IDP155

THE WORLD BANK

Internal Discussion Paper

EUROPE AND CENTRAL ASIA REGION

Report No. IDP-155

Costs and Effectiveness of Retraining in Hungary

by
Indermit S. Gill
and
Amit Dar

October 30, 1995

Office of the Vice President

Europe and Central Asia Region

ACKNOWLEDGMENTS

This paper is based on notes prepared for the seminar on "Evaluating Active Labor Programs," held on April 20, 1995, and was prepared under the general guidance of Kathie Krumm, Principal Economist, ECAVP, and Jane Armitage, Manager, Labor Markets and Social Insurance Group, PSP. It draws on material in papers by Christopher O'Leary ("An Impact Analysis of Employment Programs in Hungary"), Gyula Pulay ("Measuring the Efficiency of Employment Policy in Hungary") and background notes prepared for the seminar by David Fretwell. The paper also reflects comments by Alan Abrahart, Christine Allison, Mary Canning, Michal Rutkowski, Zafiris Tzannatos, Michael Walton and seminar participants.

TITLE: COSTS AND EFFECTIVENESS OF RETRAINING IN HUNGARY

AUTHORS: Indermit S. Gill and Amit Dar

ABSTRACT

As an early reformer, Hungary provides valuable lessons for other transition economies. Since 1989, sharp declines in GDP have been accompanied by rising unemployment and falling real wages. The unemployment rate appears to have stabilized since 1993 but employment continues to decline, reflecting continuing withdrawal from the formal labor market. In response to rising unemployment, the government has instituted labor programs since 1990. Among these programs are unemployment benefits and placement services, retraining, public service employment (public works), and wage subsidies. O'Leary (1995) has evaluated the effectiveness of retraining and public service employment, using a range of econometric techniques. This paper discusses the different techniques that can be used to evaluate these programs, provides guidance on which techniques are suitable and sufficient in transition economies, summarizes and explains the evaluation results for retraining, and incorporates cost information available from other sources for these programs to illustrate how simple cost-benefit analysis can be used to guide decision-making on active labor programs.

TABLE OF CONTENTS

COSTS AND EFFECTIVENESS OF RETRAINING IN HUNGARY

SUMMARY

For Hungary, evaluations of retraining and training programs and cost data are available. Evaluations using different quasi-experimental techniques - matched pairs, regression-adjusted, and selectivity-corrected - yield different results for retraining. When unadjusted for differences in characteristics of program participants and control groups, estimates of effectiveness indicate that retraining significantly raises the probability of re-employment. But when observable characteristics of treatment and control groups are taken into account, retraining - and training, since about 40 percent of trainees had not worked previously - is only marginally successful, at best increasing the probability of finding employment by six percent. Further controls for unobservable attributes lead to ambiguous results. Retraining is not at all successful in raising earnings.

Preliminary analysis indicates that retraining is a substitute for attributes that lead to higher re-employment probabilities in the absence of any intervention, e.g., being younger, more educated, and from more dynamic regions. That is, the program's value-added (in terms of ***improving*** labor market outcomes) is greater for relatively disadvantaged job-seekers. This finding may be country-specific, and it is necessary to determine this for other countries or regions before implementing large-scale active labor programs. For the programs evaluated in Hungary, focussing more on job-seekers who lack these attributes would appear to better serve ***both*** equity and efficiency objectives than simply ensuring support for programs whose retrainees have high re-employment probabilities and earnings gains.

This finding also highlights the usefulness of rigorous impact evaluations, i.e., those which net out the effects of such attributes in determining program effects, and underscores the need to agree upon a reliable, easily interpretable and feasible technique to evaluate the efficiency and equity effects of labor programs. Using both analytical rigor and feasibility as criteria, we argue that the preferred evaluation technique is matched pairs analysis, where trainees are compared with a subset of the control group whose characteristics most resemble their own.

Private cost-benefit analysis that brings together costs of retraining in Hungary (an average of approximately $900 per trainee in 1994) and effectiveness results reveals that at under reasonable assumptions regarding the durability of program effects and the amount and duration of unemployment benefits, it would likely take over 30 years to recover program costs at a five percent discount rate. That is, it seems difficult to justify retraining programs as currently designed on economic considerations alone, even in countries such as Hungary where governments provide unemployment benefits.

COSTS AND EFFECTIVENESS OF RETRAINING IN HUNGARY

I. INTRODUCTION

Rising unemployment and falling real wages are a vivid and costly aspect of transition. In large part, the problem is a fall in labor demand, though since the skill-mix required has changed, mismatches between workers and jobs also exist. Long-run unemployment is particularly pernicious, and active labor programs (e.g., retraining, public service employment, wage subsidies) are being used to deal with this problem. Little reliable information exists as a guide to their effectiveness: these programs have not been rigorously evaluated even in OECD countries (except the US), and the Bank entered this area with limited experience. An exception is Hungary, for which rigorous evaluations and cost data are available (O'Leary, 1995, and Pulay, 1995). This paper brings together effectiveness and cost information to arrive at crude cost-benefit measures for retraining in Hungary.

SELECTED ECONOMIC INDICATORS, HUNGARY 1989-94

Indicator	**1989**	**1990**	**1991**	**1992**	**1993**	**1994**(est)	**1989-93**
	(Real % change over previous year)						
GDP (constant prices) growth	-0.2	-3.5	-11.9	- 3.0	-0.8	2.0	-18.0
Real wages	0.9	-3.7	- 4.0	1.4	-0.8	3.1	- 7.2
Labor income (gross)	0.4	-2.9	-11.7	-4.6	-4.6	0.4	-17.4
Cash transfers (gross)	7.3	-2.9	0.7	4.1	-2.4	-6.3	0.0
Employment	-0.5	-0.6	- 3.2	-10.6	-1.2	-3.6	-20.9
	(Annual % rate)						
Unemployment rate	0.5	1.0	4.1	10.7	13.4	12.5	-18.0

Source: World Bank (1995).

Table 1

As an early reformer, Hungary provides valuable lessons for other transition economies. Since 1989, sharp declines in GDP have been accompanied by rising unemployment and falling real wages. The unemployment rate appears to have stabilized since 1993 but employment continues to decline, reflecting continuing withdrawal from the (formal) labor market. In response to rising unemployment, the government has instituted labor programs since 1990. Among these programs are unemployment benefits and placement services, retraining, public service employment (public works), and wage subsidies. O'Leary (1995) evaluates the effectiveness of retraining and public service employment, using a range of econometric techniques. In this paper, we discuss alternative techniques that can be used to evaluate these programs, provide some guidance on which techniques are suitable and sufficient in transition economies, summarize the evaluation results for retraining, and incorporate cost information to illustrate how simple cost-benefit analysis can be used to guide decision-making.

II. EVALUATION TECHNIQUES

State-of-the-art evaluation techniques for labor programs can be broadly classified into classically designed experiments and quasi-experimental studies. Classically designed experiments require selection of both the "control" and "treatment" groups - those who receive the assistance and those who do not, respectively - prior to the intervention. In quasi-experimental evaluations the treatment and control groups are selected after the intervention. Quasi-experimental evaluations are of three types - regression-adjusted for observables, selectivity-corrected (regression-adjusted for both observables and unobservables) and matched pairs.

1. CLASSICALLY DESIGNED (RANDOMIZED) EXPERIMENTS

This technique was originally developed to test drug-effectiveness. Program participants and those excluded from treatment are randomly selected prior to the intervention. If large samples are randomly assigned to treatment and control groups, observable and unobservable characteristics of the two groups should not differ on average and so any difference in outcomes can be attributed to program participation. Hence, the main appeal lies in the simplicity of interpreting results - the program impact is computed as the simple difference between the means of the samples of program participants and control group members on the outcome of interest. The main pitfalls are failure to assign randomly (e.g., because of nepotism, or excluding high-risk groups to achieve better results), changed behavior upon learning of assignment to either group (e.g., enrolling in private programs, intensifying job search), high costs due to large sample sizes, and ethical questions about excluding some people randomly from the intervention.

While randomization is thought to ensure the absence of selection bias among participants, proponents of randomized experimentation make an important assumption: that randomization does not alter the program behavior being studied. This may not be the case, and in fact, the bias induced by randomization may be quite strong. For example, individuals who might have enrolled in a nonrandomized regime may make plans anticipating enrollment in training. With randomization they may alter their decision to apply or undertake activities complementary to training. Thus risk-averse persons will tend to be eliminated from the program.

2. QUASI-EXPERIMENTAL TECHNIQUES

In these techniques, the treatment and control groups are selected *after* the intervention. To get the effect of the program, econometric techniques are used to correct for the differences in characteristics between the two groups. The main appeal lies in relatively low costs, and that they can be done at any time. The main drawback is that these techniques - if done properly - are statistically complex. The reasons for complexities arising are that attributes of individuals in treatment and control groups are different: techniques for adjusting for differences in observable attributes (e.g., sex, education, age, region) are relatively straightforward but subject to specification errors; correcting for unobservable characteristics (e.g., motivation, family connections) requires a convoluted procedure that can yield wildly different results depending upon specification.

(a) Regression-adjusted for observables

This technique is used to assess the impact of participation in a program when the observable characteristics (e.g. age, education, gender) of the participant and comparison groups are different. This method is appropriate for computing program impact estimates when the difference between the participant and comparison samples can be explained by observable characteristics.

(b) Regression-adjusted for observed and unobservable variables (selectivity-corrected)

When selection into programs is not random, and participation in a program is due to both observable and unobservable characteristics, program impacts computed using the technique in (i) above are likely to be biased. The concern is that even if participants and non-participants have similar observable characteristics, there are some unobservable characteristics (e.g. innate ability) which would cause non-participants to have different responses to the program if they had participated. This technique uses the Heckman selectivity method to try and control for these unobservables.

(c) Matched pairs.

As observable characteristics of the individuals chosen in the control and treatment groups are bound to be different, these groups are likely to have different success rates in finding employment even in the absence of active labor market programs. To control for these spurious differences, synthetic control groups are constructed using a matched pairs approach. The synthetic control group, which is a subset of the entire control group, is composed of individuals whose observable characteristics most closely match those of the treatment group. [See Annex for an explanation of the matched pairs methodology].

3. RELATIVE STRENGTHS AND RANKINGS

Estimating the effect of trainee earnings on an employment program using randomized and quasi-experimental techniques, Lalonde (1986) has shown that randomized experimentation yields significantly different results than quasi-experimental techniques. Policymakers should be aware that available non-experimental evaluations of training programs may contain large biases. However, while randomized experimentation is theoretically the best technique to estimate the effects of interventions, quasi-experimental techniques may be superior in practice.

The main weakness of randomized experiments is that among those who participate in the evaluation, it is difficult - if not impossible - to ensure that individuals in the control group do not alter their behavior in a way that contaminates the experiment. For example, people who were denied participation in public training may enroll in private programs, which would bias the results of any evaluation of public programs.[1] The second weakness of randomized experiments is that it may be difficult to assure that assignment is truly random. For example, applicants may be selected into the

[1]Heckman (1992) documents other limitations of this technique when applied to social experiments that arise due to selectivity biases, as randomization may differentially affect the decision of people to participate in such programs. For example, relatively risk-averse persons may decide not to enroll for randomized experiments, preferring instead to enroll in private training.

program due to nepotism, or administrators of the program may purposefully exclude high-risk applicants to achieve seemingly better results. The third problem concerns ethical questions about treating humans as experiment subjects. Finally, experimental evaluations are only possible for future programs, since the control and treatment groups have to be selected *ex ante*.

Using the dual criteria of rigor and feasibility, then, it is clear that randomized experiments are not necessarily superior to quasi-experimental techniques. Factoring-in the high costs of setting up such experiments, and given that the labor market programs which need to be evaluated are already in place in almost all ECA countries, randomized evaluations should perhaps be the last alternative.

Within quasi-experimental techniques, selectivity-correction may not add much, especially when information is available for a considerable number of observable individual and labor market characteristics (such as education, age, sex, household wealth, and region of residence). Besides being cumbersome and somewhat unintuitive, this method often gives arbitrary results depending on the selectivity-correction specification used.

This leaves the matched pairs and the regression-adjusted techniques. Among the two the matched pairs technique is preferred for the following reasons: First, the procedure is less arbitrary because, since the observable differences between the treatment and comparison group are minimized, functional form assumptions become less important. Second, because program impact measures are a simple difference of means of the variables of interest (re-employment probabilities and wages) between the control and treatment groups, they are easier to interpret by non-statisticians.

A weakness shared by both the experimental as well as non-experimental evaluations is that they do not take into account displacement effects of the retraining program. Thus, for example, in countries where demand for labor is constrained, retrainees may simply "bump" or displace previously employed workers so that aggregate unemployment may not change despite the size of the program. In general, displacement implies that the social benefits - from higher re-employment probabilities due to the retraining program - are lower than indicated by the evaluation, however well done.[2]

III. EFFECTIVENESS

1. BRIEF DESCRIPTION OF RESULTS

The main outcomes of interest are rate of re-employment, costs of re-employment, earnings upon re-employment, relevance of assistance to re-employment, reduction of dependence on state-support. The policy decisions based on these feedback mechanisms are to cancel, continue or modify the program.

---

[2]The use of aggregate employment data to determine the overall effect of a training program on outflows from unemployment may also be indicative of the effectiveness of the program. The program can be considered effective if there is a statistically significant positive correlation between the measure and outflows from unemployment. If a matching function is built properly with hiring of the unemployed into regular jobs regressed on the stock of vacancies and unemployment, and the stock of program participants, the methodology can bring about credible evaluation results. If the program is sufficiently large, this measure may be superior to other evaluation techniques when displacement effects are important.

The very existence of these mechanisms seems to create an ethos of cost awareness, and through that, of cost-effectiveness. This finding - if found to be general - strengthens the case for building in-country evaluation mechanisms.

Using a sample that is large by statistical standards (more than 850 observations), retraining - and training, since 40 percent of those trained had not worked before - at first sight, seems ***very successful*** in raising the probability of finding employment, raising it by almost 20 percent (Table 1). When controlling for observed worker and labor market characteristics through regression techniques, these programs appear to be ***only marginally successful*** in facilitating return to work : the gain drops to 6 percent. Using the matched pairs method, retraining is ***unsuccessful*** in raising employment probabilities as the gain is statistically insignificant. Over the period studied, about the same percentage of people with similar characteristics (matched pairs) were getting jobs regardless of participation in training. When controlling for both observed and unobservable characteristics (through econometric techniques), retraining can be either highly successful or ineffective, depending upon the selectivity-correction specification chosen. However, the effect on earnings is unambiguous - all the techniques show that retraining is ***not at all successful*** in raising earnings.

SUMMARY OF IMPACT ESTIMATES FOR PARTICIPATION IN RETRAINING, HUNGARY 1992

Estimation Methodology	**Re-employment** Probability (%)	**Earnings Gain** ($/month)
Unadjusted for Characteristics	19.2**	14.87
Regression-adjusted: Observable Characteristics	6.3*	4.93
Matched Pairs	1.2	20.52

Notes: * significant at 10 percent level, ** significant at 5 percent level. Earnings gain for all methods, and re-employment gains for matched pairs methodology are not significant, and should be treated as 0. Re-employment probabilities are computed during the year after the training program.

Selection bias corrected impact estimates are unreliable as they are "extremely sensitive to empirical specification" (O'Leary, 1995, page 18), and are not reported here.

Source: O'Leary (1995).

Table 2

It should be noted that these results are for 1992, and these programs have been substantially redesigned since. Preliminary results for later years indicate that retraining may be associated with modest but statistically significant wage gains.

2. EXPLANATION OF RESULTS

Why do the estimates for retraining impact on re-employment vary wildly, depending on methodology? That is, evaluations using different quasi-experimental techniques - matched pairs, regression-adjusted, and selectivity-corrected - yield different results for the impact of training on reemployment probabilities. These results are probably because the more ambiguous the impact of the intervention, the more the variance in results yielded by different techniques. As an illustration in contrast, evaluations of public service employment in Hungary - which is clearly associated with lower earnings and lower probability of re-employment in a normal job (O'Leary, 1995) - using all four techniques (unadjusted, regression-adjusted, matched pairs, and selectivity corrected) show uniformly adverse results.

This finding underscores the need to agree upon a reliable, easily interpretable and feasible technique to evaluate the efficiency of labor programs.

O'Leary (1995) and Allison (1994) suggest that to improve effectiveness, retraining programs in Hungary should target older, less educated males who work in non-manual occupations, because the results show that training is a *substitute* for attributes that are associated with higher re-employment probabilities: being younger, more educated, working in manual occupations, and being from certain regions. These results point to the importance of netting out the effects of such observable attributes in determining program effects, especially when equity considerations are important (e.g., encouraging the targeting of workers in worst-affected regions or most disadvantaged groups). Otherwise, programs that target people with attributes correlated with low *absolute* re-employment probabilities - for whom it turns out training is most effective in *increasing* re-employment probabilities - will be penalized relative to those that target people who are well-equipped in finding employment.

Preliminary analysis by O'Leary (1995) indicates that participation in retraining increased the probability of re-employment by:

(i) 26 percent for individuals above 40 years, 14 percent for those between 26 and 40 years, but not at all for job-seekers aged less than 25 years,
(ii) 22 percent for those with less than 8 years of schooling, but not at all for retrainees with more education,
(iii) 20 percent for those without specialized skills in non-manual occupations, and only 8 percent for those with such skills, and
(iv) 16 percent for retrainees from the worst-off region, but not at all for participants from better-off regions.

It should be stressed that this analysis is not conclusive. Lehmann (1995), for example, suggests that these results may be due to program administrators selecting the most motivated job-seekers.

In any case, this finding underscores the need to agree upon an easily interpretable and feasible technique to evaluate the equity-enhancing effects of labor programs, so that labor programs that target the most disadvantaged are not unfairly penalized relative to those that target those who are most likely to find jobs anyway, in order to make the program "look good.".

IV. COSTS

Stopping at rigorous evaluation of what labor programs achieve is not enough: cost data must also be examined to assist governments allocate resources sensibly.

Table 3 presents data on number of individuals who completed retraining, the proportion of those who found employment, and data on retraining costs for 19 of the 20 counties in Hungary. In 1994, the cost of retraining was $900 per person and less than 50 percent of individuals were successful in finding employment - thus the cost of retraining per successful labor force re-entrant was about $1920. In the three counties where the retraining programs have been evaluated (Borsod, Hajdu and Somogy), retraining costs were slightly lower than the national average $830 but, due to the lower reemployment rate of 37.6 percent, the cost per successful trainee - $2345 - was higher than the national average.

COSTS AND COMPLETION INDICATORS, HUNGARY 1994

County	**Number of trainees**	**Completed training** (percent)	**Placement ratio** (percent)	**Cost of Retraining** ('000 forints)	**Cost per placement** ('000 forints)
BUDAPEST	2384	99	51	128	249
BARANYA	379	96	39	65	149
BACS-KISKUN	234	82	57	60	127
BEKES	527	95	49	31	83
BORSOD*	**257**	**94**	**39**	**53**	**145**
CSONGRAD	147	96	38	64	176
FEJER	110	95	35	86	262
GYOR	25	100	71	55	78
HADJU*	**700**	**92**	**36**	**98**	**296**
HEVES	408	93	36	82	254
KOMARON	336	94	46	68	156
NOGRAD	130	93	34	71	226
PEST	421	92	40	66	179
SOMOGY*	**360**	**89**	**40**	**64**	**179**
SZABOLCS	257	89	49	57	131
TOLNA	212	97	48	40	87
VAS	251	62	51	45	143
VESZPREM	341	93	42	63	163
ZALA	481	93	71	48	73
Total/Average	**7960**	**93**	**47**	**90**	**192**

Notes: * indicates counties included in effectiveness evaluations in O'Leary (1995). In 1994, $1 = 100 forints.

Source: Pulay (1995).

Table 3

V. COST-BENEFIT ANALYSIS

In this section, we compare the higher re-employment probability that retrained workers have - relative to those who only receive unemployment benefits (UB) - with the higher costs of providing retraining rather than unemployment benefits. In Hungary, where employment offices offer higher unemployment benefits to the unemployed to participate in approved training programs, the costs of public retraining programs therefore involve both the direct costs of administering the program and the increase in unemployment benefits for trainees (Allison, 1994).

1. METHODOLOGY

The methodology we use can be summarized as follows:

(i) Assuming that retraining lasts an average of 12 months, and that unemployment benefits are paid for either 6 or 12 months, we compute the total financial costs of retraining and UB.

(ii) In the case where retraining programs are longer than eligibility for UB, we add the *expected* foregone earnings by retraining program participants to the cost of retraining; hence the difference between retraining and UB costs are higher when UB duration is shorter than retraining programs.

(iii) Assuming a discount rate of 5 percent, we calculate the expected benefits due to the higher chances of re-employment (and higher earnings) of retrained workers as compared to those who receive only unemployment benefits.

(iv) Using different measures of the impact of retraining reported in O'Leary (1995), we calculate benefit-to-cost ratios and years needed to recoup additional economic costs, i.e., the point in time where accumulated benefits equal costs.

Box 1 illustrates this methodology for a hypothetical active labor program and unemployment benefits scheme. The box shows that the critical assumption is how persistent the effects of retraining are on re-employment probabilities and earnings. That is, if retraining results in a higher probability of having a job in every subsequent period, the results are quite different from the case where retraining may simply help in finding a job quicker in the period(s) after program completion, after which the re-employment probabilities are the same for trainees and the control group. Retraining may also help in raising wages; the more permanent this wage differential, the more cost-beneficial is the active labor program.

BENEFIT-COST ANALYSIS OF LABOR MARKET INTERVENTIONS: A HYPOTHETICAL EXAMPLE

The following hypothetical example illustrates the type of information required to allow analysis of costs and benefits associated with administering an "active" labor program, e.g., retraining, instead of a "passive" program that only pays unemployment benefits (UB) for the same period. We assume that the retraining program and eligibility for unemployment benefits both last one year. The annual earnings gain due to the training received is about 5 percent, and remains constant over the subsequent working life of 30 years.

Under the first scenario, 90% of UB recipients and all the retraining beneficiaries find jobs within one year of searching, and this probability does not change over time. At a real discount rate of five percent over the 30 years, the benefits of the active program exceed the additional costs by a factor of almost five. Under the second scenario, 90% of the UB recipients and all retrained workers find jobs within the first year, but after the first year the remaining 10% of UB recipients also find employment implying that after the first year all UB recipients and retrainees are employed. Under this scenario the benefits of retraining exceed the additional costs by a factor of less than two. The critical difference between the two scenarios is that while retraining results in a 10% higher probability of having a job *for the remainder of working life* under scenario A, in scenario B it results in a 10% greater probability of finding a job *only in the first year after completing retraining*; in both scenarios, the earnings of retrained workers exceed those of UB recipients by 5% in every year of their working lives.

Costs, Earnings and Employment Indicators (Assumptions)	**Unemploy. Insurance**	**Active Labor Program**	**Difference**
1. Costs:			
Total financial costs	$1000	$2000	$1000
Duration of program	12 months	12 months	0 months
2. Annual earnings after program completion	$2000	$2100	$100
3. Scenario A:			
Probability of finding and holding a job			
- in the first year	0.9	1.0	
- in every subsequent year	0.9	1.0	
Relative benefit to cost ratio*			4.6
Time needed to break-even*			4 years
4. Scenario B:			
Probability of finding and holding a job			
- in the first year	0.9	1.0	
- in every subsequent year	1.0	1.0	
Relative benefit to cost ratio*			1.7
Time needed to break-even*			11 years

* Assuming a discount rate of 5% and that earnings differences remains constant over the next 30 years; "break-even" year is when the discounted present value of accumulated benefits equals economic costs (program costs plus foregone earnings).

Box 1

2. RESULTS FOR HUNGARY

Using evaluation results in O'Leary (1995), cost data for retraining in Pulay (1995), and information on unemployment benefits in Burda (1995), we apply this methodology to Hungary. Besides the effectiveness results and the financial costs of training, the benefit-to-cost ratios vary according to the assumptions regarding the duration of unemployment eligibility (which affects the economic costs), and the persistence of training impact over time (which affects the present value of cumulative benefits).[3] Box 2 summarizes the main results.

(a) Matched pairs estimates

Training will never pay off, regardless of whether the average duration of unemployment benefits is 6 months, 12 months, or longer. Both increases in re-employment probability and earnings gains from retraining are insignificant.

(b) Regression-adjusted estimates

Scenario A: Persisting differences. While the re-employment probability gain from retraining is significant (6.3 percent), the earnings gain is not. If the average UB duration is six months, regression-adjusted estimates yield a benefit-to-cost ratio of 0.97; that is, it takes longer than 30 years for retraining to pay off. However, if the average duration for which unemployment benefits are paid is increased to a year, retraining pays off within 10 years.

Scenario B: Transient differences. Training will never pay off, regardless of the duration of unemployment benefits. Over 30 years, about six percent of additional costs of retraining are recouped if unemployment benefits are to be paid for an average of six months, and 13 percent if they are paid for a year on average.

Since these results are based on a snapshot evaluation of the retraining program, it is difficult to determine how long the effects of retraining last, i.e., for how many subsequent years retrained workers have a edge over non-participants in terms of re-employment probabilities (or job stability) and earnings. The true results are likely to lie somewhere between scenarios A and B.

[3]The cumulative benefits can be calculated as the discounted present value of this stream over 30 years (the expected working life since the average age of training and control group participants is about 25 years).

HUNGARY: SIMULATED COST-BENEFIT ANALYSIS OF RETRAINING VIS-A-VIS UNEMPLOYMENT BENEFITS

Evaluation Technique	Persisting differences in reemployment probability*		Current period differences in reemploy. probability**	
	Average UB duration of:		Average UB duration of:	
	6 months	12 months	6 months	12 months
Differences in economic costs ($)	1964	900	1964	900
- Difference in financial costs	1525	900	1525	900
- Differences in foregone earnings	439	0	439	0
Matched pairs				
- Benefit-to-cost ratio	0	0	0	0
- Time taken to recover costs (years)	Never	Never	Never	Never
Regression-adjusted				
- Benefit-to-cost ratio	0.97	2.11	0.06	0.13
- Time taken to recover costs (years)	33 years	10 years	Never	Never

Notes: * equivalent to scenario A in Box 1; ** equivalent to scenario B in Box 1.
Unemployment benefits (UB) are assumed to be $1250 annually; this is based upon a minimum UB of $1030 in 1993 (Burda, 1995).
Retraining program costs are assumed to be $900 per participant, and the duration is assumed to be one year (Pulay, 1995).
Differences are based on monthly earnings and reemployment rates of retraining participants relative to those who only received UB.
Economic costs include financial costs (fixed at $900 for retraining, and at $104 per months for unemployment benefits) and foregone earnings ($0 if the program length is the same as UB benefits, negative if UB benefits are paid longer than retraining program length).
Foregone earnings for retrainees are computed as follows: The average earnings of UB recipients ($148 per month) are multiplied by the number of months that retraining duration exceeds UB duration and the probability of being employed during these months (0.494)

Box 2

VI. IMPLICATIONS

The main findings of this note are: first, when costs are a consideration, it is difficult to justify investments in retraining in Hungary even when unemployment benefits are the alternative: benefit-to-cost ratios rarely exceed one under reasonable assumptions regarding the duration of program effects, and amount of unemployment benefits. Second, even when costs are not a consideration, it is difficult to justify these investments in Hungary since the effectiveness of retraining programs - designed and targeted as they were in 1992 - is zero, as evidenced by the results of matched pairs analysis. Third, more promisingly, there are indications that retraining may be more effective in helping narrowly targeted groups such as older job-seekers, people from backward regions, and the less educated.

Given the first two findings, the question whether retraining programs should be given a high priority is pertinent, especially in countries where such expenditures divert attention and resources from critical areas such as education, health, and infrastructure. Additionally, it can be argued that if the problem is caused by falling labor demand, supply-side interventions such as retraining will not result in either sustainable employment creation or higher wages. Demand-side interventions such as public service employment and wage subsidies will also only lead to temporary (perhaps a year or two for any worker) relief, but may have other labor-related benefits such as preventing decay of work habits during the long unemployment spells that would otherwise result. In Hungary, though, it seems difficult to justify large-scale retraining programs on economic considerations alone.

However, since transition countries are undergoing upheavals that entail unprecedented levels of unemployment, governments may be justified in using active labor programs to counter the adverse psychological aspects of job loss. In these difficult circumstances, many countries may opt for programs to help maintain political and social stability, and should be helped in managing the whole package of programs better. This involves guidance on both effectiveness and cost-effectiveness considerations, both of overall program performance and for assisting specific target groups. Using these programs because governments must be "seen to be doing something" may not even buy time if these programs are not effective. Additionally, there is the matter of *relative* cost-effectiveness of labor programs, information on which can be used to select from the menu of programs available when resources are scarce. On these grounds alone, transition countries should be persuaded of the value of rigorous evaluations in deciding which programs are best used and where.

Building in-country monitoring systems and evaluation capacity may also - by promoting ownership and an ethos of cost-consciousness - strengthen links between evaluation results and policy decisions.

This leaves the matter of what constitutes a sufficiently rigorous evaluation. The effectiveness-evaluation issues that need to be resolved are:

(a) whether we should advocate experimental or quasi-experimental techniques, and

(b) if we decide that quasi-experimental methods are better suited - because of reasons listed above - then we need to choose between matched pairs, regression-adjusted, and selectivity-corrected techniques.

Matched pairs and regression-adjusted estimates are sufficient evaluation devices. Randomized

experiments are not likely to add much information. The ranking - from most to least preferred - using the dual criteria of rigor and feasibility is: matched pairs, regression-adjusted controlling only for observables, general indicators plus labor force survey, and randomized assignment experiments. Controlling for unobservables when many observed characteristics have been controlled-for is both difficult and can lead to erratic estimates depending on the specification of the selectivity equation.

Evaluations, when done well, can be a reliable guide for gauging program effectiveness. But they need to be supplemented by cost information to be of use to policymakers. But these evaluations will not inform us about who should bear the costs of the programs, and about the displacement effects of government-sponsored labor programs. That is, the information provided by evaluations, even when augmented by cost data, is not sufficient for a rigorous *social* cost-benefit analysis. That would require quantification of the net employment and earnings impact, and of social benefits (e.g., political stability, reduced crime) of these programs, which these evaluation techniques are incapable of doing.

VII. SELECTED REFERENCES

Allison, Christine. 1994. "Hungary Human Resources Project: Supervision Report on Unemployment, Training and Re-employment." December 13, 1994, World Bank.

Burda, Michael. 1995. "Market Institutions and the Economic Transformation of Central and Eastern Europe", in S. Commander and F. Corricelli (eds.), *Unemployment, Restructuring, and the Labor Market in Eastern Europe and Russia*, World Bank.

Heckman, James. 1992. "Randomization and Social Policy Evaluation", in C. Manski and I. Garfinkel, (eds.), *Evaluating Welfare and Training Programs*, Harvard University Press.

Lalonde, Robert. 1986. "Evaluating the Econometric Evaluations of Training Programs with Experimental Data", *American Economic Review*, Vol. 76, No. 4.

Lehmann, Hartmut. 1995. "Active Labor Market Policies in the OECD and in Selected Transition Economies". Policy Research Working Paper no. 1502, World Bank.

O'Leary, Christopher. 1995. "An Impact Analysis of Labor Market Programs in Hungary". Upjohn Institute for Employment Research.

Pulay, Gyula. 1995. "Measuring the Efficiency of Employment Policy in Hungary". Paper prepared for seminar on Evaluating Active Labor Programs, April 20, 1995, World Bank.

World Bank. 1995. *Hungary Country Economic Memorandum*. June, 1995.

ANNEX

MATCHED PAIRS METHODOLOGY

The aim of this method is to create a "synthetic" control group that closely matches the characteristics of the treatment group. These characteristics should be chosen using the following criteria:

(a) They should be ***observable***, e.g., education, sex, age, and region of residence, and not innate ability or family background that may be difficult to measure or categorize.

(b) They should be ***relevant***, i.e., these attributes should be correlated with the outcomes of interest, which in this case are the probability of finding a job and the earnings in the job found.

A member of the treatment group is chosen. The control group is searched until a member who is closest to this person in all the observable and relevant attributes is found. These persons then form a "matched pair". This procedure is repeated for all members of the treatment group. In practice, a computer program that uses a "distance measure" makes this task quite painless. When this task is completed, each member of the treatment group has a "identical twin" in the control group who is "identical" in his/her observable and relevant attributes, other than the fact that he/she does not receive training. Two or more persons in the treatment group may end up having the same "identical twin" in the control group; this replication of observations can be limited by including a kick-out clause in the computer program.

The effort in setting up this matching pays off in easily interpretable program impact measures. The impact of the retraining program (or any other intervention) is computed simply as the difference in outcome sample means for the treatment group and synthetic control sub-group.

For example, in the case of Hungary, O'Leary (1995) uses ten observable characteristics (e.g., age, sex, education, occupation, and region of residence) to form the matched pairs sample, and then simply computes the difference between sample means of the re-employment probability and post-training earnings to evaluate the impacts of the program. It should be noted that the differences in sample means of the observable characteristics - other than program outcomes - should be insignificant.