Training for the Urban Unemployed: A Reevaluation of Mexico's Training Program, Probecat

Wodon, Quentin and Minowa, Mari

World Bank

July 2001

Training for the Urban Unemployed: A Reevaluation of Mexico's Training Program, Probecat

Quentin Wodon and Mari Minowa

Abstract

For many years the government of Mexico has implemented a large training program for the unemployed. The program has been evaluated twice before with similar methodologies. These two evaluations yielded encouraging results in that the program apparently reduces unemployment and increases earnings. This article suggests that both evaluations may suffer from inappropriate controls for the endogeneity of program participation. Using the availability of the program at the state level as a determinant of individual participation, the article uses the data of the second evaluation to indicate that Probecat does not decrease unemployment, nor does it increase wages.

In 1984, the government of Mexico implemented a training program, Probecat (*Programa de Becas de Capacitación para Desempleados*), for the unemployed who live mostly in urban areas. The program has been evaluated twice, first by the World Bank using data from 1992 (Revenga, Riboud, and Tan 1994), and next by the Mexican Ministry of Labor using data from 1994 (STPS 1995).[1] Both evaluations used longitudinal data, comparing a sample of Probecat participants (the treatment group)

1. An earlier evaluation was attempted by Carlson (1991), but this evaluation did not consider the problem of endogeneity of program placement or selection bias, which is crucial for good evaluations.

with a sample of unemployed individuals from Mexico's urban employment survey (the control group). The second evaluation closely followed the method used in the first evaluation. The two evaluations yielded positive results. Controlling for other characteristics, such as education, family situation, and professional experience, program participants find employment faster than nonparticipants and they earn higher wages. These encouraging results have been used to support the extension of the program, which now serves about 550,000 beneficiaries per year (as of 1996–98), as compared with only 50,000 in the first decade of the program (1984–93).

This article reassesses the results of past evaluations of Probecat. We suggest that past evaluations suffered from inappropriate controls for the endogeneity of program participation. For example, past evaluations concluded from Cox regressions that the time necessary to find employment after the training was reduced for program participants, but this might have been due to sample selection because program participants may also be those most eager to work. We argue that the matching techniques used in past evaluations of Probecat were not sufficient to handle this type of problem. Following Ravallion and Wodon (1998), we use an alternative econometric method for evaluating the impact of Probecat. Specifically, the availability of Probecat at the state level is used as an instrumental variable to control for the endogeneity of program placement. We then find that Probecat has no impact on unemployment and wages.

The disappointing results of Probecat in terms of raising wages and employment should not be surprising. Most retraining programs in Organisation for Economic Co-operation and Development (OECD) countries have been found to have limited impacts, and when programs have been found to have some impact, this impact tends to vanish after a few years (Dar and Gill 1998). The fact that Probecat may not be beneficial in the medium to long run for participants does not mean that it should be suppressed. The program could be considered a temporary safety net (through the minimum wage stipend) rather than a training program. Alternatively, it could be improved to provide training with longer-lasting effects. To motivate an inquiry into how to improve the program, however, it must first be recognized that contrary to the results of earlier evaluations, the program may not be satisfactory.

The structure of the article is as follows. After a description of the program and the changes made over the years, we review the results of past evaluations. We then present our own new results. A conclusion with policy implications follows.

Program Description

Probecat, the Mexican Job Training Program for Unemployed Workers, was established in 1984 in response to rising unemployment and deteriorating living standards in the aftermath of the 1982 economic crisis. Despite a comparatively low open unemployment rate (6 percent in four major cities in 1984), Mexico suffered (then as now) from chronic underemployment. (For a discussion of unemployment and underemployment data in Mexico at that time, see, for example, Fleck and Sorrentino 1994.) Moreover, there was a recognition of the shortages of adequately trained labor in selected growing sectors of the economy. Thus, the stated objectives of the program was to improve the productivity of unemployed workers to help them find employment. The program has become massive in recent years. In the first decade of its existence (1984–93), Probecat provided training to roughly 50,000 people per year, but this rose to 199,000 people in 1994, 412,000 people in 1995, and about 550,000 people per year between 1996 and 1998.

Probecat is administered by the Mexican Ministry of Labor. The program provides publicly funded job training and a subsistence allowance during the training period to participating unemployed workers. Initially in 1986 Probecat provided training in high schools and other training centers. This first module of the program is referred to as school-based training (*Cursos Escolarizados*). Later, to strengthen the link between the training provided under the program and the actual needs of the productive sector in the economy, a new module of so-called in-service training was added (*Cursos Mixtos*). Under this module, local employers provide training, whereas the workers' stipend is paid by the government. Upon completion of the training, the employers are required to hire at least 70 percent of the trainees. A third module of the program consisting of training for the self-employed, PILEOT (*Programa de Iniciativas Locales de Empleo y Ocupación Temporal*), was created in 1995 in response to rapidly increasing unemployment after the 1994 financial crisis.

For the school-based and in-service modules, program beneficiaries are selected from the unemployed workers who register in the State Employment Service offices. The applicants' job skills and interests are evaluated against the needs of the local market. Apart from basic requirements for all, the selection procedure gives variable weights to different criteria. Only individuals with a total composite score exceeding a threshold level are eligible to participate. Participants can obtain training only once. The training lasts for two to three months on aver-

age, and participants receive stipends equivalent to the value of the minimum wage, plus transportation costs to the training site and basic health insurance coverage while on training. More details on the eligibility and features of each of the three modules are provided in table 1. It is worth mentioning that the share of in-service training within the Probecat beneficiaries increased from about 5 percent in 1987–92 to 20 percent in 1993, but dropped to 13 percent in 1996 as a result of the rapid expansion of the new PILEOT module created in 1995.

Past Evaluations

As mentioned in the introduction, two primary evaluations of Probecat have been conducted so far. The first evaluation was prepared by the World Bank (Revenga, Riboud, and Tan 1994), and used data from a survey administered in 1992 to Probecat trainees from the 1990 cohort. Data were gathered on unemployed individuals in the 1990–91 National Urban Employment Survey (*Encuesta Nacional de Empleo Urbano, or* ENEU) to construct the control group. The second evaluation was conducted in 1995 by the Ministry of Labor with a similar methodology (STPS 1995) and data for employment in 1993–94. Both studies sought to evaluate the impact of Probecat according to two main indicators: (a) the time to find a first job after the training and (b) monthly earnings. The studies also contain information on hours worked and hourly wages, as well as cost-benefit simulations, which will not be discussed here. Whereas Probecat participants were drawn only from the school-based module for the 1992 study, the 1995 study evaluated the impact of both the school-based and in-service modules.

Both studies used the posttraining labor market experiences of randomly selected program participants who graduated in the previous year, and they compared these experiences to those of a control group. For this control group, the studies used panel data on unemployed individuals drawn from the quarterly ENEU covering the main urban areas of Mexico. (Probecat is now available in rural areas, especially through the PILEOT module, but this was not the case before.) The ENEU used a quarterly rotation system so that each rotation group of individuals remained in the survey for five consecutive quarters. The choice of the quarterly data in the ENEU matched that of the survey of the Probecat participants. The ENEU included detailed information on employment status, monthly earnings, and hours worked per week. The Probecat survey administered to program participants was designed to match the questions in the ENEU, so that the information for the two

TABLE 1. MODULES OF PROBECAT'S MAIN TRAINING MODULES

	School-based training	*In-service training*	*PILEOT*
Eligibility rules	Unemployed, registered with SES, aged 16 to 55, having completed primary school and having at least 3 months of experience.	Unemployed, registered with SES, aged 18 to 55, having completed primary school (this can be waived by firm). No prior experience required.	Unemployed aiming at self-employment, aged 16 to 55, literate, no upper secondary schooling. Special module for community activities.
Training provider	Training schools or centers.	Participating firms.	Training centers or instructors.
Training duration	1 to 6 months.	1 to 3 months.	1 to 3 months.
Benefits received	Training, minimum wage, transportation costs, and health insurance.	Training, minimum wage, transportation costs, and health insurance.	Same, plus a set of tools for self employment module.
Training costs	Probecat program.	Firm (cost of instructors, equipment, and materials).	Probecat program.
Training content	Set by training provider with little customization.	Set by firm. Mostly hands-on training.	Set by the training provider with customization.
Placement	Required to register with SES. No placement.	Firms required to employ 70 percent of the trainees.	No particular follow-up.

Source: Created by the authors from information provided by the Ministry of Labor (STPS).

groups would, to a large extent, be comparable. That is, in addition to standard sociodemographic characteristics, the Probecat survey elicited information on the current or last job since completing Probecat, the first job after the training, the second job after the training, and the last job prior to participating in the program.

The two studies acknowledged that there may be unobservable characteristics influencing program participation and outcomes, and that these characteristics may be different between the treatment group (Probecat participants) and the control group (ENEU unemployed workers). To control for selection bias, the studies used two different approaches. The first approach was used for analyzing the length of the employment search, and the second for monthly earnings.

For the analysis of the length of the employment search, the studies relied on probit regressions for the probability of participating in Probecat to construct the control group. In the 1992 study, the individuals below a certain probability (0.6) of participating were eliminated from the sample. It is less clear what exactly was done in the 1995 evaluation, but the principle was the same. Next, using the (slightly) reduced sample of pooled treatment and control group members, Cox proportional hazards models were estimated to assess the impact of training on the time necessary to find employment. In 1995, for example, the regressors included dummy variables for the individual's age, the level of education, marital status, household structure, work experience, and characteristics of employment before being unemployed. The dependent variable was the number of months needed for finding employment, counted from the first month when the individual entered the control group (ENEU) or completed the training for participants.

For the analysis of earnings, both studies used Heckman's (1979) sample selection model. In the 1992 study, the first equation referred to the logarithm of monthly earnings, and the second equation to the probability of participating in the Probecat training. In the 1995 study, while the participation equation was the same, the first equation measured the difference in the logarithms of earnings before and after program participation with the logarithm set to zero in case of unemployment (neither of the studies actually provided the participation equations).

Overall, the results obtained in the two studies were encouraging for the program. In 1992, Probecat was found to reduce the length of unemployment for men, but not for women. In the 1995 study, both the school-based and in-service modules reduced the time needed to find employment for both men and women. As for monthly earnings, the 1992 study found positive impacts for men, but again not for women.

The 1995 study found positive impacts for men and negative impacts for women. Both studies also conducted cost-benefit analyses. The program costs were calculated using administrative data from the state and national program offices. Direct costs included the remuneration for instructors, the costs of training materials and facilities, and administrative costs for program operation. The training stipends paid to the program participants during the training were not included in the costs because they were considered safety net transfers. The treatment of indirect opportunity costs for participating in the program differed, but overall the two studies concluded that the program was cost effective and performing relatively well.

Although both studies were carefully implemented, several critiques can be made about the methodology used. Manski (1996) mentioned a few. First, in using the unemployed individuals in the ENEU to form the control group, it is assumed that none of the ENEU individuals have benefited from the program. This is not the case because every individual in the ENEU has some probability of having participated in Probecat. Fortunately, given that the program was small until 1993, only a very small minority of the individuals in the control group are likely to have participated in the program (this would not be true for future evaluations). Next, Manski pointed out that the combination of two random samples (Probecat trainees and ENEU unemployed individuals) was not actually a random sample, so that in the absence of the standard properties for the residuals, the results of regressions might not yield consistent parameter estimates, especially because the models used were sensitive to the assumption of bivariate normality (Goldberger 1983). In the absence of better data, not much could be done about this using standard techniques. Third, Manski argued that no theoretical proof exists that matching methods do indeed provide for a solution to the sample selection problem, although this could be debated. Finally, Manski noted that in the estimation of earnings, while participation in Probecat was controlled for, the sample selection bias resulting from the decision to work (which originally motivated the Heckman model) was not accounted for, which was recognized by the authors of the 1992 study.

In our own evaluation of Probecat to be presented in the next section, we do not solve all the above problems, but we try to solve some. Consider first the analysis of the duration of unemployment. The method used to control for endogeneity in past evaluations was rough because the matching was imprecise. Typically, when using matching procedures, one matches every participant with one, or perhaps a few, nonpartici-

pants by minimizing the distance between the probability of participation of each participant and that of his match among the nonparticipants (Rosenbaum and Rubin 1983, 1985). Here the procedure had to be reversed because there were more participants than nonparticipants in the pooled sample, but the logic was the same: every nonparticipant should be matched on an individual basis with one, or very few, participants. This is apparently not what was done in past evaluations of Probecat. The matching was apparently not done on an individual basis. Instead, individuals with low probabilities of participating were excluded from the pooled sample. There is in fact evidence in both the 1992 and 1995 studies that there remains a problem of sample selection after the frequency matching because the coefficients of the inverse Mill's ratios in the Heckman regressions used for earnings tend to be statistically significant. This problem is not recognized in presenting the Cox regressions.

Consider next the earnings regressions. As noted by Manski, although the Heckman procedure allows in principle to control for the endogeneity of program placement, there is no control for the endogeneity of the decision to work. The procedure proposed in the next section provides a way to control for both types of endogeneity in the estimation.

Alternative Evaluation

This section provides our alternative evaluation results. After discussing the model for participation in Probecat, we analyze the impact of the program on the length of the employment search and on the earnings.

Participation Equations

To evaluate the impact of Probecat, we used the same data as that used in the 1995 study, but with an alternative methodology. Following Ravallion and Wodon (1998, 2000), we used the program availability in a geographic area as a determinant of program participation at the individual level, assuming that program availability does not influence outcomes (unemployment duration and earnings) that are conditional on individual participation. To measure geographical availability, we use 1,000 times the number of Probecat participants in a given state (Mexico is a federal entity with 32 states) as a proportion of the urban population in that state. (Up to recently, the program was targeted mainly to urban areas, and this holds for the 1994 data.) We used the same

measure of program availability at the state level for the two modules (school-based and in-service) because we did not have data on the availability of each module separately.

As noted by Ravallion and Wodon (1998, 2000), if the program availability at the state level is to be used as an instrumental variable for determining program participation, apart from individual-level variables, it is important to include in the regressions state-level variables that may affect program participation, as well as outcomes. A full set of state dummies could not be included because in this case the program availability at the state level would be collinear with the state dummies. State variables, however, could be included. We used 13 such variables as controls, which are not shown in the regression tables. They include primary and secondary schooling and spending indicators, population density, shares of the urban and indigenous populations, state-level income, and variables related to wealth and consumption (cars, water, and electricity).

To tackle the problem of sample selection for the participation in the program, we first ran probit regressions for men and women separately to analyze the determinants of participation in the two modules of Probecat available at the time of the survey (school-based and in-service). The probability of participating in the school-based module is denoted by $P1_{ij}$, and the probability for individual i living in state j to participate in the in-service module is denoted by $P2_{ij}$. X_{ij} is a vector of individual level variables, and Z_j is a vector of state-level variables for state j. The relative availability of Probecat in state j is denoted by AP_j. We estimated the following:

$$P1^*_{ij} = \gamma_{P1}'X_{ij} + \delta_{P1}'Z_j + \mu_{P1}AP_j + \mu_{P1}AP_j^2 + \varepsilon_{P1ij}$$
$$\text{with } P1_{ij} = 1 \text{ if } P1^*_{ij} > 0 \text{ and } 0 \text{ if } P1^*_{ij} \leq 0 \qquad (1)$$

$$P2^*_{ij} = \gamma_{P2}'X_{ij} + \delta_{P2}'Z_j + \mu_{P2}AP_j + \mu_{P2}AP_j^2 + \varepsilon_{P2ij}$$
$$\text{with } P2_{ij} = 1 \text{ if } P2^*_{ij} > 0 \text{ and } 0 \text{ if } P2^*_{ij} \leq 0 \qquad (2)$$

The results of these probits are given in table 2. Individuals between 15 and 55 are more likely to participate in the program than younger and older individuals, which corresponds to the eligibility rules. Individuals with at least the primary level of education completed also tend to participate more (in comparison with the excluded category in the survey, which corresponds to the illiterate and those not having completed primary school), again because having completed primary school

TABLE 2. PARTICIPATION REGRESSIONS (PROBITS) FOR PROBECAT BY MODULE

	School-based				*In-service*			
	Men		*Women*		*Men*		*Women*	
	dF/dx	*P > |z|*	*dF/dx*	*P > |z|*	*dF/dx*	*P > |z|*	*dF/dx*	*P > |z|*
Aged between 15 and 25	0.183	0.004	0.288	0.243	0.283	0.997	0.973	0.000
Aged between 16 and 55	0.229	0.000	0.419	0.084	0.378	0.997	0.984	0.000
Primary education completed	0.158	0.000	0.174	0.058	0.902	0.997	0.068	0.146
Secondary or postprimary training	0.321	0.000	0.317	0.000	0.325	0.997	0.021	0.593
Higher level	0.248	0.000	0.318	0.001	0.627	0.997	–0.075	0.030
Household head	0.080	0.010	0.153	0.033	0.000	0.606	–0.008	0.796
Married individual	0.024	0.404	0.104	0.043	0.001	0.092	0.012	0.564
Household with 2 workers	0.050	0.022	0.029	0.505	0.000	0.416	0.053	0.009
Household with 3^+ workers	0.042	0.078	0.027	0.601	0.000	0.915	0.042	0.078
Previous professional experience	–0.068	0.103	–0.067	0.366	–0.004	0.035	–0.217	0.000
Working experience in past year	0.108	0.006	0.115	0.061	0.000	0.891	0.021	0.434
Past year in firm 1–15 workers	0.103	0.002	0.243	0.000	0.001	0.309	0.026	0.433
Past year in firm 16–100 workers	0.035	0.355	0.156	0.055	0.003	0.041	0.094	0.048
Past year in firm 101–250 workers	0.121	0.006	0.056	0.609	0.001	0.301	0.123	0.072
Past year in firm 251^+ workers	0.005	0.897	–0.048	0.535	0.001	0.303	0.159	0.002
Worked 35–48 hours in past year	0.087	0.000	0.084	0.075	0.000	0.451	0.006	0.763
Worked 48^+ hours in past year	0.075	0.000	0.091	0.046	0.000	0.550	0.037	0.076
Income past year	0.000	0.000	0.000	0.001	0.000	0.664	0.000	0.911
Income squared past year	0.000	0.019	0.000	0.114	0.000	0.394	0.000	0.752
Program availability	0.107	0.000	0.092	0.014	0.003	0.005	0.155	0.015
Program availability squared	–0.004	0.001	–0.004	0.233	0.000	0.004	–0.018	0.020
Number of children	n.a.	n.a.	0.025	0.463	n.a.	n.a.	0.017	0.229
Number of children squared	n.a.	n.a.	–0.002	0.617	n.a.	n.a.	–0.003	0.198
Number of observations	2,160	n.a.	947	n.a.	2,160	n.a.	947	n.a.
Pseudo R^2	0.297	n.a.	0.233	n.a.	0.351	n.a.	0.331	n.a.

n.a. Not applicable.

Source: Authors' estimation from pooled Probecat and ENEU panel surveys (1993–94). Thirteen state-level control variables are included in the regression

is a requirement. Being married has an impact on participation only for women in the school-based module. There is also evidence that individuals with several workers in their household (apart from the individual considered in the sample) have higher participation rates. Having previous work experience is negatively correlated with participation (maybe because these individuals need less training), but having worked in the previous year is positively correlated (maybe because these individuals remain actively seeking employment). Having worked in firms with other workers (instead of having been self-employed) facilitates participation, as does the fact of having worked a large number of hours in the previous year. Higher incomes in the last year also influence positively the probability of participating. Finally, individual participation is positively correlated with the program availability at the state level. This is important because it confirms that state-level availability of the program is a valid instrumental variable. There is no need to discuss here the impact of the other 13 state-level variables that were included in the regressions, but not shown in the tables.

Impact of Probecat on the Length of Employment Search

We now consider the impact of Probecat on the length of the employment search. In the terminology of survival analysis, the survivor function $S(t)$ represents the length of unemployment after training (t is measured in months). Given $S(t)$, the hazard function $\lambda(t)$ denoting the chance of becoming employed (or the risk of remaining unemployed) at time t among the individuals who were not yet employed at that time is $\lambda(t) = -d(\log S(t))/dt$. The survivor curve can be specified as a function of program participation, individual characteristics, and state characteristics, so that $\lambda = \lambda(t; X, Z, P1, P2)$. In Cox's proportional hazard model, we have:

$$\lambda(t; X, Z, P1, P2) = \lambda_0(t) \exp(\gamma' X_{ij} + \delta' Z_j + \mu_1 P1_{ij} + \mu_2 P2_{ij}) \qquad (3)$$

Cox proposed a partial maximum likelihood estimation of this model in which the baseline function $\lambda_0(t)$ does not need to be specified. The relative chance of two individuals being employed (or the risk of remaining unemployed) could then be compared. Consider two identical individuals, except for their participation in the school-based and in-service Probecat modules. At any given time t, the ratio of the hazard rates for the two individuals, also referred to as the relative risk ratio, is $\exp(\mu_1/\mu_2)$. If μ_2 is larger than μ_1, all other things being equal, the individual having received the in-service training has a higher probability

of finding employment than the individual having received the school-based training. If both μ_1 and μ_2 are positive, the individuals receiving any one of the two forms of training are likely to find employment before the individuals who received no training.

The results of the Cox regressions are given in table 3. The first columns under the "naïve" heading were obtained by using the hazard function $\lambda_0(t)\exp(\gamma' X_{ij} + \delta' Z_j + \mu_1 P1_{ij} + \mu_2 P2_{ij})$ where $P1_{ij}$ and $P2_{ij}$ denote the actual participation in the two programs. The columns under the heading "control" were obtained by using instead $\lambda_0(t)\exp(\gamma' X_{ij} + \delta' Z_j + \mu_1 IP1_{ij} + \mu_2 IP2_{ij})$, where $IP1_{ij}$ and $IP2_{ij}$ denote the index values obtained from the estimation of the probits. The index values are the right-hand sides of the probit equations less the residuals (not the expected probabilities of participation, which can be computed using the normal distribution).

If Probecat participants are among the more dynamic individuals who are willing to make sacrifices in order to be trained and to find employment, we would expect that in comparison with naïve estimates, the impact of Probecat would be smaller when suitable controls are introduced for the endogeneity of program participation. This is exactly what we observed. In the naïve estimates, because we used a model that was very similar to that used in the 1995 study, we got coefficient estimates that were fairly close to those obtained in that study (see table 3). These naïve estimates indicate that training reduces the length of unemployment for both men and women. The impacts are apparently statistically significant at the 10 percent level for the school-based module, and at the 5 percent level for the in-service module. Yet, once we use the index values from the probits instead of the actual program participation indicators, these positive impacts vanish. We even observe negative impacts in the case of men for the school-based training, although these are not statistically significant at the 5 percent level.

The results obtained for the school-based module may not sound too surprising given the short period during which individuals received training. How then can it be that despite the requirement for employers to hire 70 percent of the trainees in the in-service module, we observe no positive impact on employment after controlling for the endogeneity of program placement? It must be that without the stipend (wage subsidy) provided by the government, firms participating in the in-service module would have hired the same workers anyway. This is referred to as a deadweight loss in the literature. In OECD countries, the impact of such deadweight losses has been shown to represent from 40 to 90 percent of all hires (Foley 1992).

Impact of Probecat on Earnings

Consider now the impact of Probecat on monthly earnings. For this, we used a standard Heckman sample selection model, but in a different way than in past evaluations of Probecat. Denote by log w the logarithm of the expected wage for an individual. This wage is not zero if and only if it is larger than the individual's reservation wage (otherwise, the individual would choose not to work). Denote the unobserved difference between the individual's expected wage and his reservation wage by Δ^*. The individual's expected wage is determined by a number of individual (essentially the individual's education and experience) and state variables. The difference between the individual's expected wage and his reservation wage is determined by the same variables, plus the number of children, the fact of being a household head, and the fact of being married. If we split the individual level variables into those that influence both expected earnings and the reservation wage (vector E) and the demographic variables that influence the reservation wage only (vector D), the standard Heckman model is

$$\Delta_{ij}^* = \phi_\Delta' E_{ij} + \pi_\Delta' D_{ij} + \eta_\Delta' Z_j + \alpha_\Delta P1_{ij} + \chi_\Delta P2_{ij} + \nu_{ij}$$
$$\text{with } \Delta_{ij} = 1 \text{ if } \Delta_{ij}^* > 0, \text{ and } 0 \text{ if } \Delta_{ij}^* < 0 \qquad (4)$$

$$\text{Log } w_{ij}^* = \phi_w' E_{ij} + \eta_w' Z_j + \alpha_w P1_{ij} + \chi_w P2_{ij} + \kappa_{ij}$$
$$\text{with Log } w = \log w^* \text{ if } \Delta = 1 \text{ and } 0 \text{ if } \Delta = 0 \qquad (5)$$

The above model controls for the endogeneity of labor force participation. We estimated this model first with the actual values of the participation dummies $P1_{ij}$ and $P2_{ij}$ (this is the naïve estimation in table 4), and next with the index values $\hat{P}1_{ij}$ and $\hat{P}2_{ij}$ from the probit regressions (this is the control estimation in table 4). By estimating the program participation equations first, and then using the standard Heckman model, we were able to control for both sources of bias at once (note that all the coefficients of the inverse Mills' ratios are statistically significant in table 4, which highlights the importance of controlling for endogenous participation in the labor force). Our parameter estimates are in principle consistent, although they may not be efficient because the first probit equation is estimated separately rather than with the Heckman model. How do the results of the naïve and control estimates compare? There are fewer differences than with the Cox model. In both the naïve and control estimations, the impacts of Probecat are negative

TABLE 3. COX REGRESSIONS FOR NUMBER OF MONTHS UNEMPLOYED

	Naïve				*Control*			
	Men		*Women*		*Men*		*Women*	
	Coef.	*P* > |z|	*Coef.*	*P* > |z|	*Coef.*	*P* > |z|	*Coef.*	*P* > |z|
Program impact in 1995 study								
Probecat—School-based	0.107	0.115	0.361	0.002				
Probecat—In-service	0.476	0.000	0.776	0.002				
New estimates of impact								
Probecat—School-based	0.121	0.094	0.221	0.055	–0.156	0.063	0.360	0.074
Probecat—In-service	0.436	0.000	0.742	0.000	–0.023	0.091	0.045	0.284
Other variables in estimation								
Aged between 15 and 25	1.031	0.000	0.664	0.203	1.363	0.000	0.388	0.546
Aged between 16 and 55	1.032	0.000	0.705	0.179	1.376	0.000	0.264	0.693
Primary education completed	0.054	0.672	0.726	0.001	0.362	0.031	0.569	0.016
Secondary or postprimary training	0.096	0.427	0.403	0.059	0.510	0.007	0.113	0.679
Higher level	0.000	0.997	0.434	0.049	0.370	0.045	0.045	0.868
Household head	0.221	0.004	0.204	0.170	0.285	0.001	0.082	0.630

Married individual	0.073	0.323	–0.292	0.012	0.126	0.090	–0.350	0.007
Household with 2 workers	0.035	0.540	0.506	0.000	0.059	0.323	0.511	0.000
Household with 3+ workers	0.132	0.037	0.505	0.000	0.170	0.009	0.511	0.000
Previous professional experience	0.041	0.769	–0.186	0.273	–0.096	0.502	–0.208	0.260
Working experience in past year	0.436	0.000	0.124	0.354	0.523	0.000	0.029	0.850
Past year in firm 1–15 workers	–0.028	0.748	0.024	0.871	0.094	0.339	–0.169	0.417
Past year in firm 16–100 workers	–0.099	0.327	0.172	0.334	–0.008	0.935	0.099	0.629
Past year in firm 101–250 workers	–0.159	0.262	0.060	0.814	–0.006	0.967	0.037	0.890
Past year in firm 251+ workers	–0.140	0.131	0.017	0.922	–0.097	0.303	0.111	0.524
Worked 35–48 hours in past year	0.078	0.131	0.154	0.135	0.135	0.022	0.088	0.425
Worked 48+ hours in past year	0.045	0.375	0.112	0.252	0.091	0.097	0.057	0.601
Income past year	0.000	0.066	0.000	0.196	0.000	0.653	0.001	0.046
Income squared past year	0.000	0.234	0.000	0.532	0.000	0.300	0.000	0.225
Number of children			0.043	0.587			0.013	0.872
Number of children squared			–0.014	0.268			–0.010	0.419
Number of observations	2,160		947		2,160		947	
Log likelihood	–13,407		–4,200		–13,412		–4,217	

Source: Authors' estimation from pooled Probecat and ENEU panel surveys (1993–94). Thirteen state-level control variables are included in the regression, but not shown in the table. See text for more details.

TABLE 4. HECKMAN REGRESSIONS FOR WAGES AND EMPLOYMENT

	Naïve				*Control*			
	Men		*Women*		*Men*		*Women*	
	Coef.	*P* > |z|	*Coef.*	*P* > |z|	*Coef.*	*P* > |z|	*Coef.*	*P* > |z|
Logarithm of wage								
Probecat—School-based	–0.020	0.763	–0.083	0.460	–0.204	0.000	–0.080	0.421
Probecat—In-service	–0.115	0.269	–0.300	0.021	–0.021	0.082	–0.032	0.432
Aged between 15 and 25	–0.070	0.698	–0.237	0.619	0.224	0.259	0.006	0.992
Aged between 16 and 55	0.159	0.376	–0.074	0.876	0.507	0.012	0.224	0.699
Primary education completed	0.049	0.684	–0.026	0.899	0.337	0.017	0.010	0.961
Secondary or postprimary training	0.182	0.107	0.117	0.552	0.562	0.000	0.169	0.418
Higher level	0.309	0.007	0.316	0.121	0.662	0.000	0.410	0.054
Previous professional experience	0.107	0.331	–0.287	0.036	0.164	0.136	–0.241	0.083
Working experience in past year	–0.199	0.012	0.054	0.645	–0.141	0.080	0.041	0.731
Constant	5.517	0.000	6.637	0.000	5.754	0.000	6.284	0.000
Lambda	–0.918	0.022	–1.138	0.041	–0.914	0.021	–1.175	0.044

Probability of working								
Probecat—School-based	0.018	0.842	0.211	0.046	0.075	0.312	0.291	0.004
Probecat—In-service	0.207	0.308	0.455	0.001	–0.038	0.061	0.033	0.407
Aged between 15 and 25	0.892	0.000	0.541	0.187	1.073	0.000	0.116	0.819
Aged between 16 and 55	0.810	0.000	0.482	0.245	0.937	0.000	–0.047	0.927
Primary education completed	0.041	0.813	0.574	0.003	0.241	0.274	0.469	0.014
Secondary or postprimary training	–0.043	0.796	0.191	0.286	0.106	0.637	–0.007	0.971
Higher level	–0.083	0.620	0.420	0.024	0.049	0.821	0.218	0.253
Previous professional experience	0.108	0.426	0.175	0.182	0.068	0.611	0.122	0.343
Working experience in past year	0.561	0.000	0.058	0.599	0.523	0.000	0.017	0.875
Household head	0.165	0.063	0.312	0.009	0.174	0.048	0.198	0.069
Married individual	0.039	0.655	–0.145	0.069	0.070	0.421	–0.167	0.013
Number of children			–0.023	0.635			–0.028	0.492
Number of children squared			0.000	0.978			0.002	0.674
Constant	0.700	0.299	–2.573	0.004	–1.001	0.324	–1.636	0.085
Number of observations	2,167		948		2,160		947	
Log likelihood	–3,044		–1,332		–3,020		1,328	

Source: Authors' estimation from pooled Probecat and ENEU panel surveys (1993–94). Thirteen state-level control variables are included in the regression but not shown in the table. See text for more details.

instead of being positive. While the levels of significance differ according to the method, these results at the least shed doubts on the positive impact of Probecat observed on earnings in past evaluations.

Conclusion

During the last 15 years, Probecat has been implemented by the Mexican government as a training program for the unemployed. Past evaluations of the program suggest that it is effective in reducing the length of unemployment and increasing earnings for participants. It could be, however, that these results were obtained because of inadequate consideration of the problem of sample selection. Using the data of the last evaluation of Probecat conducted by the Mexican Ministry of Labor, we have proposed another methodology for assessing impacts. According to our results, Probecat does not have large positive effects for participants. From a policy point of view, this finding suggests that the program may not be delivering its promise. From a scientific point of view, our results point to the sensitivity of evaluation results to the methodologies used to generate the results, which is of concern when these results are used for policy recommendations.

New initiatives have been taken in recent years to try to improve the functioning of Probecat. This includes the implementation of the PILEOT module since 1995, which is targeted at economically disadvantaged communities. The main requirement for participation in PILEOT is that the applicant must have basic literacy and numeracy skills, and be unemployed or underemployed. The module provides training for individuals and groups who intend to engage in self-employment or community-based productive activities. The training contents are demand-driven. It will be important to evaluate this new module rigorously, and to test the sensitivity of the evaluation results to the assumptions made and the techniques used in the evaluation process.

References

Carlson, S. 1991. "Mexico Labor Retraining Program: Poverty Alleviation and Contribution to Growth." Report No. 6. World Bank, Latin America and the Caribbean Technical Department, Regional Studies Program, Washington, D.C.

Dar, A., and I. S. Gill. 1998. "On Evaluating Retraining Programs in OECD Countries." *World Bank Research Observer* 13:79–101.

Fleck, S., and C. Sorrentino. 1994. "Employment and Unemployment in Mexico's Labor Force." *Monthly Labor Review*, November: 3–31.

Foley, P. 1992. "Local Economic Policy and Job Creation: A Review of Evaluation Studies." *Urban Studies* 29(3–4):557–98.

Goldberger, A. 1983. "Abnormal Selection Bias." In T. Amemiya and I. Olkin, eds., *Studies in Econometrics, Time Series and Multivariate Statistics*. San Diego, London, and Toronto: Academic Press.

Heckman, J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47(1):54–61.

Manski, C. 1996. "Review of the Probecat Program Evaluation Study." World Bank, Washington, D.C.

Ravallion, Martin, and Q. Wodon. 1998. *Evaluating a Targeted Social Program When Placement Is Decentralized*. Policy Research Working Paper 1945. World Bank, Washington, D.C.

———. 2000. "Does Child Labor Displace Schooling? Evidence on Behavioral Responses to an Enrollment Subsidy." *Economic Journal* 110(462):C158–75.

Revenga, A., M. Riboud, and H. Tan. 1994. "The Impact of Mexico's Retraining Program on Employment and Wages." *World Bank Economic Review* 8(2):247–77.

Rosenbaum, P., and D. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70:41–55.

———. 1985. "Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score." *American Statistician* 39:35–39.

STPS (Secretaría del Trabajo y Previsión Social). 1995. *Evaluación del Programa de Becas de Capacitación para Desempleados*. Mexico, D.F.: STPS.