large net savings from serving higher risk families (single mothers and low socioeconomic status) but that there were no net savings associated with serving lower risk families (mothers who were married and of higher socioeconomic status):

> For the . . . higher-risk families of the Elmira PEIP, our best estimates of the savings to government are much higher than the costs. . . . In the case of the lower-risk participants of the Elmira PEIP, the savings to government are unlikely to exceed the costs. In fact, our best estimate of the net savings is that they are negative: The government savings, while positive, are not enough to offset program costs. This result illustrates the importance of targeting programs to those who will benefit most if the hope is to realize government savings that exceed costs. (Pp. xviii-xix)

ADDRESS OVERWEIGHTNESS DIRECTLY?

Deciding whether WIC works requires identifying the problem it should be addressing. When WIC was being planned almost 30 years ago, "hunger and malnutrition constitute[d] a national emergency," in the words of the 1969 White House Conference on Food, Nutrition, and Health (Rush, Horvitz, et al. 1988, 389-90). Experts disagree about the extent of hunger—the issue has long been politically contentious—but few doubt that it was a serious problem in the past and that it is much less serious now.

Conversely, for the past three decades, there has been a steady increase in the prevalence of overweightness among Americans. Today, about one third of adults are considered overweight, up one third from the 1970s (Kuczmarski et al. 1994; Borrud, Enns, and Mickle 1996). More than one fifth of all children are now considered overweight (Mei et al. 1998; Ogden et al. 1997; Troiano et al. 1995).

Overweightness appears to be most serious among the poor and those in lower socioeconomic groups (Galuska et al. 1996; U.S. Center for National Health Statistics 1990). For example, 33% of adults reported being overweight, whereas only 4% of households reported having one or more adults

who repeatedly experienced reduced food intake that resulted in "the physical sensation of hunger."[9] Only 13% of poor households reported "hunger," even with this very broad definition (Hamilton et al. 1997, viii-ix). Similarly, about 22% of children ages 6 to 17 were overweight, compared to just 1% of children younger than age 18 who reported hunger (Hamilton et al. 1997, 48).

Recently, Agriculture Secretary Dan Glickman (1998) commented on the implications of this striking change from "too little" to "too much" food:

> The simple fact is that more people die in the United States of too much food than of too little, and the habits that lead to this epidemic become ingrained at an early age.
>
> Everyone here knows the statistics: Obesity and overweightness affect 10 million U.S. children. That's a record, and there's no real sign that it won't be broken again soon. In the past 20 years, the number of obese children has doubled, placing more Americans at risk of high cholesterol, blood pressure, heart disease, diabetes, arthritis and cancer—all at an earlier age.
>
> Obesity contributes to 300,000 deaths each year. That's close to 1,000 lives lost each day at a cost to our health care system of $70 billion a year, or 8% of all medical bills.

These trends are reflected among WIC recipients. Barbara Devaney (1997), using data from the third National Health and Nutrition Examination Survey (NHANES 3), a national study of the health and nutrition status of the civilian population conducted between 1988 and 1994, found that a substantially larger proportion of WIC children and women are overweight compared to the general population. Devaney writes, "Between one quarter and one half of low-income [WIC] eligible individuals selected by even fairly generous cutoff value are at risk of being overweight" (p. 25). (They are also less likely to be underweight.)

Obesity is linked with many adverse health outcomes, including increased mortality, coronary heart disease, hypertension, diabetes, gallbladder disease, osteoarthritis, and some cancers (Troiano et al. 1995). In children, it can lead to serious health problems in later life. According to physicians at the Children's Hospital at Dartmouth, who conducted an extensive review of the long-term health consequences of early behaviors,

> Childhood and adolescent obesity is increasing in prevalence, and obesity has profound medical and psychosocial consequences for children and adolescents. Obese children are at greater risk for hypertension, respiratory disease, diabetes, and various types of orthopedic problems, particularly slipped capital femoral epiphyses. Individuals who are obese as adolescents experience approximately twice the rate of mortality on long-term follow-up compared with those who are not obese. Morbidity from causes as diverse as coronary artery disease, colorectal cancer, gout, and arthritis is increased in cohorts that are obese as adolescents. Obesity in childhood increases the risk of subsequent morbidity regardless of whether obesity persists into adulthood. Psychosocial morbidities of ado-

> lescent obesity can be profound. In a recent U.S. study, young people 16-24 years of age who were obese completed fewer years of school, were less likely to get married, had lower household incomes, and were more likely to be living in poverty than those were not obese when followed-up after 7 years. (Hedberg, Bracken, and Stashwick 1999, 141)

The WIC program is not oblivious to the dangers of obesity and overweightness to low-income Americans. Indeed, overweightness is one of the nutritional-risk criteria used in determining eligibility for benefits. However, the plain fact is that the program's basic orientation is toward increasing food consumption, albeit healthier consumption. Consider, for example, current WIC regulations that, tracking the underlying legislation, describe the "general purpose and scope" of the WIC program as follows:

> Section 17 of the Child Nutrition Act of 1966, as amended, states in part that the Congress finds that substantial numbers of pregnant postpartum and breastfeeding women, infants and young children from families with inadequate income are at special risk with respect to their physical and mental health by reason of inadequate nutrition or health care, or both. The purpose of the Program is to provide supplemental foods and nutrition education through payment of cash grants to State agencies which administer the Program through local agencies at no cost to eligible persons. The Program shall serve as an adjunct to good health care during critical times of growth and development, in order to prevent the occurrence of health problems, including drug and other harmful substance abuse, and to improve the health status of these persons. (7 CFR 246.1)

If combating overweightness were an expressed element of WIC's mission, the USDA would be encouraged to manage toward that objective, and evaluations of the program would be more likely to assess WIC's impact on obesity. That is why we conclude that efforts to combat overweightness should become an express element of WIC's mission.

7. PROGRAMMATIC FLEXIBILITY

How should responsible policy makers respond to the fact that WIC's impacts are probably modest, at best. Certainly not by simple-mindedly cutting or abandoning the program. The problems that WIC addresses are serious and require attention. In an age when so many government programs for the poor seem to have no effect, and may even make things worse, WIC's possible beneficial effects should not be slighted. But just as certainly, the program should not be continued without trying to improve it. Nutrition-related health problems are still serious and widespread, although they now also involve high levels of overweightness. If WIC can be made more effective, we should strive to make it so.

Unfortunately, no one really knows how to make WIC more effective. Partly because it has been so politically incorrect to acknowledge the limitations of past research, the necessary studies and planning have not been conducted. Thus, a good way to begin would be to open the program to innovation and experimentation.

WIC was established almost 30 years ago, and since then, its essential shape has been frozen in place. State and local officials are relatively straight-jacketed by various federal rules, even though they are presumably in the best position to judge whether various program adjustments are needed. Simply put, local health departments, hospitals, public health clinics, and community health centers should have more freedom to try new approaches.

The mechanism for giving them greater programmatic flexibility could be direct statutory grants of discretion or a legislatively authorized, case-by-case waiver process. For example, the National Association of WIC Directors has called for such "regulatory waivers for the purpose of testing alternative service delivery projects and food prescriptions."

Here are just a few of the program variations that should be allowed. Any changes should be fully and rigorously evaluated, as described below in Part 8. There is no certainty that the ideas will work, so each is expressed as a question to emphasize its tentativeness.

TARGET WIC BENEFITS TO MORE NEEDFUL FAMILIES?

As described above, WIC participants include nearly half of all infants, one quarter of all children ages 1 to 4, and the same proportion of pregnant women. Because household incomes vary from state to state, in six states, as many as 60% or more of all newborns receive benefits (authors' calculations, based on USDA and census data). About 50% of all infant formula sold in the United States is purchased with WIC dollars (U.S. General Accounting Office 1998, 1).

Implicit in current coverage of the overall WIC program is the judgment that nearly 40% of all American children are at nutritional risk, and that their mothers need corrective counseling. That cannot be right.

Essentially, WIC funding has increased far in excess of what can be effectively used under current eligibility rules. (Spending rules are also part of the problem, as described next.) Our analysis in Part 3 suggests that recent WIC expansions have primarily been among those with higher incomes and relatively less nutritional risk. As researchers at the Urban Institute pointed out in 1994, "As the Program expands, much of the new caseload will come from

increases in participation by children and postpartum women, the two categories with lowest priority now" (Ku, Cohen, and Pindus 1994, 8).

We believe that this eligibility creep has gone too far and that there should be a cutback at the upper levels of income eligibility. Actually, there should be a top-to-bottom reevaluation of WIC's eligibility criteria. According to a committee of the Institute of Medicine (1996, 8), current funding priorities may force states to deny access to some individuals to serve others who are in less need of the program's benefits.

Allow states to target more WIC resources to the most needful families. The funding that now goes to lower risk recipients could be redirected to make improvements in the basic program for more needful families, primarily by either intensifying services or expanding their duration. Both are discussed below.

SELECTIVELY INTENSIFY WIC BENEFITS?

As the research discussed in this article suggests, WIC's positive effects are probably concentrated among its most disadvantaged recipients. As Part 2 describes, the result of federal rules is essentially to require that all recipients in the same target group receive basically the same set of WIC benefits, notwithstanding their differences in need. States and local agencies administering the program cannot make real changes in benefits to meet the palpably different needs of clients. Remember that WIC agencies are prohibited from expanding the amount of the food package.

Why these limitations? They were perhaps understandable when the program was much smaller, and funding was limited. But surely, a family on welfare might need more food than a family at the top of income eligibility, which can exceed $50,000 for a family of eight. About one third of WIC families have annual incomes of more than $25,000; they have very different (and lesser) needs than those with incomes less than $10,000. Why not allow states, at least on an experimental basis, to increase the size of the food package for the most needful families?

And just as surely, $11 per month for nutritional counseling and education (as well as for administrative costs) is insufficient to meet the needs of some families. For example, enhanced nutrition education services could be directed to families with obese children or children who have a predisposition to cardiovascular disease or high cholesterol, or for parents with a drug or alcohol abuse problem.

Making WIC counseling sessions more relevant would involve major, and expensive, programmatic changes. A comprehensive review of nutrition education programs for pregnant women and mothers with infants, led by Isobel Contento (1995) of Columbia University and published in a special issue of the *Journal of Nutrition Education*, concluded, "Such interventions are effective when they focus on the pregnant woman's specific needs and provide follow-up sessions to reinforce and maintain behavioral changes" (p. 331). For example, the study reported gains in maternal and infant weight from nutrition education classes that "were augmented with multiple individual counseling sessions" (p. 331).

Janet Schiller, the USDA project officer for many WIC studies, and Mary Kay Fox (1999), a longtime WIC researcher at Abt Associates, describe suggestions from the WIC community for improving the program, including using incentives to change behavior, giving clients greater choice in the selection of classes, improving communication with health care providers, enhancing training opportunities for nutrition educators, and including dads and grandmothers. In our own visits to WIC centers, we have been told of the need for more cooking demonstrations, with special attention to the needs of recent immigrants.

Local agencies are legally authorized to implement many of these ideas, but WIC's current spending restrictions plus federal mandates on the content of services (such as voter registration) constrain their ability to do so. Although states can theoretically vary the amount they spend on counseling per participant, total spending for "nutrition services and administrative" costs is capped. Because all WIC recipients must get a minimum level of services, less than the $11 cap, there is not enough funding to intensify services in any meaningful way.

Allow states to spend more on WIC benefits and services. Instead of adding more people to the rolls, it might make sense to change WIC's rules to allow local agencies to provide more food benefits and educational services to poorer families who palpably need more aid than those at higher incomes. Once again, these expansions need not, and in fact should not, be across the board. States do not need a larger, but still one-size-fits-all, program.

ADD A FOCUS ON PREVENTING OVERWEIGHTNESS?

A number of factors seem to have increased overweightness. The most prominent is a decline in physical activity, poor diet choices, and (for adults)

a decrease in smoking (Galuska et al. 1996, 1729; Kuczmarski et al. 1994, 209). Most experts also believe that a good, if not the best, preventive approach is to develop more healthy habits among children and adolescents. Although many individual WIC agencies and staffers work hard to combat overweightness among WIC participants, a more sustained focus would require adding the prevention of overweightness to WIC's mission.

Right now, despite the efforts of many individual WIC agencies and staffers, the overall program is largely irrelevant to these concerns. Agriculture Secretary Glickman (1998) has observed,

> Currently, for example, there's nothing in the WIC program that says anything about physical activity, even though this is the number 1 reason for the rapid rise in childhood obesity. Through WIC, we encourage parents to stop smoking, to get their children immunized, to eat healthy. We also should encourage active lifestyles. I've asked Shirley [Under Secretary for Food and Nutrition and Consumer Services] and her staff to take a formal look at all our nutrition programs to see if there's a way to link diet and exercise and address the whole problem, instead of simply the food angle.

Add a focus on preventing overweightness to the WIC program. Shifting WIC's focus will be no simple matter. Researchers at the CDC, concerned about the increasing prevalence of overweightness among 4- to 5-year-old children, emphasize the importance of prevention activities in the preschool years, such as

> encourage physical activity to maintain a healthy weight, eating at least five servings of fruits and vegetables per day, and after the age of 2 years gradually decreasing dietary fat to a level of no more than 30% of energy. (Ogden et al. 1997)

However, these are the very ages when WIC participation drops off sharply.

Not only will counseling need to be expanded and reoriented, but the very foundation of the program may have to be reconsidered; but this will be a tricky tradeoff. A major reason why mothers participate in WIC is for the food package. But the package is designed to encourage consumption, albeit more healthy consumption. Deny them the package, or make it less attractive, and participation in the children's program may fall off even more.

SERVE CHILDREN OLDER THAN AGE 4?

Current rules cut off WIC eligibility when a child reaches 5 years of age. Nutritional needs may change as children mature, but they do not end. In fact, in regard to overweightness, they increase. As described above, over the past three decades, there has been a steady increase in the degree of over-

weightness among children and adolescents, which is associated with later overweightness and various health problems in adulthood.

Many researchers believe that school-age children should receive nutritional counseling because

> behaviors, or the predispositions and skills to enact them are important. . . . And chronic disease processes begin early. Risk factors for chronic disease tend to begin in youth, so that those at the high end in terms of total cholesterol, blood pressure, or weight maintain their ranking in relationship to their peers over time. (Contento 1995, 298-99)

Why not have WIC provide this counseling and education? WIC's age cutoff is usually defended on the grounds that the school-meals programs already address the needs of school-age children. But these programs are designed to increase food consumption, when the problem for many children is too much consumption or consumption of the wrong foods. In fact, to address this serious problem, various programs have been developed for specific communities (see Contento 1995).

Allow states to experiment serving children older than age 4. Why not engage the energies and resources of WIC for this effort? Rather than continue to raise income limits yet higher, consideration should be given to keeping poor children in the program longer. Not only will those older than age 4 benefit, but so might 3- and 4-year-olds, because one would hope that additional efforts would be made to make the program more relevant to their needs.

Longer participation in WIC might make it more effective in meeting its basic objectives. As noted above, the National WIC Evaluation found that the dietary status of past WIC participants was no different than the comparison group that did not receive WIC at all. In other words, not only does the food package end, but the lessons of WIC apparently did not make a permanent change in parental behavior or the child's eating habits. More time in the program might make a more lasting impression, and even if it did not, dietary behavior would at least be better for a longer period of time.

INCREASE DIRECTIVE COUNSELING?

WIC is largely "based on the knowledge-attitudes-practice paradigm," according to Schiller and Fox (1999). In fact, parents cannot even be required to attend WIC's basic nutritional counseling, although some may not know that.

A review of the research on nutrition education programs, including WIC, suggests that these programs "were not very effective in bringing about

behavioral change" and that "the use of a systematic behavioral change process is most likely to be effective in bringing about changes in behavior" (see Contento 1995, 356-57). This suggests changing the paradigm used by WIC from providing information to motivating changed behaviors. According to Schiller and Fox (1999), summarizing the research of Contento (1995),

> Effective programs are: behaviorally focused; based on appropriate theory and prior research; interactive; aimed at facilitating voluntary adoption of desired behaviors; and based on identified needs, perceptions, motivations and desires of the target audience.

A group of projects funded by the CDC suggests how much WIC might be able to change parental behavior, in this case raising immunization rates. In New York City, for example, six volunteer WIC sites were randomly assigned to one of three immunization interventions designed to increase vaccination coverage among preschool children who were eligible for measles vaccination (Burkhead et al. 1995). In two sites, WIC parents were escorted to a clinic in the same location where children could be immunized in an "express lane"; in two other sites, parents were required to return monthly for WIC coupons, rather than every 2 months (the normal schedule), until they had the child immunized; and in the final two sites, parents were offered a vaccination assessment and education and referral services ordinarily offered in the WIC office. The children at the escort sites were five times more likely to be immunized than children at the referral sites. Similarly, the children at the sites where parents could be required to return each month for their coupons were three times more likely to be immunized. In short, WIC benefits have been used to change recipient behavior for the better.

Federal rules prohibit tying WIC benefits to the acceptance of nutritional instruction or engaging in any other responsible behavior. Pregnant smokers cannot even be required to attend smoking-cessation classes, despite the well-known connection between smoking and low birthweight and other adverse birth outcomes. And yet, to the extent that WIC was actually successful in the early studies described above, one explanation could be the nutritional and health counseling that pregnant mothers receive (however inadequate it may now seem), especially given the relatively small size of the food package for all but newborns. And that may be the result of the directive or authoritative content of WIC's counseling. (Remember that if the program staff determines that recipients suffer from high cholesterol, they can issue vouchers that can only be used for low-fat foods.)

Allow WIC counseling to be more directive. WIC participants could be required to attend life skills training and parenting classes, as a condition of

receiving their food voucher. Smokers could be required to attend smoking cessation classes. Or they could be required to have their children immunized, perhaps along the lines of the immunization projects described above.

States could be permitted to experiment with such behavior-related rules, in the same way that they can now condition welfare on specific behaviors (see Mead 1997). With notable successes, many states have imposed various behavioral rules on their welfare caseloads, such as requirements to send their children to school (Long and Bos 1998), immunize them (Kerpelman et al. 1999), and attend parenting classes. Similar requirements might bring about improved outcomes for WIC participants.

Another possible approach would be to test, on a large-scale basis, an intensive nurse home visitor program. Some WIC agencies have performed limited home visiting, often using paraprofessionals funded by other programs. The most promising of these programs is the Prenatal and Early Childhood Nurse Home Visitation Program developed by David Olds (Olds et al. 1997). WIC recipients could be assigned to either WIC or WIC plus the home visitation program. Although the two programs would share some of the same services and objectives, and both groups would receive WIC food vouchers, their approach to counseling would be quite different. A comparison of outcomes would suggest which program orientation is more effective. (Because the home visitation could be considered an enhanced version of WIC's counseling component, the ethical issue of withholding benefits does not arise.)

USE ALTERNATE SERVICE PROVIDERS?

WIC is a program of the USDA, in keeping with its focus on nutrition services. But WIC works through state health departments, which in turn fund local health departments, hospitals, public-health clinics, community health centers, and so forth.

WIC was conceived in the early 1970s, long before expansion of various health services for low-income families. There are any number of other health care providers that might be appropriate WIC agencies. In three states (California, Illinois, and Oklahoma), for example, Planned-Parenthood organizations also serve as the local WIC agency (J. L. Epstein, memorandum to Phoebe Cottingham, 1999).

Encourage states to try alternate service providers. One way to coordinate health services with WIC is to co-locate WIC services with a managed care provider, made easier now by the proliferation of Medicaid managed care. Alan Kendal of Emory University has identified five states where at least one

WIC agency has co-located WIC services with managed care providers (California, Florida, Michigan, Texas, and Washington). His research in Michigan suggests that notable improvements in health indicators can be achieved by co-location (Alan Kendal, personal conversation with Peter Germanis, 1999).

Another approach that would merit serious consideration would be to integrate WIC into the actual operations of Medicaid managed-care systems. WIC went into effect before Medicaid coverage was expanded to include, among others, the same people covered by WIC, and long before most Medicaid recipients were enrolled in managed-care programs. Allowing managed care agencies to provide WIC services might avoid costly duplication of administrative services and would make it easier for WIC participants to avail themselves of the services WIC now refers them to, such as prenatal care. (One obstacle to full integration is an almost total prohibition, contained in federal law and regulations, against having for-profit organizations, such as for-profit HMOs, serve as WIC agencies.)

Managed-care providers might eagerly accept adding WIC services to their own programs, especially if they were to receive additional funding. Offering such WIC-like services could have a positive impact on pregnancy and child outcomes. An Urban Institute study reports that in "at least a couple of states managed care plans and private physicians have already asked about offering WIC services in their offices" (Ku, Cohen, and Pindus 1994, 19).

Nevertheless, we should be mindful that managed care plans have been criticized for neglecting preventative activities. Plans and providers who believe that a substantial portion of their clients will shift to other plans before the long-term costs of failing to provide preventive services are realized will be tempted to skimp on them. Hence, any integration of WIC with managed care should be pursued with caution.

8. RIGOROUS EVALUATION

Although the foregoing ideas seem consistent with research findings and common sense, they are untried and, therefore, should be carefully evaluated. Too many social programs are launched, and policies change without a full understanding of the problems they seek to address and without a rigorous evaluation to determine whether they succeed. There should be a systematic exploration of what might strengthen the program, expanding those programmatic elements that work and discarding those that do not.

Given the relatively small effects of the program overall, evaluations of program impact should devote more attention to examining impacts on subgroups and components of the program. This would then help policy makers target benefits to those groups that can most benefit from the intervention. In other areas of social intervention, the most exciting findings are often for discrete and important subgroups. For example, the positive impacts for the Elmira nurse home visitation program, described above, were primarily limited to economically disadvantaged mothers, with few significant impacts for the sample as a whole (Olds et al. 1997, 641). Similarly, Ohio's Learning, Earning, and Parenting program, an initiative providing financial incentives for teen parents to enroll and attend school, did not have a statistically significant impact on high school graduation or GED receipt for teen welfare mothers overall but did have a positive impact of 18% for those who were still in school at the time the project started (Bos and Fellerath 1997, 38).

THE LIMITS OF NONEXPERIMENTAL STUDIES

Although we believe that most WIC researchers have been sincere in their efforts to discern WIC's impacts, the plain fact is that methodological problems (such as selection and simultaneity bias) undercut their findings and make them too unreliable for policymaking. As Michael Puma, when he was with Abt Associates, and his colleagues (1992, vi) warn, "The risks inherent in the quasi-experimental design" are "results that are either obviously incorrect or, if plausible, subject to grave suspicion." Even in the absence of the state-based experimentation proposed above, therefore, the USDA should mount a systematic research effort to determine WIC's impacts.

Nonexperimental studies should not be dismissed out of hand. Often, they provide the only objective information available about a program's impact. But their findings are extremely difficult to assess when key statistical tests to determine the validity of their findings are not performed (or reported fully enough to make a judgement). Hence, nonexperimental studies should pay much more attention to methodological problems such as those identified in this article. This means testing a range of control variables and instruments, conducting the necessary tests to determine the strength of the instruments, and examining various specifications of the underlying statistical model.

Regrettably, most WIC studies simply do not provide sufficient information with which to judge the strength of their instruments and the adequacy of their statistical models. This is not the place to repeat an entire literature on the subject, but we would note the recurrent problems of (a) omitted

variables, that is, limited or poorly crafted control variables that do not capture enough data about the participants, and (b) inadequate instruments, that is, the absence of variables that explain program participation independent of factors that affect the outcomes of interest.

Greater attention to methodological rigor, however, would not necessarily result in more reliable estimates of WIC's impact. The findings of any nonexperimental study are sensitive to the sample selected, the variables considered, and the statistical models used. In addition, subgroup analysis often is complicated by the apparent differential impact of instrumental variables and other methods of controlling for selection effects.

As an illustration, consider the analyses performed by Brien and Swann (1997, 1999). They used the 1988 NMIHS to estimate the impact of WIC on Blacks and Whites. They estimated WIC's impact by varying the criteria used to select the sample, the variables included in the statistical model, and the approach to correcting for selection bias. Their research was grounded in sound theory and recognized evaluation methods, but with each modification, the findings fluctuated, sometimes considerably. For example, Tables 4 and 5 summarize their 1999 findings from eight different models. The estimated impacts on birthweight ranged from a negative 248 grams to 82 grams for Whites and from 146 grams to 393 grams for Blacks. Some were statistically significant, and others were not. Even larger variations were found with respect to the estimated impacts on the incidence of low birthweight, ranging from a negative 6% to (a highly implausible) 181% for Whites and a negative 29% to a negative 55% for Blacks. Thus, even in this very careful work, there remains a great deal of uncertainty.

Indeed, randomized experiments may be the only way to develop valid estimates of WIC's impact, because they ordinarily do not require uncertain statistical adjustments to eliminate differences between treatment and control groups.

> If properly planned and implemented, an experimental design should result in treatment and control groups that have comparable measurable and unmeasurable aggregate characteristics (within the limits of chance variation). And, from the moment of randomization, they will be exposed to the same outside forces, such as economic conditions, social environments, and other events—allowing any subsequent differences in average outcomes to be attributed to the intervention. (Besharov, Germanis, and Rossi 1997, 42)

The need for randomized experiments is becoming increasingly clear to policy makers. Random assignment was the method of choice for measuring the impact of dozens of welfare reform experiments that eventually led to the passage of the 1996 welfare reform law. It also was the strategy recom-

mended by the GAO (1997, 15) in its review of the Head Start program, outlining a strategy similar to the one we propose here.

RANDOMIZED EXPERIMENTS

As described above, many policy makers assume that it is not ethically appropriate to conduct randomized experiments in the WIC program because such high percentages of eligible recipients are already enrolled. There is, however, still room for ethical experimentation.

Random assignment for the children's program. In the early 1990s, USDA solicited proposals for a large-scale experiment designed to measure WIC's impact on children's nutritional status, cognitive development, and other outcomes. In 1992, however, Congress explicitly prohibited the USDA from undertaking the study. Because many advocates felt that the battle for full funding had been won, they may have felt that such a study was unnecessary and could only cause mischief. (After all, a rigorous evaluation of the children's portion of the program might reveal that impacts fall short of the advocates rhetoric.) The $20 million price tag may have also contributed to the study's demise. Unlike studies of the prenatal portion of WIC, which can rely on relatively inexpensive data sources (such as birth certificates and administrative records), data on children's long-term physical and cognitive development are expensive to collect.

It is still possible to initiate a straightforward randomized experiment for children, because they are not yet fully covered by WIC. In 1997, WIC's coverage of children was just 75% of eligibles. (Moreover, this may decline in the near future, as the number of potentially eligible children has been expanded by recent Medicaid expansions, which increasingly include children in families with incomes above 185% of poverty [the WIC income maximum].) It is even lower for the children in the upper end of this age group. Although we do not have estimates of the number of children eligible by age, the number of children participating declines rapidly as children grow older. For example, in 1996, 1.4 million 1-year-olds participated in WIC, compared to 1 million 2-year-olds, 900,000 3-year-olds, and only 600,000 4-year-olds (see Randall, Bartlett, and Kennedy 1998, 36). This drop in participation is surely much more rapid than the decline in the number of eligible children, suggesting that perhaps as few as one third of income-eligible 4-year-olds participate in the program. Thus, it might be possible to randomly assign 3- and 4-year olds, because the participation of eligibles is still relatively low.

Because WIC has reached full coverage among infants, and pregnant, breastfeeding, and postpartum women, it is generally assumed that a randomized experiment for these groups would be unethical to undertake, because some individuals would be denied a service or a benefit to which they would otherwise be legally entitled. But, with a little creativity, it should still be possible to mount a number of ethical randomized experiments.

Random assignment to those at the "policy margin." One way to obviate concerns about denying benefits to eligible individuals would be to study WIC's impact on those who are not now eligible—but who, in important ways, are similar to those who are. For example, the experiment could involve families with incomes slightly more than income eligibility or children slightly older than age 4. (Such experiments could be conducted, for example, in the states that have not availed themselves of the option to expand Medicaid to children with family incomes above 185% of poverty.)

Although it would be technically incorrect to apply findings from such studies to all WIC recipients, the results would be suggestive about the current program's impact and highly suggestive concerning those at the upper edges of eligibility. More important, the results would inform policy makers about the wisdom or direction of continued expansions of the program to those at the "policy margin."

These studies could be complex, randomized experiments, in which impacts are measured across various expansions in eligibility. For example, if no significant improvements were observed from extending WIC to higher income children, but measurable improvements resulted from extending it to 5-year-olds, a reconsideration of the program's target groups would be in order. It might then make sense to experiment further by testing the impact of reducing the income thresholds while raising the age limit for children's eligibility.

Random assignment to dose-response and planned-variation studies. Assuming that WIC produces positive outcomes, it is unclear which of its benefits or services are responsible. As Rossi (1988) cautions,

> Although the effects of WIC have been attributed to the dietary supplements, they are more properly viewed as the joint effects of the supplements, WIC administration, and WIC nutrition education. Because all WIC participants get all three, the effects of each cannot be estimated separately. (P. 101)

Increasing WIC's impact is best accomplished with a knowledge of which of its elements seem to have the greatest effect on recipients. That is the only

responsible basis for deciding whether the intensity of the entire program should be increased or only some element of it, such as the food packages or the nutritional counseling.

Randomly assigning individuals to WIC and to what could reasonably be characterized as enhanced versions of WIC would also obviate the ethical dilemma of denying services to eligible populations. Such randomized trials would indirectly measure WIC's effectiveness while assessing the utility and impact of various program enhancements.

Experimental designs could be used to compare the impact of WIC's nutrition education and counseling services to a more comprehensive and aggressive intervention to improve the dietary intake of participants and to modify known risk behaviors, such as smoking and substance abuse. Or there could be experiments with increasing the size or value of the food package, or replacing the WIC voucher with cash. A cash-out experiment could examine the extent to which the current WIC food voucher actually increases food consumption and nutrient intake. The 1997 review of the Head Start program by the GAO also suggested this approach for Head Start:

> Another strategy that could be used to study specific parts of the program would be to use an alternative treatment design. In this case, some randomly assigned participants would receive the full Head Start program, while others would receive partial services. For example, if the study interest is in school and cognitive issues, the control group might receive only nutritional and health services. (P. 15)

Unlike GAO, however, we recommend that the regular WIC program be compared to an enhanced version, rather than to one offering partial services. (One exception to this rule could be experiments involving those currently not eligible for the program, such as children age 5 and older. Nevertheless, such an experiment would involve an expansion of services for them.)

Random assignment to alternate health-related service providers. Finally, there is no ethical problem in varying provider of services. For the reasons described above, there is reason to believe that other service providers might be more effective in achieving WIC's goals.

Experimental designs could be used to test the relative efficacy of various health-related or family service providers. Furthermore, if large enough, the experiment could also involve a variation in services provided.

We do not mean to suggest, however, that randomized experiments would be easy to carry out. There are often daunting organizational and implementation problems. Even when ethical issues are overcome, randomization can

create hostility among service staff, thus impeding successful implementation. Larry Orr (1998, 13) of Abt Associates describes how random assignment appears to have been compromised in one WIC evaluation by program staff, who viewed denying WIC benefits as unacceptable:

> In a pilot test of an evaluation of the Women, Infants, and Children's [sic] feeding program (WIC), local staff recruited women in health clinic waiting rooms and randomly assigned them on the spot, using an algorithm based on the women's Social Security number. Proper application of the algorithm would have produced equally sized treatment and control groups. In one site, nearly two-thirds of women were assigned to the treatment group; it seems clear that recruiters falsified Social Security numbers to allow women who should have been controls to be assigned to the program. (see also Puma et al. 1991)

9. CONCLUSION

Our analysis of the research challenges the conventional wisdom that WIC is a uniquely successful program. As we have seen, many claims about WIC's effectiveness are simply misleading exaggerations.

Some have argued that the weaknesses in WIC research (and presumably WIC as well) should not be discussed, lest public and political support for the program be undermined. But unless WIC is assessed honestly, it is difficult to see how the program could ever be improved. And wouldn't that hurt poor children even more?

On the other hand, WIC provides what could be important benefits to more than 7 million individuals. In many inner-city neighborhoods, the program's $100 worth of food has become an important component of the economic safety net for mothers with newborns. And we admit to being moved by the testimony of frontline health care workers who describe the importance of WIC packages for newborns. (Besharov remembers when, as late as the early 1970s, many low-income, mothers would sell the infant formula given them so that they would have more cash. With the flood of infant formula in most poor neighborhoods, there simply is no market for the resale of infant formula, so that the formula is now much more likely to be given to the newborn.)

Based not just on this body of research but also on what is known about the impact of similar programs, we conclude,

1. Studies of WIC's impact are almost entirely nonexperimental in nature and, so, are based on statistical comparisons made between those who received WIC benefits and those who did not. As a result, many are subject to severe

problems of selection and simultaneity bias. Moreover, most studies are of limited applicability to assessing the current program, because they are based on the program as it existed more than a decade ago and thus do not reflect the composition of the caseload today.

2. WIC probably makes at least a small improvement in the diets and behaviors of some pregnant women, especially the most disadvantaged, and that in turn, may improve the birth outcomes for some number of infants.
3. WIC probably increases the nutritional intake of some infants, especially those who would not have been breastfed, but the health consequences of the increases are not clear. Moreover, WIC may reduce breastfeeding, which can have negative health consequences.
4. Overall, WIC probably makes little significant difference in the diets of 1- to 4-year-olds, but these average impacts may miss important effects for subgroups, especially for those children whose intake of nutrients might otherwise be considered inadequate.
5. WIC has expanded beyond the truly disadvantaged, even though new participants are unlikely to need or benefit from the services it provides.
6. WIC is largely irrelevant to the most serious nutritional problem facing disadvantaged Americans: overweightness.
7. Consequently, WIC does not result in the major cost savings that its advocates claim, and it may not even pass a basic test of cost effectiveness.
8. As the summary in the appendix suggests, existing WIC research, at least when read in the most favorable light, provides some (and perhaps substantial) support for the proposition that WIC has socially and policy significant impacts on particular subgroups of participants. The research has not clearly established the makeup or identity of these subgroups, but the subgroups seem to be composed of the most needful families (in some substantial measure, the equivalent of the poorest of the poor).

In the future, much greater attention should be paid to such differential effects, especially because they might suggest more focused service strategies. As Rossi (1988, 110) notes in *Feeding the Poor: Assessing Federal Food Aid*, this is one of the shortcomings of most current research: "Currently available evaluation studies place too much emphasis on central tendencies—means and medians—and do not give enough attention to measures of the distributions of responses and differentials among subgroups."

To increase WIC's positive impacts, we propose a series of possible reforms, each to be thoroughly evaluated. To emphasize the tentative nature of our recommendations, we stated them in the form of questions.

- Target WIC benefits to more needful families?
- Selectively intensify WIC benefits?
- Add a focus on preventing overweightness?

APPENDIX
WIC at a Glance

Target Group	*Commonly Reported Nutritional Risks*	*Benefits*[a]	*Coverage*[b]	*Funding Fiscal Year 1997*[c]	*Range of Evaluation Findings*
Pregnant women	General obstetrical risks, inappropriate growth or weight gain pattern, prepregnancy high weight for height, hematocrit or hemoglobin below state criteria, and inadequate or inappropriate nutrient intake.	Food: milk, eggs, iron-fortified dry cereal, vitamin C-rich juice, and dry beans or peanut butter. Services: nutrition education, and referrals to substance abuse counseling, OB/GYN care, family planning services, and other health and social services.	845,000 69% of eligibles 28% of pregnant women	$497 million Average food package: $38	Average birthweight: 0 to 4% (6% for Blacks); after correcting for selection bias: –11% to 14% (for Blacks only) Low birthweight rate: 0 to –30% (–40% for Blacks) Very low birthweight rate: 0 to –55% Preterm birth rate: 0 to –30% Infant mortality rate: 0 to –66% Neonatal mortality rate: 0 to –66% Postneonatal mortality rate: 0
Breastfeeding and postpartum women	General obstetrical risks, hematocrit or hemoglobin below state criteria, inadequate or inappropriate nutrient intake, and high weight for height.	Food: cheese, milk, juice, dried beans or peas, peanut butter, canned tuna fish, and carrots. Services: nutrition education, breastfeeding promotion, and referrals to family planning services and other health and social services.	866,000 122% of eligibles 22% of women with infants	$468 million Average food package: $34	Breastfeeding initiation and duration: insufficient evidence Pospartum women; subsequent birthweight: 3% to 4% (1 study)
Infants (0-12 months)	Infant of a WIC-eligible mother or mother at risk during pregnancy, and	Food: concentrated, liquid, iron-fortified formula (or powdered or other formula), iron-	1,863,000 122% of eligibles	$670 million Average food package: $22	Anemia: reduction; not possible to quantify Adequately immunized: 0 to

	breastfeeding mother and infant dyad.	fortified dry infant cereal, and vitamin C-rich infant juice. Services: referrals to pediatric care, immunization services, and other health services.	49% of infants	($80 prerebate)	36% (1 study) Mean nutrient intake: vitamin C (59%) and iron (32%) (1 study)
Children (1-4 years)	Inadequate or inappropriate nutrient intake, hematocrit or hemoglobin below state criteria, and high weight for height.	Food: milk, eggs, iron-fortified dry cereal, vitamin C-rich juice, and dry beans or peanut butter. Services: nutrition education, and referral to EPSDT and other health services.	3,835,000 75% of eligibles 25% of children	$2,175 million Average food package: $34	Anemia: reduction; not possible to quantify Adequately immunized: 0 to 25% (1 study) Mean nutrient intake: positive for 1/3 to 2/3 of nutrients studied, most notably iron (about 20%) (2 studies)

SOURCE: Bonnie Randall, Susan Bartlett, and Sheela Kennedy. 1998, August. *Study of WIC participant and program characteristics: 1996*, p. 85. Cambridge, MA: Abt Associates Inc.

NOTE: This table reports the most commonly reported nutritional risks, affecting at least 15% of WIC participants in 1996.

a. Food packages are tailored to meet the individual needs of participants, so the food packages described identify the foods most commonly provided for each target group. The services provided reflect those offered in WIC clinics.

b. The percentage of eligibles is reported in U.S. Department of Agriculture, "Special Supplemental Nutrition Program for Women, Infants and Children (WIC): Eligibility and Coverage Estimates—1997 Update." Because pregnant women are unlikely to participate in WIC for a full 40 weeks, their participation rate is expected to be less than 100%. For example, if all eligible pregnant women enrolled in WIC for 6 months, their participation rate would be 65%. According to USDA, the high participation rates for some groups are due to differences between the way the number of income-eligibles is estimated and the certification practices applied in local WIC agencies. In addition, some imprecision is present in any survey-based estimate. However, according to this USDA report, "These data do strongly suggest that the program has likely achieved virtually full coverage of persons in this category at the national level." The percentage of the population covered is estimated using population data from the U.S. Bureau of Census. The number of pregnant women was estimated by assuming that they equal three fourths of the number of infants, because pregnancy lasts 9 months, while infancy lasts 12 months.

c. U.S. Department of Agriculture, "Special Supplemental Nutrition Program for Women, Infants and Children (WIC): Eligibility and Coverage Estimates—1997 Update." Funding for each target group was estimated by adding the average administrative cost of $11 per participant to the average cost of food for each target group and then multiplying by the average monthly number of recipients. This is then multiplied by 12 to arrive at an annual cost.

- Serve children older than age 4?
- Increase directive counseling? and
- Use alternate service providers?

We think that reforms along these lines have a good chance of making WIC more effective. But even if they do not, that does not mean that such expansions of the program, so long as they are more carefully targeted than current services, are not socially worthwhile. Making even a small number of children, especially poor children, healthier—without harming others and without exorbitant spending—would be a sufficient ethical benefit not captured in purely economic benefit-cost calculations. As Jane Huntington and Frederick Connell (1994) write,

> We should consider whether cost savings is the appropriate criterion by which to judge prenatal care programs. It is tempting to assume that in order for these programs to be valuable, they really should save more than they cost. Yet when we require prenatal care, and other preventive health care, to pay for itself, we may be inadvertently denying valuable benefits to society. It may be better to ask not "How much does this save?" but, rather, "How much is this worth?" (Pp. 1306-07)

Hence, this article does not argue that WIC's weaknesses justify abandoning or even cutting the program. On the contrary, there should be a sustained effort to make the program more effective. This effort should start with a policy debate about WIC's role and impact coupled with a grant of greater flexibility to state and local WIC agencies to open the program to innovation and experimentation. And as described at the close of this article, any changes should be carefully evaluated. Furthermore, even in the absence of a waiver-based experimental strategy, the federal government should conduct a series of randomized demonstrations to determine more definitively the impact of each of WIC's program components, with particular attention to relevant or key subgroups.

If evaluations prove these to be sound ideas, we realize the result could be a major shift in who gets served and how. But that should not prevent needed reform. Clearly, restraint and sensibility are called for. But that is not the same as inaction.

NOTES

1. The seven categories are (a) infants from birth to 3 months, (b) infants from 4 to 12 months, (c) women and children with special dietary needs, (d) children from 1 to 5, (e) pregnant and

breastfeeding women (basic), (f) postpartum nonbreastfeeding women, and (g) breastfeeding women (enhanced). The amount and type of food within each category may vary, depending on the type of recipient and their nutritional need.

2. This provision was intended to simplify eligibility determination. However, several states have expanded Medicaid eligibility, primarily through waivers and the State Children's Health Insurance Program, beyond 185% of poverty that currently is the maximum income limit specified for WIC. This has effectively increased WIC eligibility in these states as well.

3. The 1997 spending level for each target group was estimated by adding the average monthly administrative cost of $11 to the average monthly cost of food in 1997 for each target group and then multiplying the total by the average monthly number of recipients in each target group. This monthly cost is multiplied by 12 to arrive at an annual cost. Data for these estimates are from the U.S. Department of Agriculture, Food and Nutrition Service (1999, 2).

4. The nutritional-risk estimates are based on health survey data and approximate the percentage of those income-eligible who also have at least one nutritional risk. It is estimated that infants are most likely to have a nutritional risk (95%), and older children are least likely to have one (75%).

5. A number of studies often cited as part of literature documenting WIC's success are not included in this review but are summarized in Anne Gordon and Lyle Nelson's (1995) *Characteristics and Outcomes of WIC Participants and Nonparticipants: Analysis of the 1988 National Maternal and Infant Health Survey.* They suffer from the same methodological problems as the studies reviewed in this article, and because they are older, they are less representative of the current WIC program than the newer studies.

6. The U.S. General Accounting Office "statistically combined" results from 17 studies and concluded that WIC reduced the incidence of low birthweight by 25%. The overall effect size was estimated by weighting the evaluation effect sizes by evaluation sample size, with no apparent consideration that some studies might be more reliable than others.

7. Although this memorandum was not part of the Gordon and Nelson (1995) report, it was prepared for USDA to explain the methodological difficulties of adjusting for selection bias.

8. See Lopez 1999, stating,

> Technically, all three methods first estimate the chances that an individual will participate in WIC based on observed characteristics and a set of variables called instruments. Usually, instruments reflect policy differences, with the main purpose of the instruments to identify the impact of policy variables which affect the chances that an individual pregnant woman participates in WIC. . . . These same instruments, however, should not influence the birth outcomes of an infant, except through the policy effects the mother faces as a result of participating in WIC. For each of these methods to work, the policy instrument must have enough variation in the provision of the policy to identify why pregnant women participate in WIC. If this does not happen, then the estimated effects, addressing selection, may not be reliable, and hence, not useful. (P. 4)

9. This definition of hunger reflects a "social" rather than a "medical" problem. The 1984 President's Task Force on Hunger linked medical definitions of hunger to "a weakened, disordered condition brought about by prolonged lack of food" that was associated with conditions such as serious underweight or stunting of growth in children. In contrast, hunger as a social problem was defined as "a situation in which someone cannot obtain an adequate amount of food, even if the shortage is not prolonged enough to cause health problems" (Hamilton et al. 1997, 3).

REFERENCES

Besharov, D. J., P. Germanis, and P. H. Rossi. 1997. *Evaluating welfare reform: A guide for scholars and practitioners*. College Park: The University of Maryland, School of Public Affairs.

Borrud, L., C. W. Enns, and S. Mickle. 1996. USDA surveys food consumption changes. *Food Review* (September-December)

Bos, J. M., and V. Fellerath. 1997. *Final report on Ohio's Welfare Initiative to Improve School Attendance Among Teenage Parents: Ohio's Learning, Earning, and Parenting Program*. New York: Manpower Demonstration Research Corporation.

Brien, M. J., and C. A. Swann. 1997. Prenatal WIC participation and infant health: Selection and maternal fixed effects. Discussion Paper 295, Thomas Jefferson Center, University of Virginia.

Brien, M. J., and C. A. Swann. 1999. Prenatal WIC participation and infant health: Selection and maternal fixed effects. Unpublished manuscript.

Burkhead, G. S., C. W. LeBaron, P. Parsons, J. C. Grabau, L. Barr-Gale, J. Fuhrman, S. Brooks, J. Rosenthal, S. C. Hadler, and D. L. Morse. 1995. The immunization of children enrolled in the Special Supplemental Food Program for Women, Infants, and Children (WIC): The impact of different strategies. *Journal of the American Medical Association* 274 (4): 312-16.

Caan, B., D. M. Horgen, S. Margen, J. C. King, and N. P. Jewell. 1987. Benefits associated with WIC supplemental feeding during the interpregnancy interval. *American Journal of Clinical Nutrition* 45:29-41.

Contento, I. (senior author), with G. I. Balch, Y. L. Bronner, D. M. Paige, S. M. Gross, L. Bisignani, L. A. Lytle, S. K. Maloney, S. L. White, C. M. Olson, and S. S. Swadener. 1995. Nutrition education for pregnant women and caretakers of infants. *Journal of Nutrition Education* 27 (6): 277-418.

Devaney, B. 1992. *Very low birthweight among medicaid newborns in five states: The effects of prenatal WIC participation*. Alexandria, VA: U.S. Department of Agriculture, Food and Nutrition Service.

Devaney, B. 1997. *Yield of risk for three WIC nutrition risk criteria*. Princeton, NJ: Mathematica Policy Research, Inc.

Devaney, B. 1998. Commentary. In *Feeding the poor: Assessing federal food aid*, edited by P. H. Rossi, 119-23. Washington, DC: American Enterprise Institute.

Devaney, B., L. Bilheimer, and J. Schore. 1990. *The savings in Medicaid costs for newborns and their mothers from prenatal participation in the WIC Program: Volume 1*. Alexandria, VA: U.S. Department of Agriculture, Food and Nutrition Service.

Devaney, B., L. Bilheimer, and J. Schore. 1991. *The savings in Medicaid costs for newborns and their mothers from prenatal participation in the WIC program: Volume 2*. Alexandria, VA: U.S. Department of Agriculture, Food and Nutrition Service.

Devaney, B., L. Bilheimer, and J. Schore. 1992. Medicaid costs and birth outcomes: The effects of prenatal WIC participation and the use of prenatal care. *Journal of Policy Analysis and Management* 11 (4): 573-92.

Devaney, B. L., M. R. Ellwood, and J. M. Love. 1997. Programs that mitigate the effects of poverty on children. *The Future of Children* 7 (2): 88-112.

Devaney, B., and A. Schirm. 1993. *Infant mortality among Medicaid newborns in five states: The effects of prenatal WIC participation*. Alexandria, VA: U.S. Department of Agriculture, Food and Nutrition Service.

Fox, M. K., N. Burstein, J. Golay, and C. Price. 1999. *WIC Nutrition Education Assessment Study: Final report.* Alexandria, VA: U.S. Department of Agriculture, Food and Nutrition Service.

Fraker, T. M., A. R. Gordon, and B. L. Devaney. 1995. The use of selection-bias models in program evaluation. Unpublished manuscript. Washington, DC: Mathematica Policy Research.

Galuska, D. A., M. Serdula, E. Pamuk, P. Z. Siegel, and T. Byers. 1996. Trends in overweight among U.S. adults from 1987 to 1993: A multistate telephone survey. *American Journal of Public Health* 86 (12): 1279-735.

Glickman, D. 1997. Press Conference on Women, Infants, and Children Program. U.S. Department of Agriculture.

Glickman, D. 1998. Remarks at USDA Symposium on Childhood Obesity: Causes and Prevention, Washington, DC.

Goldenberg, R. L., and D. Rouse. 1998. Prevention of premature birth. *The New England Journal of Medicine* 339 (5): 313-20.

Gordon, A. 1993. Effects of prenatal WIC participation on birth outcomes estimated using alternative selection-bias correction methods (Memorandum to Janet Tognetti). Princeton, NJ: Mathematica Policy Research, Inc.

Gordon, A., and L. Nelson. 1995. *Characteristics and outcomes of WIC participants and nonparticipants: Analysis of the 1988 National Maternal and Infant Health Survey.* Alexandria, VA: U.S. Department of Agriculture, Food and Nutrition Service.

Hamilton, W. L., J. T. Cook, W. W. Thompson, L. F. Burton, E. A. Frongillo Jr., C. M. Olson, and C. A. Wehler. 1997. *Household food security in the United States in 1995*. Cambridge, MA: Abt Associates.

Hedberg, V. A., A. C. Bracken, and C. A. Stashwick. 1999. Long-term consequences of adolescent health behaviors: Implications for adolescent health services. *Adolescent Medicine* 10 (1): 137-51.

Huntington, J., and F. A. Connell. 1994. For every dollar spent—The cost-savings argument for prenatal care. *The New England Journal of Medicine* 331 (19): 1303-07.

Institute of Medicine. 1996. *WIC nutrition risk criteria: A scientific assessment*. Washington, DC: National Academy Press.

Karoly, L. K., P. W. Greenwood, S. S. Everingham, J. Hoube, M. C. Kilburn, P. Rydell, M. Sanders, and J. Chisea. 1998. *Investing in our children: What we know and don't know about the costs and benefits of early childhood interventions*. Santa Monica, CA: RAND.

Kerpelman, L. C., D. P. Connell, W. Gunn, R. Schmitz, and J. Wilwerding. 1999. *Preschool immunization project evaluation: Executive summary.* Cambridge, MA: Abt Associates.

Krause, M. 1999. *States have expanded eligibility through Medicaid and the State Children's Health Insurance Program*. Washington, DC: NGA Center for Best Practices, Health Policy Studies Division.

Ku, L. 1999. Debating WIC. *The Public Interest* (135): 108-12.

Ku, L., B. Cohen, and N. Pindus. 1994. *Full funding for WIC: A policy review.* Washington, DC: The Urban Institute.

Kuczmarski, R. J., K. M. Flegal, S. M. Campbell, and C. Johnson. 1994. Increasing prevalence of overweight among U.S. adults: The national health and nutrition examination surveys, 1960 to 1991. *Journal of the American Medical Association* 272 (3): 205-11.

Lewis, K., and M. Ellwood. 1998. *Medicaid policies and eligibility for WIC*. Cambridge, MA: Mathematica Policy Research, Inc.

Lewit, E. M., L. S. Baker, H. Corman, and P. H. Shiono. 1995. The direct cost of low birth weight. *The Future of Children* 5 (1): 35-56.

Long, D. A., and J. M. Bos. 1998. *Learnfare: How to implement a mandatory stay-in-school program for teenage parents on welfare*. New York: Manpower Demonstration Research Corporation.

Lopez, M. H. 1999. Evaluating WIC. Comments on recent research. Paper on file with authors.

Macro International, Inc., and the Urban Institute. 1995. *Final report—WIC dynamics: Volume 1*. Alexandria, VA: U.S. Department of Agriculture, Food and Consumer Service.

Martin, J. A., B. L. Smith, T. J. Mathews, and S. J. Ventura. 1999. Births and deaths: Preliminary data for 1998. *National vital statistics reports*. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control, National Center for Health Statistics.

Mead, L. M. 1997. *The new paternalism: Supervisory approaches to poverty*. Washington, DC: Brookings Institution.

Mei, Z., K. S. Scanlon, L. M. Grummer-Strawn, D. S. Freedman, R. Yip, and F. L. Trowbridge. 1998. Increasing prevalence of overweight among US low-income preschool children: The Centers for Disease Control and Prevention Pediatric Nutrition Surveillance, 1983 to 1995. *Pediatrics* [Online] 101 (1).

Metcoff, J., P. Costiloe, W. M. Crosby, S. Dutta, H. H. Sandstead, D. Milne, C. E. Bodwell, and S. H. Majors. 1985. Effect of food supplementation (WIC) during pregnancy on birth weight. *The American Journal of Clinical Nutrition* 41 (May): 933-47.

NAWD web site. http://www.wicdirectors.org/legagenda8.htm.

Ogden, C. L., R. P. Troiano, R. R. Briefel, R. Kuczmarski, K. M. Flegal, and C. L. Johnson. 1997. Prevalence of overweight among preschool children in the United States, 1971 through 1994. *Pediatrics* [Online] 99 (4).

Ohls, J. 1998. Remarks. In *Workshop on evaluation of food assistance programs in an ERA of welfare reform: Background materials*. Washington, DC: National Research Council.

Olds, D., J. Eckenrode, C. R. Henderson, H. Kitzman, J. Powers, R. Cole, K. Sidora, P. Morris, L. Pettitt, and D. Luckey. 1997. Long-term effects of home visitation on maternal life course and child abuse and neglect: Fifteen-year follow-up of a randomized trial. *Journal of the American Medical Association* 278 (8): 637-43.

Oliveira, V., and C. Gunderson. 2000. WIC and the nutrient intake of children. In *Food Assistance and Nutrition Research Report*. Washington, DC: US Department of Agriculture, Economic Research Service.

Orr, L. 1998. Implementation and data collection. In *Social experimentation: Evaluating public programs with experimental methods*. Washington, DC: U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation.

Paneth, N. S. 1995. The problem of low birthweight. *The Future of Children* 1:19-34.

Puma, M. J., N. R. Burstein, and M. K. Fox. 1992. *Study of the impact of WIC on the growth and development of children: A field test*. Bethesda, MD: Abt Associates.

Puma, M. J., J. DiPietro, J. Rosenthal, D. Connell, D. Judkins, and M. K. Fox. 1991. *Study of the impact of WIC on the growth and development of children. Field test: Feasibility assessment*, Final report: Volume I. Cambridge, MA: Abt Associates.

RAND. 1998. *Preventing very low birthweight births: A bundle of savings*. Santa Monica, CA: Author.

Randall, B., S. Bartlett, and S. Kennedy. 1998. *Study of WIC participant and program characteristics: 1996*. Cambridge, MA: Abt Associates.

Rose, D., J. P. Habicht, and B. Devaney. 1988. Household participation in the food stamp and WIC programs increases the nutrient intakes of preschool children. *Journal of Nutrition* 128 (3): 548-55.

Rossi, P. H. 1988. *Feeding the poor: Assessing federal food aid*. Washington, DC: American Enterprise Institute.

Rush, D., D. G. Horvitz, W. B. Seaver, J. M. Alvir, G. C. Garbowski, J. Leighton, N. L. Sloan, S. S. Johnson, R. A. Kulka, and D. S. Shanklin. 1988. Background and introduction. *The American Journal of Clinical Nutrition* 48 (2): 389-93.

Rush, D., J. Leighton, N. L. Sloan, J. M. Alvir, D. G. Horwitz, W. B. Seaver, G. C. Garbowski, S. S. Johnson, R. A. Kulka, J. W. Devore, M. Holt, J. T. Lynch, T. G. Virag, M. B. Woodside, and D. S. Shanklin. 1988. Study of infants and children. *American Journal of Clinical Nutrition* 48 (2): 484-511.

Rush, D., N. L. Sloan, J. Leighton, J. M. Alvir, D. G. Horvitz, W. B. Seaver, G. C. Garbowski, S. S. Johnson, R. A. Kulka, M. Holt, J. W. Devore, J. T. Lynch, M. B. Woodside, and D. S. Shanklin. 1988. Longitudinal study of pregnant women. *American Journal of Clinical Nutrition* 48 (2): 439-83.

Schiller, J. T., and M. K. Fox. 1999. WIC nutrition education: Current issues and opportunities. A synthesis of recent research. Presented at the National Association of WIC Directors annual meeting, New Orleans, LA.

Schramm, W. F. 1985. Participation and its relationship to newborn Medicaid costs in Missouri: A cost/benefit analysis. *American Journal of Public Health* 75 (8): 851-57.

Schramm, W. F. 1986. Prenatal participation in WIC related to medicaid costs for Missouri newborns: 1982 update. *Public Health Reports* 101 (6): 607-15.

Schwartz, J. B., D. K. Guilkey, J. S. Akin, and B. M. Popkin. 1992. *The WIC breastfeeding report: The relationship of WIC program participation to the initiation and duration of breastfeeding.* Alexandria, VA: U.S. Department of Agriculture, Food and Nutrition Service.

Sherry, B., D. Bister, and R. Yip. 1997. Continuation of decline in prevalence of anemia in low-income children. *Archives of Pediatric Adolescent Medicine* 151 (September): 928-30.

Sullivan, L. W. 1995. One for our children. *The Washington Post*, 28 February 1995.

Troiano, R. P., K. M. Flegal, R. J. Kuczmarski, S. M. Campbell, and C. L. Johnson. 1995. Overweight prevalence and trends for children and adolescents: The National Health and Nutrition Examination Surveys, 1963 to 1991. *Archives of Pediatric Adolescent Medicine* 149:1085-91.

U.S. Center for National Health Statistics. 1990. Health Promotion and Disease Prevention: United States. *Vital Health Statistics* 10 (185).

U.S. Code of Federal Regulations, 7 CFR 246.

U.S. Department of Agriculture. 1998. *WIC Policy Memorandum 98-9 Nutrition Risk Criteria.* Washington, DC: Author.

U.S. Department of Agriculture. 1999. Profile of overweight children. *Nutrition Insights*, Insight 13.

U.S. Department of Agriculture, Food and Nutrition Service, Office of Analysis and Evaluation. 1997. *Special Supplemental Nutrition Program for Women, Infants, and Children (WIC): Eligibility and coverage estimates 1997 update.* Washington, DC: Author.

U.S. Department of Agriculture, Food and Nutrition Service, Office of Analysis and Evaluation. 1999. *Fiscal Year 1997 WIC Food Package Costs.* Washington, DC: Author.

U.S. General Accounting Office. 1992. *Early intervention: Federal investments like WIC can produce savings*, GAO/HRD-92-18. Washington, DC: Author.

U.S. General Accounting Office. 1997. *Head start: Research provides little information on impact of current program*, GAO/HEHS-97-59. Washington, DC: Author.

U.S. General Accounting Office. 1998. *Information on WIC sole-source rebates and infant formula prices*, GAO/RCED-98-146. Washington, DC: Author.

U.S. General Accounting Office. 1999. *Efforts to control fraud and abuse in the WIC program can be strengthened*, GAO/RCED-99-224. Washington, DC: Author.

Watkins, S. R. 1998. Remarks delivered at Nutrition Education and Breastfeeding Promotion Conference, Atlanta, GA.

Yip, R. 1989. The changing characteristics of childhood iron nutritional status in the United States. In *Dietary iron: Birth to two years*, edited by L. J. Filer Jr. New York: Raven.

Yip, R., N. J. Binkin, L. Fleshood, and F. L. Trowbridge. 1987. Declining prevalence of anemia among low-income children in the United States. *Journal of the American Medical Association* 258 (12): 1610-23.

Douglas J. Besharov is a resident scholar at the American Enterprise Institute and a professor at the University of Maryland School of Public Affairs.

Peter Germanis is assistant director of the University of Maryland's Welfare Reform Academy.