

Nonimpact Research for Assessing Ongoing Programs

Although the central focus of this report is on estimating program impacts, it is important to recognize that other aspects of program assessment can also be valuable in evaluating ongoing programs. Research that is not specifically focused on program impact typically employs different research methods than those described previously. Especially important are monitoring studies, participation studies, and program integrity studies, described briefly below.

Monitoring Studies

Monitoring studies (also called surveillance studies) periodically measure outcomes for the participant or target population without specifically attempting to attribute the outcomes to the program. For example, there is considerable interest in altering the nutrient content of school meals to promote compliance with the Dietary Guidelines for Americans. Sampling school meals on a periodic basis can establish whether the trends in meal preparation are in conformity with expectations. Finding that trends are not improving might indicate the need for more programmatic activity, such as an expanded effort to train school food service staff. Alternatively, positive trends might indicate that no program changes are needed. Of course, neither finding substantiates any program effect. An improving situation may mean that the policymaker has less to worry about, which is important information, but additional information is needed before the program can be credited with the improvement.

Participation Studies

Participation studies are typically concerned with issues of targeting and target population penetration. Targeting studies measure the extent to which the actual clients of a program are limited to those intended to be served by the program. Penetration studies address the issue of whether the program reaches a sufficient percentage of its intended target population. Programs that are serving unintended clients are wasting resources, relative to the program objective. And programs that do not have high enough participation levels may not be able to achieve the effects they seek.

Targeting Studies

Targeting studies typically involve surveys of program participants to determine whether their characteristics are as intended. For example, Glantz and his colleagues (1997) surveyed the families of children served in the CACFP. A major finding from this study was that the families of children in the family child care portion of the program had average incomes far above the poverty level. This finding indicated that this portion of the program was weakly targeted. (The Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (PRWORA) contained provisions designed to strengthen the targeting.)

Penetration Studies

Penetration studies have to be designed to represent the entire potential target population in order to ascertain what proportion and types of units are participating in the program. Accordingly, penetration studies tend to be more extensive and expensive than targeting studies. An exception to this rule occurs for programs that are included in ongoing periodic national surveys, such as the Current Population Survey (CPS) or the Survey of Income and Program Participation (SIPP), both run by the Bureau of the Census. Trippe (1995) studied trends in participation in the FSP as shown in the SIPP,¹² and showed that participation rates as well as the overall number of eligible households increased during the period 1988-92.

Other examples of penetration studies reveal serious program flaws. A 1987 provision in FSP legislation extended the use of food stamps to pay for meals in soup kitchens serving the homeless. Surveying a national sample of soup kitchen operators, Burt and Cohen (1988) found that few kitchens had applied to be authorized to accept food stamps, and individuals who used soup kitchens were largely unwilling to pay for food, which was otherwise free, by using food stamps. Because soup kitchens relied heavily on surplus food given without charge to kitchens that did not charge their customers, these organizations could have lost an important subsidy if they had accepted food stamps.

¹²Both the CPS and SIPP have consistently underestimated participation in the Food Stamp Program and other assistance programs.

Program Integrity Studies

A program's potential for impact can be affected by any diversion of program funds from their intended purpose, and program integrity studies focus on the question of whether such diversions exist. Investigative or enforcement arms of the government often pursue program integrity through detailed investigation of a small number of suspect situations, as in a recent examination of the CACFP carried out by the USDA Office of the Inspector General (USDA, 1999). Such investigations can lead to prosecutions and remedial actions, but they usually do not produce general estimates of the prevalence of problems or the overall percentage of funds diverted. Complementary research using surveys or administrative data from representative samples is therefore often needed.

Impact Evaluation of Demonstrations

The preceding sections considered potential strategies for evaluating ongoing food assistance and nutrition programs, with an emphasis on entitlement or saturation programs that have been operating at substantial volume for over two decades.

We turn now to evaluations of “demonstration” or “pilot” programs. These demonstrations typically represent policy initiatives that are to be tested and examined on a limited scale before full-scale implementation. The intervention may be an entirely new program, but it is more commonly a significant modification to an existing program. Past examples include demonstrations of cashing out food stamps, requiring education and training for food stamp recipients, and delivering food stamp or WIC benefits through electronic benefit transfer.

Perhaps the largest set of examples of demonstration impact evaluations consists of the waivers obtained by more than 40 States from the U.S. Department of Health and Human Services to demonstrate the feasibility and effectiveness of State-proposed changes in rules for the Aid for Families with Dependent Children (AFDC) program. The changes ranged from imposing time limits on AFDC benefits to capping benefits upon the birth of additional children. Many of the waiver changes involved requiring preparation for employment and mandating job searches. Most of the waivers were evaluated for impact by randomized experiments

in which the experimental groups proceeded under the changed rules and the controls continued under existing AFDC regulations.

Three distinguishing features of a demonstration lead to evaluation strategies that differ from those for ongoing programs:

- The intervention is new. In principle, evaluation activities can begin at the same time as implementation of the demonstration, or even before.
- The intervention has not been mandated by law for the entire program or service population.
- The intervention is applied to a restricted number of participants. During the relevant periods, some potential targets will be subject to the intervention and some will not.

These features generally make it much easier to identify a Counterfactual in a demonstration than in ongoing programs. In particular, the absence of a legal entitlement and saturation volume remove the main obstacles to randomized experimentation, which make this the preferred impact evaluation design. Nevertheless, some circumstances require quasi-experiments, as discussed below.

Randomized Experiments

In evaluating a demonstration intervention that modifies an existing program, the intervention's impact is normally defined as the difference between outcomes with the new intervention and outcomes with the pre-existing version of the program. The Counterfactual is the status quo; the control subjects experience the usual program services but are not offered the new services incorporated in the intervention. For example, the several demonstrations of cashing out food stamps estimated the effects on food purchases of receiving benefits in the form of checks rather than in the form of food stamps. They did not estimate the overall impact of subsidizing food purchases.

Strengths and Limitations of Randomized Experimentation in a Demonstration

The randomized experiment is the strongest design available for evaluating demonstration interventions. The findings of such an evaluation are considered substantially more reliable than findings from even the strongest of the quasi-experiments. If a randomized