Inter-American Development Bank
Banco Interamericano de Desarrollo
Latin American Research Network
Red de Centros de Investigación
Research Network Working paper #R-522

Revisiting the Employability Effects of Training Programs for the Unemployed in Developing Countries

by

Angel Calderón-Madrid

El Colegio de México

May 2006

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library

Calderón-Madrid, Angel.

Revisiting the employability effects of training programs for the unemployed in developing countries / by Angel Calderon-Madrid.

p. cm.
(Research Network Working papers ; R-522)
Includes bibliographical references.

1. Occupational retraining. 2. Employability—Developing countries— Effect of occupational retraining on. 3. Unemployment. I. Inter-American Development Bank. Research Dept. II. Latin American Research Network. III. Title. IV. Series.

370.113 C943--------dc22

1300 New York Avenue, N.W.
Washington, DC 20577

The Research Department (RES) produces a quarterly newsletter, *IDEA (Ideas for Development in the Americas)*, as well as working papers and books on diverse economic issues. To obtain a complete list of RES publications, and read or download them please visit our web site at: http://www.iadb.org/res

Abstract[1]

Data sets used for evaluations in developing countries do not lend themselves to measuring the impact of training programs on the re-employment dynamics of trainees. An exception is a data set collected for an evaluation conducted in 1994 on participants in a training program targeting the unemployed in Mexico. In addition to having a control group of eligible individuals who did not participate in the program, this data set is the only one with longitudinal data covering not only the length of unemployment episodes after the training of the respondent, but also the duration of his/her employment spells. We use this data and estimate the additional weeks individuals work as the result of training, relative to what would be the case without it. Based on hazard functions, we calculate a program's impact on both the time spent searching for a job and the time spent in that job. We show that a failure to distinguish between finding a "sustained" job versus finding "a job" can lead to misleading conclusions about a program's effectiveness. We also illustrate the need to correct for unobserved heterogeneity across individuals in hazard functions to avoid misleading implications in an evaluation.

[1] This paper grew out of a research project coordinated by James Heckman and Gustavo Márquez and sponsored by the Inter-American Development Bank's Research Network project initiative "Evaluating Training Policies in Latin America and the Caribbean Countries." Research assistance from Gonzalo Rangel is acknowledged, as well as the helpful comments received from the coordinators of this project and from Jeffrey Smith and Petra Todd at the IDB Research Network's meetings. The author also thanks Robert Lalonde for his comments at the UIA seminar in Mexico City on September 2003 and the participants of the 2002 LACEA Meeting held in Madrid, Spain. Angel Calderón-Madrid: acalde@colmex.mx.

1. Introduction

As a consequence of the democratization process in a number of developing countries, there has been a call for transparency in and critical assessment of public spending. Sometimes the same demand has been made by international agencies, such as the Inter-American Development Bank (IDB), Asian Development Bank (ADB) and World Bank, when specific public spending programs have been financed by them. This practice has led to a growing consensus that the future of active labor market programs in particular should be decided after adequate measurements of their impact have been conducted.

Unlike in developed countries, the lack of unemployment insurance and the very low savings of workers in developing countries make open unemployment unaffordable for most participants in the labor market. Faced with poor employment opportunities, unemployed workers are forced to engage in low-paying activities in which they are less productive than they would be otherwise. Because of this, the impact of programs targeted to the unemployed depends to a great extent on their effectiveness in achieving good job matches.

These adverse economic conditions also account for a public policy emphasis on programs that help individuals find a "sustained" job—as opposed to "any job." Therefore, an assessment of the impact of programs seeking to help the unemployed regain employment—regardless of how long they hold on to it—is only part of the feedback needed by policymakers to determine its effectiveness. Indeed, as suggested by Ham and Lalonde (1996), "[They] may prefer to fund a program that lengthens employment durations as opposed to one that shortens unemployment durations, because the former program is likely to lead to more stable job histories and greater human capital accumulation."[2]

In spite of the need for this kind of feedback, and for rigorous evaluations that may lead to insights regarding the successful implementation of programs in other countries, data sets used for evaluations in developing countries do not lend themselves to measuring the impact of training programs on the re-employment dynamics of trainees.[3] An exception is a data set collected for an

[2] Ham and Lalonde (1996), p. 176.

[3] Although discrepancies in the methods and interpretation of results still exist in developing countries, most empirical studies addressing the impact on hourly wages provide estimates with a minimum of required rigor. By contrast, the impact on employability, if addressed at all, has not adequately been dealt with. The methodological rigor followed in corresponding studies for developed countries (Ham and Lalonde, 1996; Bonnal, Fougere and Sérandon, 1997; and Eberwein, Ham and Lalonde, 1997 and 2002) has not permeated evaluations of programs in developing countries.

evaluation conducted in 1994 on participants in a training program targeting the unemployed in Mexico.[4]

In addition to having a control group of eligible individuals who did not participate in the program, this data set is the only one with longitudinal data covering not only the length of unemployment episodes after the training of the respondent, but also the duration of his/her employment spells. In spite of the richness of this data set, previous evaluations that have used it have concentrated exclusively on the program's impact on the reduction of time required to find a job and/or on the probability of finding one.[5] That is, they have addressed the problem of the employability of workers as is commonly done in other developing country studies.

In this paper we use this data set to calculate the training program's impact on the number of weeks needed by participants to find a job and on the time they spend in that job. We estimate the number of additional weeks within a year individuals would work, relative to what would be the case if they did not participate in the training program. These calculations are based on hazard functions. Our analysis allows for non-parametric correction of biases due to unobserved heterogeneity across individuals.[6]

We illustrate how a failure to eliminate biases originating from this source might lead to erroneous conclusions regarding the effectiveness of the program. In addition, with a nonparametric estimation of the time dependence of our hazard functions, we capture how hazard rates **out of employment** first increase and then start declining—as suggested in theoretical economic models of job-matching and turnover.[7]

The structure of the paper is as follows: the data set is described in Section 2, with an emphasis on information about re-employment dynamics. Section 2 also explains the procedure that matches a trainee with a member of the non-participant group. The statistical framework used in this evaluation is presented and discussed in Section 3. Section 4 presents the results and Section 5 discusses the impact of the program on its beneficiaries and presents the cost-benefit analysis, as applied to this type of program. The final section offers some conclusions.

---

[4] This program is called PROBECAT (*Programa de Becas de Capacitación para Desempleados*). It had a yearly registration of less than 200,000 people in 1994 (Revenga, Riboud and Tan, 1994) but expanded two-fold when the country entered a major recession in 1995. The number of trainees increased in subsequent years to reach a peak number of 590,000 in 2000.

[5] An exception is the multi-spell analysis of Calderón-Madrid and Trejo (2001), who examined how training affects the frequency or duration of up to three post-training job status spells.

[6] Meyer (1990) illustrates the development and advantages of this estimation strategy to labor market analysis. His strategy, in turn, is an extension of Heckman and Singer's (1984) approach.

[7] Jovanovic (1979).

2. Characteristics of the Data Set

2.1 Re-Employment Dynamics

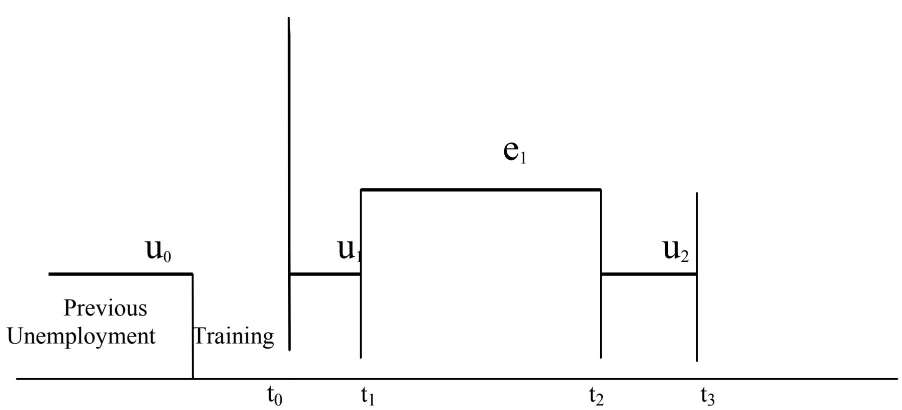
We use a data set that registered the re-employment dynamics of participants in a training program and of members of an appropriate control group. The program consisted of courses and training offered to unemployed individuals with previous working experience, lasting for an average of three months, in publicly funded institutions associated with either the Ministries of Education or Labor.[8]

In addition to time needed to find a job after their training, the survey registered whether respondents were still in their first post-training job at the moment of the interview (September 1994).[9]

If that was not the case, it registered the length of time they kept that job. The survey also captured the number of weeks in unemployment already experienced prior to joining the program.

The key feature of this survey is that it enabled us to construct job histories for each interviewed person. Figure 1 illustrates the example of an individual interviewed at point in time t_3. His job history shows that after the end of his training (t_0), he experienced an initial unemployment spell, u_1, (which captures how fast he found a job). At time t_1, he found an initial job, e_1, but left it after a number of weeks (which are represented by the difference between t_2 and t_1). The figure also shows that, prior to joining the program, this person had already experienced a number of weeks of unemployment, u_0.

Figure 1. Re-Employment Dynamics of Trainees



[8] Participants received allowances equivalent to the minimum wage while enrolled in the program, plus transportation and partial health insurance coverage.
[9] Their training began in the first quarter of 1993.

The survey was applied to 1,786 trainees and to 437 members of a control group. Of the participating trainees, 1,432 were male and 354 were female; the corresponding figures in the control group were 273 and 164, respectively.

2.1.1 Survivor Rates in Employment and Unemployment

Data on the time spent in each job category revealed that after finishing their training, 34 percent of male participants had already found a job within a month; one out of three was still unemployed by the end of the fourth month; one out of four remained unemployed by the middle of the year and 12 percent spent more than 360 days unemployed.

Table 1.
Kaplan Meier Survivor Functions
(Proportion remaining in each job status)

Interval in days	Participants				Control group			
	Men		Women		Men		Women	
	Unemploy-ment	Employ-ment	Unemploy-ment	Employ-ment	Unemploy-ment	Employ-ment	Unemploy-ment	Employ-ment
0 30	0.66	0.95	0.79	0.94	0.8	0.96	0.87	0.98
30 60	0.48	0.9	0.71	0.88	0.63	0.95	0.73	0.97
60 90	0.39	0.85	0.66	0.82	0.42	0.89	0.64	0.93
90 120	0.33	0.76	0.6	0.73	0.35	0.82	0.61	0.86
120 150	0.29	0.71	0.56	0.67	0.31	0.76	0.54	0.75
150 180	0.25	0.68	0.53	0.64	0.27	0.67	0.48	0.71
180 210	0.21	0.65	0.49	0.62	0.23	0.64	0.45	0.63
210 240	0.19	0.62	0.47	0.57	0.21	0.6	0.41	0.57
240 270	0.17	0.59	0.43	0.56	0.18	0.55	0.4	0.56
270 300	0.16	0.57	0.4	0.52	0.17	0.51	0.37	0.54
300 330	0.14	0.54	0.37	0.49	0.14	0.49	0.35	0.44
330 360	0.12	0.51	0.33	0.46	0.13	0.42	0.34	0.44
Number of observa-tions	1,432	1369	354	268	273	224	164	99
Censored spells	161	773	114	153	238	96	108	41
Completed spells	1,271	596	240	115	35	128	56	58

These data appear in the first column of Table 1, which shows the proportion of men who remained unemployed after finishing training. The second column of Table 1 indicates that while 76 percent of the men stayed in their job for at least four months, only two out of three lasted longer than six months in their first post-training job and only half stayed for at least one year. By contrast, the third and fourth columns show that the unemployment rates of female participants were significantly higher during the periods examined. Barely half of these women had found a job within six months and 32 percent remained unemployed a year after finishing the training. In addition, although the patterns of employment retention rates for these women were similar to those of the men, the survival rates for each date were relatively lower for women.

2.2 Matching Procedure

The control group for this evaluation consisted of eligible people who did not participate in the program. Unlike members of control groups used by experimental evaluations, they were not selected from the set of individuals who intended to participate but were refused enrollment into the program. Their input was needed to infer counterfactual outcomes for participants, namely what the beneficiaries of the program would have experienced had they not participated. The empirical survivor functions corresponding to members of the control group (the last columns of Table 1) indicate the percentage of individuals who stayed in each job category.

Due to the non-experimental nature of this evaluation, we used a matching method to pair each trainee with a member of the control group who had similar pre-program observable characteristics. The assumption justifying the use of matching methods for an unbiased evaluation of this program is this: conditional on a vector of observable characteristics, W, the employment and unemployment periods of the set of non-participants have the same distributions that participants would have experienced had they had not participated in the program. In view of the large number of pre-treatment observable characteristics, we applied the propensity score method variant of matching (Rosenbaum and Rubin, 1983). This variant has the advantage of reducing the dimensionality of the matching problem down to matching on one scalar, while considering the importance of all pretreatment variables included in the analysis. This scalar is the propensity score, P(W), defined as the probability of participation in the program conditional on pre-treatment observable characteristics.

A logit-regression was estimated to include each individual's propensity score, P(W). We incorporated as predictor variables in the logit regression the following pre-program observable characteristics of individuals (see the Appendix): geographic zone where the individual is located; four demographic characteristics: age, family position, education and civil status; time spent without a job previous to the training; three characteristics of the previous job: part- or full-time, formal or informal sector[10] and whether the individual was a wage earner or self-employed; three reason why the individual left the previous job: 1) marriage, care of children or relatives, 2) being fired or plant closure (identified as market reasons) and 3) unsatisfied with the job or left to study. We also included occupation in their last job.

We identified for each of the treatment group members a corresponding match from the control group, subject to the following two requirements: both of them had to be of the same sex and the absolute differences in their propensity score values could not be larger than 0.01. When there was more than one control candidate for a trainee, the matched person was randomly selected among non-participants fulfilling this criterion.

Following these criteria, we matched 1390 sets of men and 322 sets of women. Table 2 presents the parameters of the distribution of these two groups. As is shown, working with these subsets of the original sample, we end up with treatment and control groups that are distributed almost as if they were obtained from a "balanced experiment." The number of trainees that could be included in our analysis compared to those from the original data set was 89.5 percent of available men and 86.3 percent of available women. The incurred wastage of information is the cost to be paid for not having to worry about biases caused by differences in the support of the distribution of P(W) for each group, or in the shape of the distribution over the common support.

Table 2. Unemployed Individuals With Prior Working Experience

	Number of Observations	**Propensity Score**				
		Mean	**Median**	**Min**	**Max**	**Std. Dev.**
Men						
Control Group	1390	0.889	0.938	0.065	0.986	0.132
Participant Group	1390	0.891	0.94	0.065	0.996	0.133
Women						
	obs	Mean	Median	Min	Max	Std. Dev.
Control Group	322	0.848	0.912	0.126	0.977	0.169
Participant Group	322	0.849	0.911	0.123	0.986	0.169

[10] The formal sector is defined as having access to social security and other non-wage benefits.

3. Specification of the Model

To calculate the program's impact on the re-employment dynamics of the trainees, we follow the semi-parametric estimation procedure originally proposed and used by Meyer (1990)—namely, a piece-wise proportional hazard model that allows for unobserved heterogeneity correction. We discuss first the general characteristics of this model and then the specific assumptions used to address heterogeneity.

3.1 Hazard Function Specification

Survival models take as the point of departure the definition of a nonnegative continuous random variable T, which represents the spell duration with a density function, f(t). This function f(t) has a corresponding survivor function, simply defined as 1-F(t), i.e. as the probability that duration will equal or exceed the value t (where F(t) is the distribution function). The hazard function, h(t), is given by:

$$h(t)=f(t)/(1-F(t))$$

In this relationship, h(t) can be interpreted as an exit rate or escape rate from the state in consideration. It is the limit (as Δt tends to 0) of the probability that a spell terminates in interval (t, $t+\Delta t$), given that the spell has lasted t periods.[11]

We use the so-called "Mixed Proportional Hazard" (MPH) model. Its specification has two parts: a "baseline" hazard and a "systematic part." Using the assumption that the "systematic part" of the hazard takes the form of an exponential function, the hazard rate is thus multiplicative in all the separate elements of the covariates. In addition, we specify that the systematic part is composed of two parts: observed individual characteristics, X,[12] and a binary dummy variable, Z, indicating whether or not the individual is in the treatment group,[13] *viz*:

$$h(t|X,Z)=h_0(t)\exp(X\beta+Z\gamma)$$

[11] Some people who started a spell of employment/unemployment in a given job status may still have been in the same status when they were last interviewed. The data for these people are called censored, and they would constitute a problem for a standard regression model where the dependent variable was the length of the spell. If we exclude people with unfinished spells, we throw away part of the data set and introduce a serious bias against people with longer and more recent spells in each of the job statuses. Duration models have the distinct advantage of being able to handle censored data effectively (Kieffer, 1988).

[12] By construction, the duration of the first post-training job, te, starts after the moment at which the first spell of unemployment, tu1, is realized. A multi-spell specification of these models enables us to capture the dependence between states by including tu1 as an additional covariate in the hazard for te. (Van den Berg, 2000 and Calderón-Madrid and Trejo, 2001).

[13] This assumes that the different services provided by the multidimensional nature of the training program are adequately captured by a single binary variable.

The baseline hazard, $h_0(t)$, captures the common hazard among individuals in the population, while the systematic part captures the individual observed heterogeneity through the effect of a set of covariates on the hazard rate. As in Meyer (1990), we allow for a flexible duration dependence in our estimation by using a step function for $h_0(t)$. That is, we estimate a piecewise-linear baseline hazard with a number of breakpoints.

In turn, β is the vector of parameters to be estimated and so is the scalar γ. The larger the exponential of the parameters in β are, the more probable it is that the individual with characteristics represented by X will exit the job status, given that the spell has lasted t periods.

3.2 Semiparametric Specification of Unobserved Heterogeneity and Time Dependence

Also following Meyer (1990) in the specification of unobserved heterogeneity across individuals, we assume that, if this is present, it is independent of the covariates, its distribution can be approximated with two points of support and it enters the hazard function multiplicatively. Additionally, in view of the multi-spell nature of the re-employment dynamics under analysis, we assume that it is independent across spells. Hence, the hazard function to be estimated becomes:

$$h(t|X,Z)=\Omega_i h_0(t)\exp(X\beta+Z\gamma)$$

where Ω_i is a random variable that is assumed to be independent of X and Z. As it is commonly done, we assume a specific shape for the distribution of the variable Ω_i. In our estimations, the covariates of the systematic part X were: individual characteristics such as head of household, level of formal education, age and marital status, as well as: time spent without a job before the date at which training began; characteristics of his/her previous job according to whether it was in formal or informal sector, whether it was part or full time and if the person was self-employed or wage earner; type of occupation; and reasons why the previous job was left.[14]

4. Results

The hazards out of unemployment and of employment were estimated separately for males and females. The results are presented in Tables 3 through 5. We estimated two cases: one in which unobserved heterogeneity is controlled for and the other in which it is not. As stated in the previous section, we assume that:[15]

[14] These are presented and detailed in the Appendix.
[15] Our procedure maximized a likelihood function with respect to both the location and amount of probability mass at each of the points of support.

Unobserved heterogeneity is independent of the covariates, independent across spells and is drawn from a distribution with two points of support, v1 and v2, with probability p1 and p2, respectively.

As the last four rows of Tables 3 and 4 indicate, the statistically significant values of the two-point distribution v1 and v2 imply that controlling for unobserved heterogeneity was required in the case of exits from unemployment, both for female and male participants. By contrast, we rejected the hypothesis of biases due to unobserved differences across individuals in the case of hazards out of employment for male and female participants. There are two explanations why this result differs from that of unemployment. First, only individuals who found a job are included in the estimations in Table 5, after a corresponding matching procedure was conducted when required. That is, we might have been left with a group without unobserved heterogeneity across individuals; those that presented it remained unemployed. Second, unobserved heterogeneity, but dependent across spells, might prevail, but this is not captured by our estimations because we explicitly ruled it out with our statistical assumptions.[16]

The results of the coefficients for the treatment variable Z in Columns 2 and 5 in these tables indicate that ignoring unobserved heterogeneity leads to erroneous interpretations about the effectiveness of the program in helping both females and males out of unemployment. The effect of the program on hazard rates out of unemployment and of employment is given by the estimated coefficient for this dummy variable, which indicates program participation. A positive impact on re-employment dynamics is implied when the exponential value of this coefficient is above (below) unity in hazards out of unemployment (employment).

It is also interesting to mention two counter-intuitive results arising in the last columns of Table 3. According to the 11th, 12th and 14th rows, hazards out of unemployment have the following implications when biases due to unobserved heterogeneity are not eliminated: a) female heads of households do not leave unemployment faster than daughters living with their parents and b) women who left their previous job because the plant closed, their contract ended or because they were fired (market reasons) do not leave unemployment faster than the rest (except for those women who left their previous job due to marriage or care of their children or relatives). Once unobserved heterogeneity is controlled for, these results no longer hold: female heads of households find jobs twice as fast as daughters do and women who leave unemployment faster are those that were in that

[16] A multiple-spell estimation, as in Calderón and Trejo (2001) that relaxes the assumption of independence between the covariates and the heterogeneity term is left for future research, as a way to have a more general specification.

status because their contract finished, because they were fired or because the plant they were working at closed.[17]

Our nonparametric estimation of the time dependence of the hazard function indicates that no flexibility of duration dependence was required for the case of exits out of unemployment (first and second rows). By contrast, the first five rows in tables with hazards rates out of employment indicate that these first increase and then start declining, as suggested in theoretical economic models of job-matching and turnover (Jovanovic, 1979).

The main conclusions of the impact of the program on the participants' re-employment dynamics are as follows:

1. We reject the hypothesis that the impact on female participants' transitions out of unemployment is statistically insignificant, once biases due to unobserved heterogeneity are eliminated. This implies that the program is effective in helping them to find a job more quickly. The opposite is the case for male participants: once we control for unobserved heterogeneity, we conclude that the program does not help them to find a job more quickly. Our results indicate that, for females there is approximately a reduction of 35 percent in the time needed to find a job with respect to the counterfactual (not taking a training program, as captured by the comparison group).[18]
2. By increasing the time they hold on to their jobs relative to what would be the case if they had not participated, female and male participants benefited from the services provided by the program.

This result for male participants illustrates another case that could lead to erroneous interpretations about the effectiveness of the program: an evaluation might suggest that the program is not effective in improving employability of participants, when there is an exclusive concentration on the program's impact on leaving unemployment.[19]

[17] As it is shown in the second column of Table 3, they find a job more than twice as fast asthose who left it to study and almost 50 percent faster than those who left their jobs due to dissatisfaction with them.

[18] For females' hazard rates out of unemployment, the impact of the program in terms of percentage reduction in time required to find a job is calculated as follows. The exponential of the parameter affecting the variable Z (dummy for participating in the program) is 1.55. With this figure, we get an impact of approximately 0.645 (1/1.55) in the expected time to find a job, which is attributed to the training. This implies a 35 percent reduction in the time required to find a job, relative to those who did not participate in the program. In view of the unemployment hazard function specification, the standard normalization for unobserved heterogeneity sets an expected value equal to 1. Hence in this case we refer to observations whose average value of the unobserved heterogeneity has an unadjusted hazard.

[19] This is the conclusion that Wodon and Minowa (2001) reached in evaluating this program. (They only considered the

As shown in our analysis, the impact on re-employment dynamics is positive because participants retain their jobs for a longer period of time, even if the program is not effective in helping them find a job more quickly.

Table 3. Hazard Functions, Female Unemployment Spell

	Controlling for Unobserved Heterogeneity			Not Controlling for Unobserved Heterogeneity		
	β	exp(β)	pvalue	β	exp(β)	pvalue
h(1)	0.0001		0.08	0.0027		0.02
h(2)	0.0003		0.07	0.0025		0.02
Left job due to marriage or care of relative	-1.51	0.2	0.00	-1.27	0.28	0.00
Left job due market reasons	0.80	2.23	0.00	-0.01	0.99	0.48
Left job voluntarily due to dissatisfaction or change of address	0.38	1.46	0.04	-0.07	0.94	0.37
zone2	-0.05	0.95	0.42	-0.34	0.71	0.02
zone3	0.03	1.03	0.46	-0.46	0.63	0.00
zone5	-0.96	0.38	0.00	-0.43	0.65	0.01
zone6	-0.47	0.63	0.03	-0.51	0.60	0.00
Head of household	1.06	2.87	0.00	0.38	1.46	0.03
Daughter	0.35	1.42	0.13	0.44	1.55	0.04
Age	0.01	1.01	0.24	0.02	1.02	0.02
Full time wage-earner, formal sector	-0.17	0.84	0.28	0.34	1.40	0.06
Full time wage-earner, informal sector	-0.86	0.42	0.00	-0.19	0.83	0.20
Part time wage-earners	-1.36	0.26	0.00	-0.62	0.54	0.01
Full time self employed	-0.83	0.44	0.02	0.15	1.16	0.31
Single	0.10	1.10	0.36	0.04	1.04	0.42
Unempl. between one and two months	1.53	4.63	0.00	0.60	1.83	0.00
Unempl. between two and three months	0.81	2.24	0.00	0.11	1.12	0.29
Unempl. between three six months	0.34	1.41	0.10	-0.29	0.75	0.07
Unempl. more than six months	-0.36	0.70	0.09	-0.42	0.66	0.02
School2	0.40	1.50	0.15	0.36	1.43	0.11
School3	-0.20	0.82	0.35	-0.26	0.77	0.28
School4	1.50	4.50	0.00	0.42	1.52	0.10
School5	0.48	1.62	0.13	0.27	1.31	0.18
School6	0.10	1.10	0.42	0.25	1.28	0.21
School7	-0.16	0.85	0.37	0.03	1.03	0.47
School8	-1.52	0.22	0.00	-0.30	0.74	0.22
School9	-0.64	0.53	0.11	-0.63	0.53	0.07
Z (treatment group=1, 0 otherwise)	0.44	1.55	0.00	-0.25	0.78	0.02
V1	1.00		0.00	1.00		0.00
V2	26.91		0.00			
P1	0.34		0.00	1.00		0.00
P2	0.66		0.00			

impact of training on hazard out of unemployment).

Table 4. Hazard Functions, Male Unemployment Spell

	Controlling for Unobserved Heterogeneity			Not Controlling for Unobserved Heterogeneity		
	β	exp(β)	pvalue	β	exp(β)	pvalue
h(1)	0.01		0.00	0.01		0.00
h(2)	0.03		0.00	0.02		0.00
zone2	-0.25	0.78	0.00	-0.11	0.90	0.04
zone3	-0.28	0.76	0.00	-0.10	0.91	0.07
zone4	-0.15	0.86	0.13	-0.11	0.90	0.16
zone5	-0.29	0.75	0.00	-0.23	0.80	0.00
zone6	-0.22	0.80	0.00	-0.16	0.85	0.01
Left job due to marriage or care of relative	0.54	1.72	0.06	0.29	1.34	0.19
Left job due market reasons	0.35	1.42	0.00	0.32	1.38	0.00
Left job voluntarily due to dissatisfaction or change of address	0.40	1.50	0.00	0.45	1.57	0.00
Head of household	0.07	1.07	0.21	0.20	1.22	0.00
Age	-0.02	0.98	0.00	-0.02	0.98	0.00
Full time wage-earner, formal sector	0.26	1.30	0.03	0.20	1.22	0.03
Full time wage-earner, informal sector	0.26	1.30	0.03	0.31	1.36	0.00
Part time wage-earners	-0.73	0.48	0.00	-0.41	0.67	0.00
Full time self employed	0.12	1.13	0.20	0.12	1.13	0.14
Single	-0.63	0.53	0.00	-0.34	0.71	0.00
Unempl. between one and two months	0.01	1.01	0.45	0.02	1.02	0.39
Unempl. between two and three months	-0.33	0.72	0.00	-0.15	0.86	0.03
Unempl. between three six months	-0.12	0.89	0.10	-0.09	0.92	0.10
Unempl. more than six months	-0.91	0.40	0.00	-0.65	0.52	0.00
School2	-0.39	0.68	0.00	-0.15	0.86	0.09
School3	-0.30	0.74	0.19	-0.30	0.74	0.19
School4	0.13	1.14	0.16	0.03	1.04	0.38
School5	-0.20	0.82	0.05	-0.13	0.88	0.11
School6	-0.38	0.69	0.04	-0.31	0.74	0.05
School7	-0.06	0.94	0.35	-0.07	0.93	0.30
School8	-0.29	0.75	0.03	-0.26	0.77	0.02
School9	-0.16	0.85	0.16	-0.10	0.91	0.24
Z (treatment group=1, 0 otherwise)	-0.04	0.96	0.24	0.11	1.12	0.00
V1	1.00		0.00	1.00		0.00
V2	16.44		0.00			
P1	0.74		0.00	1.00		0.00
P2	0.26		0.00			

Table 5. Employment Spell

	Male			Female		
	β	exp(β)	pvalue	β	exp(β)	Pvalue
h(1)	0.017		0.02	0.05		0.11
h(2)	0.025		0.02	0.06		0.11
h(3)	0.029		0.02			
h(4)	0.018		0.02			
h(5)	0.026		0.02			
zone2	0.28	1.32	0.01	0.10	1.10	0.34
zone3	-0.06	0.94	0.29	-0.10	0.90	0.35
zone4	0.14	1.15	0.05	-0.31	0.73	0.17
zone 5	0.17	1.18	0.03	-0.16	0.85	0.26
Left job due to marriage or care of relative	0.44	1.55	0.00	-0.48	0.62	0.12
Left job due market reasons	-0.02	0.98	0.45	-1.09	0.34	0.00
Left job voluntarily due to dissatisfaction or change of address	-0.16	0.85	0.05	-0.61	0.54	0.03
Head of household	-0.18	0.84	0.03	0.89	2.43	0.00
Age	-0.22	0.81	0.00	-0.15	0.86	0.15
Full time wage-earner, formal sector	0.19	1.21	0.14	0.20	1.22	0.26
Full time wage-earner, informal sector	-0.23	0.80	0.10	0.09	1.09	0.39
Part time wage-earners	0.24	1.27	0.14	0.55	1.74	0.08
Full time self employed	0.25	1.29	0.08	-0.13	0.88	0.40
Single	-0.03	0.97	0.36	-0.12	0.89	0.35
desc2	0.36	1.44	0.03	-0.12	0.89	0.35
desc3	-0.60	0.55	0.28	-0.57	0.57	0.10
desc4	0.49	1.64	0.00	0.52	1.68	0.26
desc56	0.04	1.05	0.41	-0.26	0.77	0.28
desc7	0.12	1.12	0.34	-0.39	0.68	0.21
desc8	0.25	1.28	0.12	-0.49	0.61	0.22
desc9	0.03	1.03	0.45	0.12	1.12	0.43
desc10	-0.35	0.70	0.08	-1.26	0.28	0.13
Z (treatment group=1, 0 otherwise)	-0.18	0.84	0.00	-0.42	0.66	0.02
V1	1.00		0.00	1.00		0.00
V2	1.00		0.01	1.00		0.00
P1	0.14		0.00	0.46		0.00
P2	0.86		0.44	0.54		0.28

5. Incorporating Impact on Re-Employment Dynamics in the Cost-Benefit Analysis of a Program

The analysis presented here highlights cases in which the benefits attributed to improving the re-employment dynamics of participants might, on their own, compensate for the cost of a training program. That is, it indicates that a program’s evaluation must go beyond the impact on wages of beneficiaries or on the time out of unemployment. With estimates such as the ones conducted here, an integral cost-benefit analysis can quantify, in addition to a program’s impact on wages, its impact on earnings due to additional weeks individuals work in a year relative to what would have been the case if they had not joined the program.

To illustrate this point, we use a parameter that captures the effect of the program on hazard rates out of employment and calculate the percentage increase in the time participants retained their jobs relative to what would have been the case if they had not participated in the program. The average impact on working days attributed to program participation is presented in Table 6:[20] 69 female participants and 43 male participants. With these results, and an assumption of post-training wages remaining at pre-treatment levels, a first approximation to benefits in yearly income earnings of participants can be obtained for cost-benefit analysis.

Table 6. Average Number of Days Spent in Each Status

	Female Average Number of Days Spent in each status during the year		**Male** Average Number of Days Spent in each status during the year	
	If Treated	If Untreated	If Treated	If Untreated
Unemployment	75.48	144.94	87	130.82
Employment	198.46	129	257.83	214

6. Conclusions

We worked with a data set that registered the duration of unemployment and of employment after training was provided by a program targeted to the unemployed. This information—a set of variables associated with characteristics of their previous work and with their demographic profile—enabled us to estimate the program's impact on re-employment dynamics. Our results indicate that this program, which has been offered by the Mexican government for more than a decade, is more than a safety net providing temporary relief for the unemployed, as has been suggested in studies using the same data set but ignoring a number of issues raised in this paper.[21]

As might be the case with other training programs in developing countries, the impact on earnings of this program's trainees attributed to additional weeks working in a year—even in the

[20] The inverse of this parameter was multiplied by the average number of days that non-participants were employed during the year.

[21] See Giugale, Lafurcade and Nguyen, 2001.

absence of a major improvement in daily wage—might be a benefit large enough to compensate for a large part of the costs of the program.

Our results indicate that, once male and female participants find a job, they are employed for a longer period of time relative to what would have been the case had they not benefited from the program. They also show that women who participate in the program also benefit from shorter periods of time spent looking for a job. Our analysis showed that testing for unobserved heterogeneity must be an integral part of this type of estimations. Two misleading results were obtained when unobserved heterogeneity was not accounted for: one was hazard functions implying counter-intuitive results, such as female heads of household spending more time looking for jobs than other women in the labor market. The other was the fact that the impact of the program on unemployment spells of participants was not well estimated.

References

Abbring J.H., and G.J. Van den Berg. 2003. "The Nonparametric Identification of Treatment Effects in Duration Models." *Econometrica* 17: 1491-1517.

Bonnal, L., D. Fougere and A. Sérandon. 1997. "Evaluating the Impact of French Employment Policies on Individual Labour Market Histories." *Review of Economic Studies* 64: 683-713.

Calderón-Madrid, A. and B. Trejo. 2001. "The Impact of the Mexican Training Program for Unemployed Workers on Re-employment Dynamics and on Earnings." Inter-American Development Bank Research Network Draft Working Paper.

Eberwein, C., J. Ham and R. Lalonde. 1997. "The Impact of Being Offered and Receiving Classroom Training on Employment Stories." *Review of Economic Studies* 64: 655-682.

----. 2002. "Alternative Methods of Estimating Program Effects in Event History Models." *Labour Economics* 9: 249-278.

Giugale M., O. Lafurcade and V. Nguyen, editors. 2001. *Mexico: A Comprehensive Development Agenda for the New Era*. The World Bank, Washington D.C.

Ham, J. and R. Lalonde. 1996. "The Effect of Sample Selection and Initial Conditions in Duration Models: Evidence From Experimental Data on Training." *Econometrica* 64: 175-205.

Heckman, J., and B. Singer. 1984. "A Method for Minimizing the Distributional Assumptions in Econometric Models for Duration Data." *Econometrica* 52: 271-320.

Jovanovic, B. 1979. "Job-Matching & the Theory of Turnover." *Journal of Political Economy* 87(5): 972-990.

Kieffer, N.M. 1988. "Economic Duration Data and Hazard Functions." *Journal of Economic Literature* 26: 646-679.

Meyer, B.D. 1990. "Unemployment Insurance and Unemployment Spells." *Econometrica* 58(4): 757-782.

Revenga, A., M. Riboud and H. Tan. 1994. "The Impact of Mexico's Retraining Program on Employment and Wages." *The World Bank Economic Review* 8(2): 247-278.

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

Van den Berg, G. 2000. "Duration Models: Specification, Identification, and Multiple Durations." In: J. Heckman and E. Leamer, editors. *Handbook of Econometrics*, *Volume 5*. Amsterdam: Elsevier.

World Bank. 2000. "Securing Our Future in a Global Economy." *World Bank Latin American and Caribbean Studies*. Washington, D.C.

Wodon, Q., and M. Minowa. 2001. "Training for the Urban Unemployed: A Reevaluation of Mexico's Probecat." In: S. Devaradjan, F. Hasley, and L. Squire, editors. *World Bank Economists' Forum*. Washington D.C.: The World Bank.

Appendix 1.

Variables used as determinants of the probability of program participation functions and as co-variates in the survival models:

Reasons why the previous job was left
Marriage, childbearing care of children or other relative, equals 1, 0 otherwise.
Left their job due to market reasons (fired, end of contract), equals 1, 0 otherwise.
Left their job voluntarily because of a change of address or job dissatisfaction, equals 1, 0 otherwise.
Left their job to study, equals 1, 0 otherwise.
Geographic Region
Zone 1: In Western region of Mexico, equals 1, 0 otherwise.
Zone 2: In Northern region of Mexico, equals 1 for persons, 0 otherwise.
Zone 3: In the Coast of Mexico, equals 1 for persons, 0 otherwise.
Zone 4: In Bond (maquiladora) Northern Region of Mexico, equals 1 for persons, 0 otherwise.
Zone 5. In the South states of Mexico, equals 1 for persons, 0 otherwise.
Zone 6: In Mexico City and Central Area of Mexico, equals 1 for persons, 0 otherwise.
Unemployment duration before the beginning of the training program
Less than one month equals 1, 0 otherwise.
Between one and two months, equals 1, 0 otherwise.
More than two and up to three months equals 1, 0 otherwise.
More than three and up to six months equals 1, 0 otherwise.
More than six months equals 1, 0 otherwise.
Characteristics of previous job
Formal sector , wage earner and worked more than 35 hours: equals 1, 0 otherwise.
Formal sector, wage earner and worked less than 35 hours: equals 1, 0 otherwise.
Formal sector, non-wage earner (i.e. self-employed) and worked less than 35 hours: equals 1, 0 otherwise.
Informal sector, wage earner and worked less than 35 hours: equals 1, 0 otherwise.
Informal sector, non-wage earner (i.e. self employed) and worked less than 35 hours: equals 1, 0 otherwise.
Formal or informal sector, non-wage earner (i.e. self-employed) and worked less than 35 hours: equals 1, 0 otherwise.
Gender: Equals 1 if female, 0 if male
Age: Units of this variable are in years divided by 10.
Family position
Head of household: equals 1, 0 otherwise.
Second salary in household: equals1, 0 otherwise.
Son, daughter or other position different from the above: equals 1, 0 otherwise.
Civil Status
Single: equals 1, 0 otherwise.
Married or “free union”: equals 1, 0 otherwise
Divorced or widowed: equals 1, 0 otherwise.
Education
School1: incomplete primary equals 1, 0 otherwise.
School2: complete primary school and incomplete secondary education equals 1, 0 otherwise.
School3: post-primary courses equals 1, 0 otherwise.
School4. incomplete secondary school education equals 1, 0 otherwise.
School5: complete secondary education or incomplete post-secondary school training equals 1, 0 otherwise.
School6: complete post-secondary school training courses equals 1, 0 otherwise.
School7: incomplete high school education equals 1, 0 otherwise.
School8: complete high school education equals 1, 0 otherwise.
School9: education above the previous one equals 1, 0 otherwise.
Occupation in previous job
Ocu1: technician equals 1, 0 otherwise.
Ocu2: agricultural activities equals 1, 0 otherwise.
Ocu3: handicraft and repairing activities equals 1, 0 otherwise.
Ocu4: fix machinery operator equals 1, 0 otherwise.

Ocu5: assistant in repairing and maintenance activities equals 1, 0 otherwise.
Ocu6: drivers and assistant of machinery handling equals 1, 0 otherwise.
Ocu7: administrative activities equals 1, 0 otherwise.
Ocu8: trade and selling activities equals 1, 0 otherwise.
Ocu9: personal services in established places equals 1, 0 otherwise.
Ocu10: domestic services equals 1, 0 otherwise.

Appendix 2.

Table A1.
LOGIT Regression for the Matching Procedure

Variables	Coef	Z
Left job due to marriage or care of relative	-1.81	-4.70
Left job due market reasons	-2.02	-7.87
Left job voluntarily due to non-satisfaction or change of address	-1.21	-4.74
Zone1	1.73	5.77
Zone2	0.61	2.12
Zone3	0.32	1.10
Zone5	1.91	5.44
Zone6	1.15	3.79
Head of household	0.25	0.87
Son/Daughter	-0.04	-0.11
Single	-0.58	-2.44
Un. between one and two months	0.18	0.87
Un. more than two and up to three months	0.34	1.44
Un. more than three and up to six months	0.75	3.48
Un. more than six months	0.47	2.12
Full time wage-earner, formal sector f.f.m35	1.80	8.06
Part time wage-earners	1.65	5.29
Full time self employed	1.74	6.24
Full time wage-earner, informal sector i.f.m35	1.61	6.76
Education (5 categories)		
Previous job occupation (9 types)		
Sex	-.40	-2.25
Age	0.41	8.45
Age Squared	-0.005	-7.98

Table A1., continued

Variables	Coef	Z
School2	1.33	4.84
School3	0.87	1.44
School4	2.19	6.94
School5	2.66	9.34
School6	2.13	4.02
School7	0.67	2.02
School8	2.42	6.91
School9	2.70	7.56
Ocu1	0.33	1.06
Ocu2	2.90	4.98
Ocu3	1.01	4.38
Ocu4	1.57	4.87
Ocu5	1.50	6.09
Ocu6	0.23	0.64
Ocu7	0.81	3.23
Ocu8	1.35	5.13
Ocu9	1.24	4.33
Ocu10	1.61	3.58
Constant	-8.80	-8.51

Number of obs = 2,223
LR $chi^2(41)$ = 792.76
Prob > chi^2 = 0.0000
Log likelihood = -839.57331
Pseudo R^2 = 0.3207