SAMPLING FRAME FOR THE ELIGIBLE NON-PARTICIPANT SAMPLE
IN THE NATIONAL JTPA STUDY

Jeffrey Smith
Department of Economics
University of Maryland,
NBER and IZA
smith@econ.umd.edu

First version, February 1992
Revised, March 1994
Revised, March 10, 2003

1. Introduction

As part of the National Job Training Partnership Act (JTPA) Study, data were collected on a sample of persons eligible for, but not participating in, the JTPA program at four of the sixteen sites taking part in the study. These eligible non-participants, or ENPs, were intended to provide information about the characteristics of the population eligible for the JTPA program and to serve as the basis for comparison groups in non-experimental evaluations of the JTPA program.

Although the construction of the ENP sample is described in Bloom et al. (1990), the presence of some omissions in that account and the availability of more recent information on hit and response rates suggest the value of the present expanded and corrected discussion.

The collection of the ENP sample had two goals. The first was to gain information on the characteristics of the JTPA eligible population. This information is of use in understanding the selection process into JTPA, the effects of the JTPA performance standards system, and in formulating marketing and outreach efforts for the JTPA program. The second goal was to provide a comparison group for use in constructing non-experimental estimates of the impact of the JTPA program.

As noted in Bloom et al. (1990) these two goals imply somewhat different emphases for the sample design. The comparison group function is served when those components of the eligible population most likely to apply to JTPA are over-sampled; this is one of the justifications for excluding individuals aged over 54 years from the ENP sampling frame. On the other hand, the goal of providing a clear picture of the entire

eligible population is best served by taking a random sample, which was in fact largely done within the sub-population of 16 to 54 year-olds.

In its final form, the ENP sample comprises 3,004 individuals at four of the sixteen JTPA Service Delivery Areas (SDAs) involved in the National JTPA Study (NJS). The sample was originally intended to include ten sites, but was reduced to five when the costs of the listing required for the eligibility screener turned out higher than expected. Five sites then became four when the Fresno, California SDA dropped out of the NJS after survey operations had already begun. The four sites finally included in the sample are the Corpus Christi-Nueces County SDA, the Jersey City SDA, the Northeast Indiana (Fort Wayne) SDA and the Providence-Cranston SDA. A follow-up interview was scheduled for each member of the ENP sample; these interviews were spread over the time frame covered by the first and second follow-up interviews for the experimental samples at each site.[1]

The remainder of this note proceeds more or less chronologically through the design and fielding of the full baseline survey for the ENP sample.

2. Population Definition

The eligibility criteria for the JTPA program fall into two broad categories: economic disadvantage, as determined by participation in the AFDC, Food Stamp or General Assistance programs or by having a family income in the last six months below certain levels, and barriers to employment, which include such factors as language difficulties

---

[1] Note that this differs from the claim on page 91 of Bloom et al. (1990) that the ENP follow-up interviews would all take place during the time span of the first follow-up interview for the experimental samples at each site.

and alcoholism.[2] Service delivery areas in JTPA were required to enroll at least ninety percent of their trainees under the economic disadvantage criteria; in practice, as shown in U.S. Department of Labor (1993), the actual fraction exceeded even this high level. This fact, combined with the qualitative and difficult-to-measure character of many of the barriers to employment, led to a decision to include in the ENP sample only those persons meeting the economic disadvantage eligibility criteria.

In addition to limiting the population of interest to the economically disadvantaged, several other subsets of the overall JTPA eligible population were also excluded. First of all, individuals permanently out of the labor force due to a total disability were excluded, as they would not ever demand JTPA services. Second, in-school youth were excluded, primarily to maintain comparability with the experiment, from which they were also excluded. Finally, individuals 55 years of age or older were excluded from the population of interest on the grounds that previous studies such as Sandell and Rupp (1988) revealed an extremely low participation rate in JTPA among members of this age group. By not interviewing these older individuals, the available resources could be focused on subsets of the eligible population more likely to apply to JTPA.

The text in Bloom et al. (1990) further indicates that individuals presently participating in JTPA will be excluded from the ENP sample. It is unclear how this

---

[2] A person is economically disadvantaged if he or she receives, or is a member of a family that receives, cash welfare payments under a Federal, State or local program; receives, or is a member of a family that receives, a total family income for the six-month period prior to application which does not exceed the poverty level established by the Office of Management and Budget or 70 percent of the Bureau of Labor Statistics lower living standard income level, whichever is greater; is a foster child whose foster family is receiving State or local government payments; or is an adult handicapped individual whose own income meets the eligibility criteria, but whose family income may not. See NCEP (1987) or the detailed study of the eligibility rules and their consequences in Devine and Heckman (1994).

restriction was implemented in practice, as the screener instrument does not contain information on JTPA participation. What is known is that 104 of the 3004 ENP sample members were at some point randomly assigned as part of the experimental evaluation.

3. Sample Selection

After defining the population of interest, the next step in preparing the ENP sample involved designing and implementing a sample selection mechanism for choosing the eligible persons to be administered the full baseline survey. As part of its contractual agreement with Abt Associates, National Opinion Research Center (NORC) survey research staff designed the sampling plan for the ENP sample. Their design included the following steps:

1. Using Census Summary Tape File 3 (STF3) records with Types 15 and 16, a record was generated for each Enumeration District (ED) or Block Group (BG) at the four full baseline sites containing the number of people with household incomes at or below 124 percent of the poverty level (referred to hereafter as "in poverty") in the ED or BG and the total number of people in the ED or BG. Note that Type 15 refers to BGs and Type 16 refers to EDs. Table 95 in the STF3 data was used to determine these numbers. This Table indicates the number of people in the given geographic division with incomes at various points in relation to the appropriate poverty level. Note that the text of Bloom et al. (1990) is incorrect in stating that the cutoff was at the poverty level, rather than at 125 percent of the poverty level.

2. The total number of persons in poverty in each SDA was determined by aggregating over all of the EDs and BGs included in the SDA.

3. For each ED and BG, the fraction of individuals in the ED or BG in poverty was calculated. The EDs and BGs were then sorted by the fraction of individuals in poverty within them.

4. Beginning with the EDs and BGs with the lowest fraction in poverty, EDs and BGs were eliminated until the cumulative fraction of people in poverty in the SDA that had been eliminated was as close to five percent from below as possible. Note that this elimination of EDs and BGs with low fractions of their population in poverty is not mentioned in Bloom et al. (1990).

5. A file of records containing all the retained census blocks (not BGs!) and EDs was then produced. Each record contained the total number of Housing Units, the number of Occupied Housing Units, the total population and the total population in poverty for the given block or ED. Individual blocks were assigned a value for the fraction in poverty equal to that of the block group to which they belonged.

6. For each block or ED, the ratio of the number of individuals in the block or ED in poverty to the mean number of individuals in poverty in all of the blocks and EDs in the SDA was calculated. Call this number the poverty ratio.

7. A random sample of blocks was drawn from the set of retained blocks and EDs for each SDA. The probability that a given block or ED was chosen was proportional to the poverty ratio in that block or ED.

8.   Field staff enumerated all of the dwelling units in each of the selected blocks.

9.   A random sample of dwelling units was drawn from the list of dwelling units for each of the selected blocks. The probability of a given dwelling unit being drawn was inversely proportional to the poverty ratio in the block containing the dwelling unit. Note that the combination of the probabilities in this step and the step above means that each dwelling unit in the set of retained blocks and EDs had an equal probability of being drawn.

The total number of dwellings randomly chosen was selected to yield a sample of completed ENP baseline surveys equal in size to the expected number of controls in each SDA, based on assumptions regarding the overall incidence of JTPA eligibility at each site and the completion rates for the screener and full baseline surveys.[3]

Several features of this sampling scheme deserve note. First, this is a random sample of dwelling units, not of individuals. Second, this is a standard area probability sample. Other sampling frames, such as random digit dialing or sampling off of lists of program participants from Food Stamps or AFDC, could have been used instead. These alternative methods would have cost less, but would have had the disadvantage of excluding interesting subsets of the eligible population, such as those with insufficient resources to maintain a telephone, from the sampling frame. Third, no stratification is attempted in the ENP sample, though plans originally called for stratification both on demographic characteristics and on labor force status.

---

[3] In practice, the numbers do not match up very well. At the Corpus Christi SDA, there are 1060 ENPs and 442 controls, at Northeast Indiana (Fort Wayne) there are 892 ENPs and 1041 controls, at Jersey City there are 529 ENPs and 470 controls, and at Providence there are 523 ENPs and 423 controls.

Finally, the most interesting feature of this sampling scheme is the use of selection proportional to size sampling techniques. This last term is the formal name for the procedure by which blocks are first selected in proportion to their poverty ratio, and then, within the blocks thus selected, dwelling units are selected with a probability inversely proportional to the poverty ratio for their block. The desire for clustered sampling, of which this scheme is an instance, results from the fact that it reduces survey costs by reducing the amount of interviewer travel necessary for a given sample size. As Kish (1965) notes, the primary reason to prefer the selection proportional to size method to a regime with fixed and equal probabilities of selection at each stage is that the former reduces the amount of variation in the expected sample size, since it removes the possibility that the selection of a few large or small blocks at the initial step could drastically alter the result of the sampling.

4. Screener Interviews

The household screener survey used to locate persons eligible for JTPA is a simple, two-page instrument. On the first page, the name and relationship to the respondent of each household member is collected, along with their age, sex, labor force status, disability status and, for those aged 16-21, their school enrollment status. This information allows implementation of the exclusions of those outside the 16-54 age group and also enables the planned stratification by demographic characteristics and labor force status.

The second page collects information on program participation and income in the last six months for use in determining eligibility based on the economic disadvantage

criteria. The income categories were created separately for each site so that the categorical boundaries matched the income eligibility cutoffs for families of different sizes. These cutoffs also vary over time within sites in response to annual adjustments in the poverty line and lower living standard income level. This page also collected locating information.

Abt survey staff administered the screener at the random samples of dwelling units generated by the NORC survey staff as described above. The response rate was 94.0 percent in Corpus Christi, 91.6 percent in Providence, 80.2 percent in Jersey City and 90.0 percent in Northeast Indiana (Fort Wayne). For the four sites as a whole, the response rate is 89.4 percent, but there is surprising variation across sites, with the Jersey City site having a noticeably lower response rate than the other three sites.

Also of interest are the "hit rates" at each site. The hit rate is the number of eligible non-participants found per completed screener. The hit rates range from 40.0 percent at Corpus Christi and 40.2 percent at Jersey City, both relatively poor sites, down to 17.5 percent in Providence and only 6.1 percent in the relatively well-off Northeast Indiana (Fort Wayne) site.

5. Full Baseline Interviews

Abt attempted to administer the full baseline survey to all of the eligible individuals identified by the screener interview. These interviews took place over the period June 1988 to November 1989. At roughly the same time, from December 1987 to November 1989, NORC survey staff attempted to administer the full baseline survey to those in the

control group at the four sites. The later start for the ENP full baseline interviews results

from the time spent listing the sites and administering the screener.

Individuals in the ENP sample interviewed by Abt had the option of completing

the survey in English, Spanish or, at the Providence site, in Laotian or Cambodian, while

those in the control group interviewed by NORC all completed the survey in English.

The response rate for the baseline survey among the ENPs was 79.1 percent at Corpus

Christi, 74.5 percent at Providence, 77.7 percent at Jersey City and 78.8 percent at Fort

Wayne (Northwest Indiana), for an overall ENP response rate of 77.8 percent. In

contrast, the response rates obtained for the control group average a bit below 90 percent.

The reason for the difference is unclear, though perhaps the control group members felt a

stronger obligation to complete the survey given their participation in the experiment.

6. Follow-up Interviews

NORC survey staff attempted to complete a single follow-up interview with each

member of the ENP sample who completed the full baseline survey. These interviews

were scheduled from 12 to 36 months after the baseline interview. This period

corresponds to the union of the first and second follow-up interview periods for the

experimental treatment and control group members. Only a single follow-up interview

was scheduled for the ENPs in order to reduce survey costs. Response rates for the single

ENP follow-up survey were 80.81 percent for adult males (age 22 and older), 80.32

percent for adult females (age 22 and older), 72.02 percent for male out-of-school youth

(ages 16 to 21) and 73.67 percent for female out-of-school youth (ages 16 to 21). The

overall response rate was around 79.34 percent. Analysis of the earnings data collected

in the baseline survey reveals that non-respondents to the follow-up survey have lower

earnings during the baseline period, indicating that non-response is not independent of

important variables of interest.

References

Bloom, Howard, Larry Orr, Fred Doolitte, V. Joseph Hotz and Burt Barnow. 1990. *Design of the National JTPA Study*. New York, NY and Bethesda, MD: Manpower Demonstration Research Corporation and Abt Associates, Inc.

Devine, Theresa and James Heckman. 1994. "The Structure and Consequences of Eligibility Rules for a Social Program." In Solomon Polachek, ed., *Research in Labor Economics, Volume 15*. Greenwich, CT: JAI Press. 111-170.

Kish, Leslie. 1965. "Selection with Probabilities Proportional to Size Measures." Chapter 7 in *Survey Sampling*. New York, NY: Wiley.

National Commission for Employment Policy. 1987. *The Job Training Partnership Act.* Report by the National Commission for Employment Policy. Washington DC: NCEP.

Sandell, Steven and Kalman Rupp. 1988. "Who is Served in JTPA Programs: Patterns of Participation and Intergroup Equity." National Commission for Employment Policy Research Report RR-88-03.

U.S. Department of Labor. 1993. *Job Training Quarterly Survey: JTPA Title IIA and III Enrollments and Terminations During Program-Year 1991 (July 1991-June 1992).* Employment and Training Administration, Office of Strategic Planning and Policy Development, Division of Performance Management and Evaluation.