EVALUATING THE IMPACT OF ACTIVE LABOR PROGRAMS: RESULTS OF CROSS COUNTRY STUDIES IN EUROPE AND CENTRAL ASIA

David H. Fretwell
Senior Employment and Training Specialist
World Bank

Jacob Benus
Manager, International Safety Net Group
Abt Associates

Christopher J. O'Leary
Senior Economist
W.E. Upjohn Institute for Employment Research

Social Protection Discussion Paper No. 9915
June 1999

EVALUATING THE IMPACT OF ACTIVE LABOR PROGRAMS:
RESULTS OF CROSS COUNTRY STUDIES
IN EUROPE AND CENTRAL ASIA

Table of Contents

Page

List of Tables

Summary Abstract

Active labor programs (ALPs), commonly found in Organization for Economic Corporation and Development (OECD) countries, are being implemented in transition and middle income economies as one ingredient of labor policy intended to assist in the redeployment of workers negatively impacted by economic restructuring. Active labor polices have both social/political and economic agendas. They are a signal from government to citizens that it cares about individuals who lose their jobs as a result of economic restructuring and wants to help them reenter the labor market. They are also intended to increase productivity, reduce the demand for public income support, and reduce poverty.

The impact of active labor programs has been the subject of some debate. It is argued that the most cost-effective way to generate employment is to create a positive investment climate and reduce government intervention. It is also argued that if the unemployed do not have the appropriate human capital, or have other characteristics that impede them from competing in the labor market, the government needs to intervene. ALPs are intended to help overcome market failures by assisting those most negatively affected by economic change so they do not join the ranks of the long-term unemployed with associated ramifications for themselves and the economy as a whole.

The Study addresses the economic agenda of ALPs and was designed to answer the question: do active labor programs have a significant positive impact on employment and earnings, and if so for whom? This question is of considerable interest to middle income countries, which have limited resources to allocate to competing development priorities, as well as bilateral and multilateral development agencies. In addition, while there are findings from OECD research indicating that well targeted and designed ALPs may have positive impacts, there is little information available to indicate that such findings can be applied to middle income and transition economies.

The Study was coordinated by the World Bank, Europe and Central Asia Human Development Sector Unit (ECSHD) and implemented by National Employment Agencies in four countries (the Czech Republic, Poland, Hungary, and Turkey). It used quasi-experimental design techniques to examine net employment and earnings impact of five ALPs (e.g. training, public service employment, wage subsidy, self-employment, and general employment services). The findings are similar to, but have some deviations from, related OECD and International labor Organization (ILO) research. The Study indicates that ALPs can have a significant positive impact on post-program employment and earnings for selected target groups. Poorly designed or incorrectly targeted programs, however, may have no impact and, in some cases a negative impact; they may also be costly, ineffective, and inefficient. The results of the Study show that the impact of ALPs varies within and between countries. The Study demonstrates that middle income countries can successfully implement ongoing performance monitoring and quasi-experimental design evaluation programs with some initial outside assistance. Development of evaluation capabilities should be incorporated into Bank lending involving active labor programs. However, notwithstanding the findings of the Study, it must be emphasized that a good investment climate, and not active labor program activity, is the primary engine for job creation.

Acknowledgments

This Study was initiated by the World Bank Europe and Central Asia Human Development Sector Unit in partnership with the Czech Republic, Hungary, Poland, and Turkey.

The authors are grateful for the initiative and professional cooperation from the management and staff of: the Ministry of Labor and Social Affairs of the Czech Republic, the Hungarian Ministry of Labor and National Labor Center, the Ministry of Labor and National Labor Office in Poland, the Turkish Employment Organization, Abt Associates, W.E. Upjohn Institute for Employment Research. The authors are grateful for the significant financial contribution from the U.S. Department of Labor/U.S. Agency for International Development, the European Training Foundation, the World Bank, and the countries involved.

The authors would also like to thank the technical coordinators including Jiri Berkovska (the Czech Republic), Gyorgy Lazar (Hungary), Piotr Kolodzieczyk (Poland), and Recep Varcin (Turkey) for their support, as well as Ralph Harbison (former Division Chief, Human Resources Operations Division, Central and Southern Europe Departments, World Bank) and Michal Rutkowski (Social Protection Sector Leader, Human Development Sector Unit, Europe and Central Asia Region) for their initial and continued support without which the Study would not have been completed.

EVALUATING THE IMPACT OF ACTIVE LABOR PROGRAMS: RESULTS OF CROSS COUNTRY STUDIES IN EUROPE AND CENTRAL ASIA

I. INTRODUCTION

1.1 Active labor programs (ALPs) are common in OECD countries, and are increasingly found in transition economies and other middle income countries that are undergoing industrial restructuring and experiencing high levels of unemployment. Government operated ALPs typically include: job counseling and referral services, public works or community employment, wage subsidies, small business creation programs, and retraining. The objective of these programs is to expedite the redeployment of labor, thereby reducing the duration of frictional and structural unemployment, increasing productivity, and reducing expenditures on income support programs. There are political, social, and economic reasons for government action to quicken labor redeployment and cushion the impact of layoffs. Nevertheless, to avoid costly mistakes, investments in these labor programs need to be evaluated as soon as possible, as middle income and transition countries often face severe resource constraints and there are competing demands for government investment in other sectors.

1.2 The administration, financing, and delivery of ALP services vary somewhat by country. In general, the National Employment Service, normally linked to a Ministry of Labor, administers the programs through its provincial and local labor offices. Financing may be from State Budget resources (Turkey, Australia, United States), payroll taxes (Czech Republic and Hungary), or a combination of both (Poland). In many middle income countries financing is from payroll taxes that are used to create an Employment Fund that then finances ALPs and unemployment benefits. Though the aim of such forward financing is to provide reliable sources of money for programs, in practice, during times of high unemployment, the majority of funds are often allocated to unemployment benefits and little remains for ALPs. To offset this problem some countries are establishing minimum budget "set-asides" for ALPs, and providing base funding for operation of general employment services from the state budget. Delivery of ALP services, except for general employment services, is generally accomplished by local

service providers through contracts with local labor offices. Finally, development of comprehensive ALPs usually parallels implementation of formal unemployment benefit systems in order to stimulate job search and ensure that those receiving temporary income support are provided with services to help them quickly rejoin the labor force.

1.3 The development of ALP evaluation systems is well underway in some countries, and initial results are available. However, there is considerable work yet to be done and the nature, depth, and results of ALPs evaluation vary greatly, even within the OECD. Little has been done to evaluate ALP programs in middle income and transition economies. Evaluations that have been done primarily focus on evaluating the costs and effectiveness of programs on participants, and do not make direct comparisons between participants and non-participants with similar characteristics; nor do they look at broader societal impact.

1.4 Recent OECD reviews (Fay, 1996) of program evaluations on the effectiveness of active labor market policies reveal that ALPs differ widely in their objectives and their impacts, both across countries and within countries over time. Program evaluations attempt to determine the impact of various ALPs, both for the individual and on society at large. Individual impacts are usually measured in terms of post-program earnings and/or employment performance. Societal impacts include an estimation of the "dead-weight factor" displacement and substitution effects, along with some accounting for possible externalities. Recent evaluations suggest some ALPs can help most groups of the unemployed. Many unemployed benefit from early intervention through the provision of counseling and job search assistance. Others benefit through targeted employment subsidies, particularly in the private sector. The picture is more mixed with respect to public training programs, which account for a large share of public spending in ALPs in many countries.

1.5 The International Labor Organization also recently concluded an evaluation of ALPs for the long-term unemployed (Meager, 1998). Results of evaluation studies of measures for dealing with unemployment, such as training, public works, special job placement programs, and job subsidies were reviewed. One of the results emerging from the review, which included some 100 evaluation studies,

was that methodological differences and data deficiencies, as well as different institutional and macro-economic contexts, make it difficult to come to firm conclusions about ALP policy effectiveness for the long-term unemployed. However, the review suggests some general policy conclusions: (a) a consensus is emerging that programs, including training, which are directly linked to the labor market and economy, have a greater chance of being effective; (b) there is a growing emphasis on policies that promote reintegration and prevention of long-term unemployment; (c) there is increasing emphasis on job search assistance, as opposed to more expensive ALP measures, but the long term impact of such initiatives needs to be examined; (d) there is growing evidence that ALP measures need to be carefully targeted; (e) there is increasing recognition and evidence that the scale of schemes (i.e., smaller schemes) is critical to their performance; (f) integration of packages (i.e., counseling and training) increases their effectiveness; and (g) there is a need for more rigorous evaluation of the net impact of ALPs, rather than simple monitoring of gross impacts, as well as a need to examine the impact of the administration, management, and institutional context of the delivery of ALPs, including the role of local service providers, which may be critical in determining outcomes in practice.

1.6 In summary, there is a need for more rigorous, comprehensive and ongoing evaluations to support program management and policy decisions, in OECD as well as middle income countries. Two complementary evaluation approaches are desirable:

- The first involves defining "performance indicators" for each program, based on the objectives of the program (e.g., increased probability of employment, enhanced wages), then measuring the extent to which program participants meet these indicators. Performance indicator evaluation systems can help program managers establish targets, provide them with information regarding the degree to which programs are achieving agreed outcomes, provide comparisons between programs and regions, and improve cost-effectiveness of programs. However, performance indicator evaluations do not provide net impact data as they do not compare participants with similar non-participants.

- The second evaluation method which relies on "comparison group design" procedures, however, does provide net-impact information, and can assist program managers and policy makers to make key decisions about program design and implementation. Comparison group design evaluations accomplish this by comparing the degree to which program participants and non-participants, with similar observable characteristics, achieve program outcomes (e.g., do participants in small business assistance programs have a better success rate than non-participants).

1.7 Performance monitoring is normally an ongoing process, but the more costly comparison group design studies, which provide net impact estimates, are needed to interpret the results of performance monitoring and calibrate the targets for performance monitoring systems. In addition, when policy makers and managers consider whether to expand, curtail, or alter the design of an existing program, there is usually an interest in doing a benefit-cost analysis. Benefit-cost analyses may be considered from several perspectives: the program participant, the local labor office, the local government, the national government, and/or society as a whole. A program may have a net benefit for an individual participant, but may not be cost-effective from the perspective of government. The area for concern in benefit-cost analyses is whether the evaluation methodology captures the full income of program completers and whether the analysis has taken account of any displacement that may occur when a program participant increases his/her reemployment success at the expense of non-participants. If the latter occurs, the overall general gain of the program to society may be less than originally calculated.

1.8 In response to the perceived need to improve the knowledge of the impact of ALPs, a number of middle income and transition economies (e.g., Poland, Hungary, Mexico, Turkey, Korea, Brazil, Chile) are striving to improve the quality of ALPs through the recognition that poorly delivered programs have little chance of success. Several countries are initiating systematic evaluation programs, with support from World Bank investment projects. Evaluation activities address several questions, including: (a) what are the key indicators of success for different labor programs, (b) how can information be collected on these indicators; (c) do the results justify the investments, i.e. do observable benefits exceed program costs; and (d) how can information be collected to improve targeting, effectiveness and quality of program delivery? This report summarizes the design and results of four quasi-experimental

design evaluations of active labor programs in Poland, Hungary, the Czech Republic, and Turkey. The approximately US$1.0 million study was coordinated by the World Bank and implemented by four countries between June 1995 and December 1997. The paper summarizes the results, policy implications for ALP implementation,[1] and design and implementation issues that should be considered, if similar work is repeated.

Table 1.1:
Profile of Countries Participating in the Cross Country Study (1996) [2]

	Czech Republic	Hungary	Poland	Turkey
Population (000)	10,300	10,174	38,639	62,700
Workforce (000)	5,130	4,474	17,643	22,236
Per Capita GDP (US$)	4,740	4,340	3,230	2,838
Unemployment	3%	11.2%	13.6%	6%
Unemployment Benefit	yes	Yes	yes	no

II. STUDY OBJECTIVE, BACKGROUND, DESIGN AND DATA ISSUES

Objective

2.1 The objective of this cross country study was to determine if there was any significant difference between those individuals who participated in active labor programs and similar individuals who did not participate in the programs (the comparison group), with regard to agreed outcome measures of program success (e.g., employment, wage levels) in four countries: the Czech Republic, Poland, Hungary, and Turkey. The study evaluated five different ALPs across the four countries for several categories of program participants, grouped by demographic characteristics and geographic location.

Background

2.2 The project was implemented by the national employment services of the Ministries of Labor in

[1] More detailed "Country Reports" are available from the World Bank, Abt Associates (Czech Republic and Turkey), The Upjohn Institute (Poland and Hungary), and from the respective Ministries of Labor.

[2] Annex II contains more detailed population and labor force data on each country.

each of the four participating countries. Financing was provided by multiple partners including the US Agency for International Development (USAID), the European Union (European Training Foundation), and the World Bank. In addition, the four countries provided significant in-kind and direct financial support. The research in Turkey was financed directly from a World Bank investment project in that country. The World Bank Europe and Central Asia Human Development Sector Unit acted as the lead coordinating agency for the study. Two technical assistance contractors participated to help with the design and implementation: Abt Associates (USA) assisted Turkey and the Czech Republic, and the W.E. Upjohn Institute for Employment Research (USA) assisted Poland and Hungary. Additional details regarding the general design and implementation of the Study are contained in Annex I.

Design

2.3 The study used a quasi-experimental design (matched pairs comparison methodology) using randomly drawn samples from the program participant and non-participant populations (see Annex IV). The use of a formal classical experimental design (i.e., random assignment of control and treatment groups) was not considered as a practical alternative because of the additional costs and time that this approach would require, and the social and ethical questions raised by using random assignment to select participants for programs. The main outcomes examined were the proportion reemployed, reemployment earnings, the duration of unemployment, the duration of unemployment compensation, and secondary effects such as new jobs resulting from self-employment. To estimate program impacts the Study used: (a) simple unadjusted differences between mean outcomes, (b) difference between means using a comparison group formed by matched pairs, and (c) regression adjusted impact estimates. The results of the regression adjusted impact estimates are presented in this paper (see para. 2.14 and Annex IV for details).

2.4 The following active labor programs were included in the study: (a) retraining, (b) public works or temporary community employment, (c) wage subsidy, (d) self-employment initiatives, and (e) general employment services. There was general consensus on the programs to be studied, but there were significant differences in definitions and programs between countries. This made it difficult to make direct

comparisons. While this was an impediment to quantitative cross country comparisons, it did not preclude all qualitative comparisons and did facilitate evaluation of different alternatives for similar programs. Table 2.1 summarizes which programs were studied in each country and the approximate unit cost per participant for each ALP in 1996 funds.

Table 2.1:
Programs by Country and Approximate Unit Costs per Participant Served /1 (US $)

	Czech Republic	Hungary	Poland	Turkey
Employment Services /2	12	25	30	17
Training	265	500	300	200
Public Service Employment	625	1,200	800	N/A
Wage Subsidy	885	950	560	N/A
Self Employment /3	885	1,000	2,830 /3	N/A

/1 Direct program delivery costs in US$ 1996, provided by National Labor Offices
/2 Costs of Employment Services include all administrative costs, including administration of unemployment benefit programs, due to the difficulty of desegregating costs
/3 Figure represents the gross costs of a micro-credit program, net will be reduced since 50% of credits are repaid (with interest) by the recipients

2.5 All samples for the ALP control and treatment groups were drawn from the registrants at local labor offices who were unemployed and/or seeking work (details of sampling and control group section and contained in Annex III). The sites in each country were selected to yield a sample that was representative of the nation as a whole. Sample sizes were set to ensure precision based on considerations of tests for observing effects of a size that would be of interest to policy-makers. That is, the samples were determined to be sufficiently large to reject the null hypothesis of no effect. Furthermore, the sample sizes were also sufficiently large to provide reliable estimates of differential program effects on selected demographic.

Table 2.2:
Sample Sizes and Response Rates By Country

	Number of ALP Participants	ALP Participant Response Rate (in percent)	Comparison Group	Comparison Group Response Rate (in percent)
Czech Republic	2,211	23	2,256	15
Hungary	7,228	81	4,415	76
Poland	7,118	93	7,169	90
Turkey	1,643	43	1,748	37

2.6 The data in Table 2.2 shows the size of the participant and comparison groups and the response rates for each country. As indicated, the sample sizes and response rates varied substantially across the four countries. Specifically, the sample sizes varied between 15,048 in Poland and 3,391 in Turkey. In each country, however, the resulting sample size was sufficient to address the key question in this study: is there a significant difference on key outcomes between those who participated in active labor programs and similar individuals who did not participate in these programs. The smaller sample size in Turkey reflects the fewer number of ALPs operated in that country; in fact, only one ALP was evaluated in Turkey.

2.7 The data in Table 2.2 also reflect quite different response rates to follow-up questionnaires, particularly in the Czech Republic. A unique problem was encountered in this country where, due to privacy legislation, all individuals selected for the study first had to be contacted by the Employment Service to gain their permission to be interviewed by a private survey firm. About 18 percent (4,537) of an overall sample of 24,977 agreed to participate in the study, and of the 4,537 that agreed to participate, 4,467 actually responded to the surveys (98.5 percent). These circumstances in the Czech Republic greatly increased design costs, and lowered overall response rates, as can be seen in Table 2.2. While the low response rate could have generated response bias, any bias was likely similar for both participant and comparison groups because of the uniform pre-screening process and the high response rate in the second stage (see Annex III for a more detailed discussion of sample selection issues and procedures in each country).

2.8 Personal interviews, as opposed to mail surveys, were the primary method of obtaining follow-up information. However, implementation procedures differed across the countries. In Poland and Hungary, local labor office staff completed the interviews and, while this assisted in gaining a high response rate, it may have had some implications for the objectivity of responses. In Turkey and the Czech Republic, the approach was to use a third party contractor. While this helped guarantee objectivity, it considerably increased the cost per observation. Some telephone interviews were also conducted in Turkey to reduce costs. The impact of the differing response rates should be considered when making comparing cross-country impact estimates. However, if the data are similar in high and low response countries, it may help to validate evidence from the low response rate countries. The direct local cost per participant follow-up was US$10 in the Czech Republic, US$4 in Hungary, US$4 in Poland, and US$13 in Turkey (including interviews, supervision of interviewers, but excluding data processing and related international consultant costs for design). Procedures for follow-up interviews and questionnaires were field tested and interviewers were trained prior to full-scale implementation.

Data and Analysis Issues

2.9 *Definition of Programs:* As noted previously, it is difficult to compare ALP outcomes between countries because of differences in definitions and policies between programs. However, the differences in themselves can be useful in undertaking impact analyses (i.e., in one country the public service employment program could be operated by private and public sector institutions, and the long term employment impact was quite different depending upon type of program operator). Considerable differences were noted in definitions of self-employment, public works, and wage subsidy programs, and as such, cross country comparisons between these programs must be completed with care.

2.10 *Selection of control groups.* The problem of selecting comparison groups that are similar to the participant groups is difficult because some variables (e.g., motivation) are impossible to measure. There are several techniques for comparison that can be used to ameliorate differences, but their use is affected by the time and funds available. In the case of this study the primary approach was to select comparison groups with similar observable characteristics as the participant groups. Further

adjustments were made in estimating impacts by including these and additional variables in multivariate regression models.

2.11 *Dead-weight*. This is a key issue in evaluating program impacts. Money spent on ALPs that provides services to people who could have been reemployed without the assistance is a "dead-weight" on the program. To account for this cost there is a need to compute net program impact estimates. That is, the effect of the program net of preexisting abilities. This is done by comparing the outcome of ALP participants with the outcome for similar individuals who did not participate. For example, if 60 percent of participants were gainfully employed at the end of a small business assistance program, and 40 percent of a matched pairs group of similar non-participants were employed, the dead-weight is 40 percent and the net impact is 20 percent.

2.12 *Displacement and substitution*. Displacement occurs when ALP participants gain reemployment at the expense of other qualified workers who might have taken the job anyway, so there is no net gain in employment by using ALPs. Substitution occurs when ALP money received by a firm to expand employment, simply reduces spending which otherwise would have been made anyway. Similarly, local governments may simply use ALP resources to displace already budgeted expenditures (a common problem with public service employment programs). Some have argued that these effects are mollified-mitigated by the fact that the faster job matches which result from government spending expand the size of the economy that leads to secondary employment effects. It could also be argued that since a very small proportion of the unemployed participate in ALPs, any such impact is minimal. The development of human capital through ALPs could have a positive long-term impact even if there are short-term displacement effects.

2.13 *Creaming:* This refers to a practice whereby program operators select the best participants, as opposed to those who may benefit the most from the program, to help ensure observed program success. This problem is sometimes encountered when performance monitoring evaluation systems are used. Program managers may attempt to improve program performance by selecting job-ready candidates for participation, as opposed to others who may be less job ready, to improve job

placement rates from programs. Such actions increase dead-weight and decrease the net program impact, and this is one reason why performance monitoring systems need to be supplemented with net impact studies. Careful program targeting can help eliminate creaming.

2.14 *Program policy and quality.* This was difficult to measure with participant follow-up surveys since self-reporting is not a reliable way of measuring program quality. However, policy and quality factors often have a major influence on the impact of ALPs. Simply put, programs that are poorly designed or poorly managed will probably not have significant positive impacts. For example, there is growing evidence that if training agencies do not have a contractual obligation to place an agreed number of participants in jobs, with built-in incentives and disincentives, the quality of training and post-program placement rates may not be high. Furthermore, there is evidence that public service employment programs operated by public agencies have a very low rate of transition to regular non-subsidized employment, but the same programs run by private contractors have much higher rates of reemployment in normal jobs. For these reasons any analysis of the results of program impact must take into account the program policies and quality of program content.

2.15 *Measuring Program Impacts:* The central issue in using non-experimental methods for evaluating program impacts is how to select a comparison group that is most similar to the program participants as possible, but that does not participate in the program. The experiences of the comparison group are then used as a measure of what would have happened to participants in the absence of the program. In this Study, a non-participant sample was drawn to match the participant group as closely as possible on using demographic characteristics. Having constructed a matched comparison group for each of the ALPs, the measure of program impacts is the difference between participant group outcomes and comparison group outcomes. For any given outcome, an unbiased measure of the program impact is provided by a simple difference in participant and comparison group means. This simple difference in outcome means is referred to as the unadjusted program impact. A more precise, and still unbiased, impact estimate can be obtained through multivariate analysis, using covariates to explain some of the variation in outcomes across the sample. By including a variable that captures participant status (i.e., P=1 if the labor office registrant is in the participant group and P=0 if

the registrant is in the comparison group), it is possible to obtain an unbiased estimate of the average impact of the program on the outcome. In addition to the "dummy" variable for participant status, the regression equations include variables reflecting demographics and other characteristics. Impact estimates obtained from such multivariate regression techniques are referred to as regression adjusted program impacts. The results presented in this paper reflect these regression adjusted estimates (see Annex IV for a more detailed discussion of these statistical techniques).

2.16 *Benefit-cost analysis*: This is a difficult issue because of problems in identifying long-term fiscal return from ALPs to individuals and society, and because of possible displacement and substitution effects. This Study provides partial information for such an assessment. This information includes: program participant costs, temporary income support savings, the net impact on reemployment and average monthly earnings. However, other crucial information was not available: downstream wage impacts, and returns to society (tax revenues, productivity gains, long-term income support payments required by non-participants). Other studies, including related work in the U.S.[3], provide information to indicate that the wage and employment impact of some ALPs (e.g., training) may be long-term in nature, however, short-term measurement of outcomes may underestimate the impact. This study does provide some insight into short-term gross rates of returns to individuals, but was unable to look at longer-term societal returns.

III. FINDINGS

3.1 This chapter presents an overview of the main findings from evaluations of the most popular ALPs conducted in the Czech Republic, Hungary, Poland, and Turkey. The effects of five ALPs are considered. The impact estimates are reviewed across countries by program in the following order: training, public service employment, wage subsidy, self-employment and employment services. Not every program was operated in each country, and there were differences across countries operating similar programs.

[3] See for example, Jacob Benus, et al., "Third Annual Assessment Report of the Workforce Development Partnership program," Abt Associates, 1996. Also Jacob Benus, et al., "Self Employment Programs: A new Reemployment Strategy". Unemployment Insurance occasional paper 95-4, U.S. Department of Labor, 1995.

3.2 The reviews of results for each program begin with a summary of the main elements of how the ALP operates in each country. This is followed by an examination of the observable characteristics of the ALP participant samples. Next, a qualitative summary of program impacts is given on the important outcome measures: ever reemployed in a job (initial employment), employed in a job on the survey date (current employment), average monthly earnings at the start of the first new job (initial earnings), average monthly earnings in the current job on the survey date (current earnings), and the amount of unemployment compensation payments (unemployment compensation). Program impacts on employment are in percentage (e.g. +0.10 indicates participants had a 10 percent better chance of gaining employment as compared to non-participants. Program impacts on earnings are in dollars (e.g. +$86 indicates participants obtained eighty six more dollars of monthly income than non-participants).

3.3 The label given in parentheses after each of these outcome measures is the label provided in the table that summarizes the results. Results were judged significant for reporting in this chapter based on formal statistical tests at the 90 percent level of confidence or above. In addition to overall impact estimates on the five outcomes listed, this chapter provides a qualitative review of impacts on the two employment outcomes for important subgroups partitioned by: gender, age, education, and the duration of prior unemployment.

3.4 The summary of results provided here is necessarily brief. A wealth of additional information including additional sub-group analyses and impact estimates measured over different time periods is contained in the individual country reports for the Czech Republic (Benus, 1998a), Hungary (O'Leary, 1998a), Poland (O'Leary, 1998b), and Turkey (1998b). The interested reader is encouraged to obtain these reports and undertake further examination of the findings.

Training

3.5 Retraining programs were evaluated in all four countries (the Czech Republic, Poland, Hungary, and Turkey). The general objectives of these programs were to address problems of structural

unemployment by providing the unemployed with updated and additional skills and knowledge. There were several variants of retraining, including training in private and public institutions, on-the-job training in enterprises, and combinations of institutional and on-the-job training. In addition, training could be either "group" or "individual." That is, it may be the case that a training course was organized by, or for, a labor office, and unemployed job seekers were referred for participation in the group, or alternatively, an individual proposed to a labor office to enter an ongoing course of study in an existing educational or training institution that was financed by the labor office using a "voucher" approach.

3.6 Turkish training programs emphasized on-the-job training, averaging 4.5 months in length, as opposed to institutional training. Contracts with training agencies and enterprises were "performance-based' with pre-negotiated job placement rates and trainees were provided with a token amount for living and travel expenses. The programs in Poland, Hungary, and the Czech Republic tended to be more institutional based. The maximum length of training in Poland was 12 months. Participant stipends were up to 115 percent of the unemployment benefit; participants who left a course before completion had to reimburse the costs of training. Training contracts were not performance-based but were subject to a public procurement process. In Hungary, training was generally less than 12 months, participants were provided a stipend up to 110 percent of the unemployment benefit plus reimbursement of direct costs. Contracts were not performance-based but public procurement procedures were used. In the Czech Republic, training had the largest number of participants of all ALPs. There were two training programs, one for the general unemployed and one for youth. The former averaged two months in length and the participants got 70 percent of their previous wage during training. The latter youth training program, focused on on-the-job training (similar to Turkey); the employer received a lump sum for the training in exchange for retaining the participant for at least one year beyond the end of the training program.

3.7 Table 3.1 summarizes the observable characteristics of trainees. Trainees were slightly more male in Hungary and Poland, but more female in Turkey and the Czech Republic, relatively young in all countries (30 years or less). A considerable portion had completed only primary education (with the exception of the Czech Republic), and another major group had completed vocational secondary school

training (in Hungary, Poland, and the Czech Republic). It should be noted that many individuals completing secondary vocational programs did not actually achieve secondary school matriculation.

Table 3.1:
Characteristics of Participants of Retraining Programs (in percent)

	Czech Republic 1/	**Hungary**	**Poland**	**Turkey**
Gender				
Male	25	56	51	45
Female	67	44	49	55
Average Age	30	28	30	23
Education				
Primary	10	35	26	18
Secondary Voc.	24	49	62	N/A
Secondary	59	24	09	67
Post Secondary	01	06	03	15

1/ Adult retraining program

3.8 The overall impacts of training programs on employment, earnings, and use of unemployment benefits are presented in Table 3.2. The data indicates that, in general, training had a small but positive impact on employment, except for Turkey where the findings were not significant or were negative. It should be noted that the Turkish program was primarily a short-term on-the-job training program that had many characteristics of a wage subsidy program and there is current no-unemployment benefit program in Turkey. The employment impact appears to have been quite durable in Poland and Hungary, tended to dissipate over time for the short Czech training programs, and became negative over time in Turkey. While not shown on Table 3.2, the data also indicate that the current employment impact of individual training in Hungary was more positive than for group training, by about 0.03 percent; The reason for the difference between individual and group training in Hungary is not clear, and may be due to unmeasured characteristics of those entering individual training (e.g. personal initiative). Data from Poland and Hungary also show that shorter training (e.g., 1-6 months) can have similar, and in some cases grants impact, than longer training (e.g., 6-12 months).

Table 3.2:
Overall Impact of Training Programs

	Czech Republic	Hungary /1	Poland	Turkey
Any Employment	+0.11***	+0.17 **	+0.10**	+0.02 /2
Current Employment /3	+0.03 /2	+0.12 **	+0.14**	-0.06 ***
Initial Monthly Earnings	N/A	$10 *	N/A	N/A
Current Monthly Earnings	+$86 ***	$5	+$7**	+$32 ***
Unemployment Comp. /4	+$198 ***	-$27	+$94*	Na

na - Not available
*** Impact statistically significant at the 99 percent level of confidence
** Impact statistically significant at the 95 percent level of confidence
* Impact statistically significant at the 90 percent level of confidence
/1 Group training
/2 Impact on self -employment was positive
/3 At the time of the survey.
/4 Positive means more unemployment benefits were paid to participants.

3.9 The employment impact on sub-groups of participants is presented in Table 3.3. The impact can be seen as positive for both males and females, but was more positive for females; and was more positive for youth and middle aged workers, as opposed to older workers. The impact is also more positive for individuals with primary and secondary education, as opposed to individuals with post-secondary training. Training can be effective for both long and short-term unemployed, but was more positive for the short-term unemployed, except for in the Czech Republic and for females in Turkey. In several instances the impact of training tended to decline over time, which has ramifications for benefit-cost analyses. This was particularly true for the Turkey program, which was primarily a subsidized on-job-training program which appeared to be helping people into initial employment, but the net effect did not last. This type of program may have short term social benefits, but did not appear to produce long term employment benefits.

3.10 Although not shown on the table, the data from Poland shows significant current employment impact from both public and private training providers, but slightly higher impact from private providers (e.g.+0.10 for public, and +0.12 for private). The data also show the impact of training on current employment was slightly higher in areas of high unemployment vs. low unemployment (e.g., in Hungary +0.14 vs. +0.10, Poland +0.09 vs. +0.06). While not reported in Table 3.3, one retraining subgroup in Hungary (those who were forced out of their earlier job or were school leavers), had significantly higher reemployment rates than those who left their jobs voluntarily or were new entrants to the labor force.

Table 3.3:

Employment Impact of Training by Subgroup /1

Subgroup	Czech Republic Any/current	Hungary /2 Any/current	Poland Any/current	Turkey any/current
Gender				
Male	+0.04/+0.01	+0.10**/+0.03	+0.09**/+0.12**	-0.09***/-0.2***
Female	+0.11***/+0.04	+0.11**/+0.09**	+0.06**/+0.12**	-0.11***/-.06**
Age				
Youth	+0.06/-0.04	+0.07**/+0.06**	+0.07**/+0.10**	+0.04/+0.00
Middle age	+0.12***/+0.09***	+0.15**/+0.91**	+0.20**/+0.29**	+0.02/-0.06**
Older worker	+0.05/-0.00	+0.09/_0.00	-0.12/-0.08	+0.01/-0.07**
Education				
Primary	+0.10*/-0.04	+0.17**/+0.05	+0.15**/+0.15**	+0.11***/-0.01
Secondary	+0.08***/+0.05	+0.12**/+0.06*/3	+0.08**/+0.12** /3	+0.03/-0.03
Post Secondary	+0.11/-0.00	+0.07/+0.22**	+0.07/+0.10	-0.03/-0.10**
Unemployment				
Short <12 mos	+0.09**/-0.02	+0.11**/+0.08**	+0.14**/+0.17**	-0.06***/-0.14***
Long >12 mos	+0.08***/+0.06**	+0.09**/+0.00	+0.01/+0.06**	+0.07***/-0.01

*** Impact statistically significant at the 99 percent level of confidence
** Impact statistically significant at the 95 percent level of confidence
* Impact statistically significant at the 90 percent level of confidence
/1 Any employment and current employment at time of survey
/2 The results are for group training.
/3 Vocational secondary school

Public Service/Temporary Community Employment

3.11 Public service employment (PSE) programs were evaluated in three countries (Czech Republic, Poland, and Hungary). The general objective of PSEs is to provide temporary income support through short-term transitional employment. PSE projects usually provide support to improve public infrastructure and services, and may assist in maintaining and developing job skills.[4] All three countries have invested considerable resources in these programs. In Poland and Hungary, the programs had the highest number of participants of any ALP. In Poland the program was called "public works" and participant stipends were set at 75 percent of the average national wage which was double the 36 percent of the average wage paid to recipients of unemployment benefits. In Hungary, the Employment Fund paid for 70 percent of direct costs for PSE participants with local governments covering the remaining costs. In the Czech Republic the program was called "Publicly Useful Jobs," participation

[4] For a more complete analysis of the objectives and outcomes of public service employment see David Fretwell, Sandra Wilson, "Public Service Employment - A Review of Programs in Selected OECD Countries and Transition Economies." World Bank Discussion Paper, World Bank, Washington D.C., 1999.

was limited to six months, but could be extended, and projects generally required only low job skills. In all three countries work on a PSE project was considered bona fide employment to requalify for unemployment compensation, which caused some problems with program operation. PSEs are increasingly used as a means of determining if a person is really available for work, and if the person refuses a PSE job offer they may lose their social assistance and/or unemployment benefit.

3.12 A summary of the observable characteristics of PSE participants is presented in Table 3.4. Compared to the general population of registered unemployed PSE participants tended to have a lower level of educational attainment--the majority had only primary school education, were more likely to be middle aged, and were most likely male. The manual nature of work required on many PSE projects may have influenced the gender and educational attainment composition of participants.

Table 3.4:
Characteristics of Public Service Employment Participants (percent)

	Czech Republic	Hungary	Poland
Gender			
Male	61	66	85
Female	37	44	15
Average Age	34	36	29
Education			
Primary	88	67	87
Secondary	10	30	12
Post Secondary	01	03	01

3.13 The impact of PSE programs on employment, earnings, and use of unemployment benefits is presented in Table 3.5. The results indicate mostly negative impacts on earnings and employment. In two countries of the three countries studied, there were significant positive impacts on the amount of unemployment compensation paid (more compensation paid). These results are partially the result of program design where participants may requalify for benefits by participating in a PSE. The results for Poland indicated a significant positive impact (+0.10) for transition to regular non-subsidized employment when private contractors were used. Use of public contractors had a significant negative impact (-0.05). Table 3.4 shows the combined impact of public and private PSE contractors.

Table 3.5:
Overall Impact of Public Service Employment Programs

	Czech Republic	Hungary	Poland

Any Employment	+0.05	-0.07**	-0.05**
Current Employment	-0.10***	-0.06**	0.02
Initial Earnings	N/A	+$4.13**	N/A
Current Earnings	-$35	-$9**	-$6
Unemployment Compensation	+$114 ***	-$9**	+$103 *

N/A - Not available
*** Impact statistically significant at the 99 percent level of confidence
** Impact statistically significant at the 95 percent level of confidence
* Impact Statistically significant at the 90% level of confidence

3.14 The employment impact, following completion of PSE programs, on subgroups of participants is presented in Table 3.6. The employment impact by gender tended to be negative, or insignificant, with a positive indication for females in Hungary. In Poland, where there was a large number of young people involved in public works projects, youth reemployment was negatively impacted by participation in a PSE. Participation in a PSE does not help the long-term unemployed re-enter normal jobs, and only has a positive impact for short-term unemployed in Hungary.

Table 3.6
Employment Impact of Public Service Employment Programs by Subgroup /1

	Czech Republic Any/current	Hungary any/current	Poland any/current
Gender			
Male	0.02/-0.12**	0.01/0.01	-0.07**/0.00
Female	0.08/-0.07	0.12**/0.10**	0.02/0.04
Age			
Youth	0.01/-0.09	0.00/-0.01	-0.07**/0.01
Middle age	0.05/-0.08	0.06/0.04 0.14/0.14	-0.04/0.01
Older worker	0.06/-0.13**		-0.05/0.04
Education			
Primary	0.05/-0.11*	0.02/0.01	-0.01/-0.00
Secondary	0.00/-0.09	0.03/0.03 /2	-0.08**/0.02 /2
Post Secondary	-0.35/-0.29	0.13/0.15	-0.12/-0.22
Unemployment			
Short <12 mos	0.07/-0.12	0.08**/0.05**	0.02/0.02
Long >12 mos	0.04/-0,10**	-0.02/0.03	-0.13**/-0.01

*** Impact statistically significant at the 99 percent level of confidence
** Impact statistically significant at the 95 percent level of confidence
* Impact statistically significant at the 90 percent level of confidence
/1 Any employment and current employment at time of survey
/2 Secondary vocational school

Wage Subsidy

3.15 Wage subsidy programs were evaluated in three countries (Czech Republic, Poland, and Hungary). The general objective of these programs was to facilitate new employment by subsidizing employer wage costs for a limited period of time. These programs attempted to stimulate labor demand, provide on-the-job work experience, and establish employer-employee relationships. In Poland, a program called "intervention works" operated like a wage subsidy program and it is analyzed here as such. In intervention works, wage and social insurance costs for participants were paid for up to six months up to the level of the unemployment benefit, and for 150 percent of the average monthly wage for the subsequent six months. Projects could not compete with private companies, and were available only to enterprises that did not lay off more than 10 percent of workers over the preceding six months. In Hungary, the wage subsidy was 50 percent of wages for up to one year and was targeted at long-term unemployed or young first-time wage earners. This subsidy was only available to employers which had not laid off related workers in the preceding six months and which guaranteed to retain the participants for a period equal to the duration of the wage subsidy. In the Czech Republic the program

was called "Socially Purposeful Jobs" and provided a lump sum payment to employers who hired the unemployed for up to two years. If the job did not last two years or the participant was not retained, the employer had to return the entire lump sum subsidy. The Czech program had two streams, the first promoted wage employment, the second self-employment. Data from the first is presented in the following tables, the second is presented in the next section on "self employment." It is also useful to compare the findings on these programs with the data from Turkey on training programs presented in preceding sections, as the Turkish training programs included many elements of wage subsidy programs.

3.16 A summary of the observable characteristics of wage subsidy participants is presented in Table 3.7. Wage subsidy participants tended to have low levels of educational attainment--the majority having only completed primary school -- and were younger than the average registered unemployed (except for in Hungary where the long term unemployed were the target group). It should be noted that the Czech Republic also has a separate wage subsidy program for recent graduates; combining the data from the two programs would lower the average age. The gender, age, and educational distribution of the Turkey training program were similar to Poland (see Table 3.1).

Table 3.7:
Characteristics of Wage Subsidy Participants (percent)

	Czech Republic	Hungary	Poland
Gender			
Male	35	56	41
Female	65	44	59
Average Age	32	34	23
Education			
Primary	74	53	47
Secondary	23	43	40
Post Secondary	02	04	01

3.17 The impact of wage subsidy programs on employment, earnings, and use of unemployment benefits after program completion is presented in Table 3.8. Findings for retraining in Turkey (Table 3.2) have been included for comparison purposes because the characteristics of the Turkish on-job-training retraining program are very similar to wage subsidy programs. The data show the impact varied somewhat by country. Poland's program had a positive and lasting impact on employment, where

programs of less than or six months, had more impact than longer programs (e.g., 0.18, 0.26, 0.12). No lasting impact was found in longer Czech Republic programs, and a negative impact emerged in Hungary and Turkey. The impact on the use of unemployment benefits was positive in the Czech Republic and Hungary, with participants received more unemployment compensation; but negative in Poland, with participants receiving considerably less compensation. The data from the Czech youth wage subsidy program reflect the same trends as for the adult program presented in Table 3.8.

Table 3.8:
Overall Impact of Wage Subsidy Programs

	Czech Republic	**Hungary**	**Poland**	**Turkey /1**
Any Employment	0.09***	-0.01**	0.23**	+0.02
Current Employment	0.02	-0.03**	0.24**	-0.06 ***
Initial Earnings	N/A	$10	N/A	N/A
Current Earnings	-$18 **	-$6	$1	$32 ***
Unemployment Compensation	$48 **	$7**	-$182**	Na

na - Not available
*** Impact statistically significant at the 99 percent level of confidence
** Impact statistically significant at the 95 percent level of confidence
* Impact Statistically significant at the 90% level of confidence
/1 Turkey data from Table 3.2 for comparison purposes

3.18 The employment impact of wage subsidy programs on subgroups of participants are presented in Table 3.9. Both males and females appear to benefit and the impact appears lasting, except in the Czech Republic, where the program worked better for females. In general, all age groups benefited from the program, except in the Czech Republic, where there was no impact for older workers. The programs worked better for those with primary and secondary level education. The programs appeared to work equally well for those with short and long-term unemployment, except in the Czech Republic where there was a only a short-term benefit for the long term unemployed. Females and lower educated participants gained the most from the programs in Turkey and the Czech Republic program, where a considerable proportion of participants with these characteristics were in the programs.

Table 3.9:
Employment Impact of Wage Subsidy Programs on Subgroups /1

	Czech Republic Any/current	Hungary any/current	Poland Any/current
Gender			
Male	0.06/0.06	0.07**/0.08**	0.11**/0.11**
Female	0.10***/0.02	0.12**/0.19**	0.15**/0.11**
Age			
Youth	0.12***/0.04	0.06**/0.67**	0.13**/0.134**
Middle age	0.16**/-0.01	0.07**/0.09**	0.14**/0.16**
Older worker	na/0.03	0.14**/0.14**	0.15/0.30**
Education			
Primary	0.13**/0.04	0.12**/0,13**	0.15**/0.18**
Secondary	0.07/-0.02	0.08**/0.06** /2	0.14**/0.13** /2
Post Secondary	0.10/0.10	0.02/-0.00	0.05/-0.04
Unemployment			
Short<12 mos	0.06/-0.00	0.09**/0.08**	0.25**/0.21**
Long >12 mos	0.10***/0.02	0.12**/0.12**	-0.09**/0.01**

*** Impact statistically significant at the 99 percent level of confidence
** Impact statistically significant at the 95 percent level of confidence
* Impact statis tically significant at the 90 percent level of confidence
/1 Any employment and current employment at time of survey
/2 Secondary vocational school

Self Employment

3.19 Self-employment programs were evaluated in three countries (Czech Republic, Poland, and Hungary).[5] A small self-employment training program was also implemented in Turkey but, since the program focused on training and had very few participants, the program impacts were not investigated. Self-employment programs are often initiated to address a "lack of demand" for labor, and also address structural unemployment. In Poland, the programs provided micro-credit loans that could not exceed 20 times the aggregate monthly wage, and were offered at prevailing interest rates. If the self-employment continued 24 months in Poland, 50 percent of the original loan amount was forgiven (this program has since been supplemented with a network of small-business technical assistance centers and incubators). In Hungary, the program focused on technical assistance and additional income support including up to six months of supplemental unemployment benefits. The program also financed half the cost of technical assistance services/training, and up to half the premium on loan insurance for business

[5] Results reported here for transition countries may be contrasted with findings from classically designed random assignment field experiments on self-employment conducted in the United States (Benus et al., 1995).

start-up. In the Czech Republic the program was one part of the "Socially Purposeful Jobs" program as described in the previous section on wage subsidies. Participants were given credits, of an average of about US$900, to start a small business.

3.20 A summary of the characteristics of self-employment participants is presented in Table 3.10. A higher percentage of participants tended to be male in Hungary and Poland than in the Czech Republic, were middle aged. Both primary and secondary educated individuals participated equally, along with some individuals who had higher education

Table 3.10:
Characteristics of Participants in Self-employment Programs (in percent)

	Czech Republic	**Hungary**	**Poland**
Gender			
Male	44	62	60
Female	56	38	40
Average Age	35	36	34
Education /1			
Primary	46	51	54
Secondary	45	38	43
Post Secondary	09	11	03

/1 Level of education completed

3.21 The impact of self-employment assistance on employment, earnings, and use of unemployment benefits after program completion is presented in Table 3.11. The impact on employment was positive and lasting in two of the three countries studied. The impacts on earnings was positive in Poland, neutral in the Czech Republic, and negative in Hungary[6]. There may have been reluctance for full disclosure to public officials as part of participants' tax avoidance strategy. The negative impact on use of unemployment compensation, e.g., less compensation was paid to participants, was considerable in Hungary and Poland. This trend was in contrast with the outcome in the Czech Republic, where more benefits were paid, and may has been a function of the design of the program if the credit provided was actually charged as unemployment compensation.

[6] The control group was constructed in the same manner as for other programs, employment included both self and/or wage employment, and earnings included self and/or wage employment earnings.

Table 3.11:
Overall Impact of Self-employment Programs

	Czech Rep.	Hungary	Poland
Initial Employment	0.11***	0.17	0.28**
Current Employment	0.22*** /1	0.19	0.24**
Initial Earnings	N/A	-$40**	N/A
Current Earnings	$5	-$26**	$71**
Unemployment Compensation	$79 *	-120**	-$264**

N/A - Not available
*** Impact statistically significant at the 99 percent level of confidence
** Impact statistically significant at the 95 percent level of confidence
* Impact Statistically significant at the 90% level of confidence
/1 self employment impact was 0.75***

3.22 The impact of self-employment programs on subgroups of participants is presented in Table 3.12. In general, both genders benefited significantly with females faring better in the Czech Republic. Middle and older workers benefited more consistently than youth, although youth benefited in Hungary and in initial employment in Poland. Participants with lower education levels (e.g., primary) benefited equally, and in several cases better, than workers with more education, particularly those with post-secondary education. There is evidence that the programs can have a positive impact for those who are both long and short-term unemployed, with a slightly greater impact with the longer-term unemployed.

Table 3.12:
Employment Impact of Self-employment Programs by Subgroup /1

Subgroup	Czech Republic any/current	Hungary any/current	Poland Any/current
Gender			
Male	0.09/0.10	0.71**/0.75**	0.16**/0.06*
Female	0.14**/0.32***	0.12**/0.16**	0.37**/0.26
Age			
Youth	0.09/-0.05	0.09**/0.07**	0.15**/0.39
Middle age	0.06/0.17**	0/07**/0.09**	0.28**/0.20**
Older worker	0.19***/0.35***	0.14**/0.14**	0.38**/0.17**
Education			
Primary	0.59***/0.72***	0.12**/0.13**	0.33**/0.71**
Secondary	0.00**/0.07	0.08**/0.06* /2	0.24**/0.15** /2
Post-Secondary	0.08/0.10	0.02/0.00	0.27**/-0.04
Unemployment			
Short<12mos	0.11/0.11	0.09**/0.08**	0.23**/0.22**
Long >12mos	0.12**/0.24***	0.12**/0.12**	0.28**/-0.01

*** Impact statistically significant at the 99 percent level of confidence
** Impact statistically significant at the 95 percent level of confidence
* Impact statistically significant at the 90 percent level of confidence
/1 Any employment and current employment at time of survey
/2 Secondary vocational school

Employment Services

3.23 An attempt was made to evaluate Employment Service (ES) programs in all four countries (the Czech Republic, Poland, Hungary, and Turkey). The general objectives of employment services were to: (a) register the unemployed and monitor their employment status to validate continuing eligibility for income support payments; (b) provide placement services to assist those who are frictionally unemployed to re-enter the labor market, and; (c) for those who are structurally unemployed, to provide career assessment and screening prior to entry into another ALP (Fretwell and Goldberg, 1993). Services provided by the ES generally included job placement, career counseling and assessment, in-depth job search assistance and job clubs. Additionally, the ES administered the unemployment compensation and other active labor programs.

3.24 It should be noted that, at the time the study was conducted, employment services were in a state of development in all countries and, while all provided registration and job placement services, other services were not widespread. This is reflected in the data from Hungary and Poland which shows that 81 and 86 percent respectively, of the comparison group who used the ES, used the job referral service with the use of other services being minimal. ALP participant use of public ES offices was, however, quite widespread. This was possibly due to the recent emergence of unemployment and the need to register at the employment service to obtain unemployment benefits, and the historical propensity of individuals in these countries to look for assistance from state institutions.

3.25 Difficulties were encountered in evaluating the ES in this study because there was no clear way to define participant and comparison groups as done for other ALPs. All ALP participant and comparison group members, including those using employment services, were registered with the ES. Those who did not participate in other ALPs (e.g., were in the general comparison group), became the main sample for evaluating the ES. Each person in the control group was asked if they used or did not use services of the ES, and from this information participant and matched comparison groups were derived. The problem with this approach was that, by definition, all individuals had come in contact with the ES and, since the ES in three

countries (Poland, Hungary, and the Czech Republic) had to administer a work test to determine eligibility for unemployment compensation, all individuals got minimal placement services. Therefore, the analysis of the ES involved comparing those who got minimal services against those who sought out and used more services. Because of this type of analysis, the significance of the impact of the ES may be understated, except in Turkey, where there was no unemployment benefit and no formal requirement to provide minimal services to all registrants. One way to address this problem in a future study would be to select a comparison group from a general household survey, as these individuals would not necessarily be registered at the ES. However there was not sufficient time or resources to attempt this type of analysis during the Study.

3.26 Notwithstanding the above problem, the observable characteristics of ES users are summarized in Table 3.13. The table indicates that both genders tended to use the employment services, the majority of people using the ES had primary and secondary levels of education, and in Poland users tended to be younger than non-users.

Table 3.13:
General Characteristics of Employment Service Users

	Czech Republic	Hungary	Poland	Turkey
Gender				
Male	37	57	46	59
Female	63	43	54	41
Average Age	33	33	25	28
Education				
Primary	62	46	53	N/A
Secondary	33	41	45	N/A
Post Secondary	5	02	02	N/A

3.27 The use of ES assistance appeared to help individuals gain initial employment in Turkey, but did not have an impact in the other countries. The positive finding in Turkey (where there was no requirement for Unemployment Compensation beneficiaries to use the ES) combined with the neutral impacts found in the other countries (where there were minimal identifiable differences in usage across the groups) tends to support the conclusion that the ES can have a positive impact, and that the findings of the Study in the Czech Republic, Hungary and Poland are suspect due to the previously described

sampling problem. The downstream impact of the ES on employment and earnings was less pronounced, perhaps because individuals do not use the ES to change jobs after gaining initial reemployment. The negative impact on current earnings in Hungary may be attributable to individuals taking lower paying jobs proposed by the ES, as opposed to staying out of the labor market longer and gaining reemployment at a higher wage rate. Survey results indicate that participants in Poland and Hungary consistently reported that the ES did help them gain employment.

3.28 Detailed data on subgroup impacts for the ES is not reported here because of the previously outlined data problems. However, the findings do indicate that while the impact on males is not significant, the impact of the ES for females is positive. Use of the ES appears to help both long and short-term unemployed re-enter the labor market. In addition, in Hungary and Poland, the ES had a positive impact on initial employment in areas of both low and high unemployment.

3.29 A secondary objective of the ES is to screen and counsel the unemployed prior to entry into other more expensive and extensive ALPs. The Study did attempt to analyze the interaction between the ES and other ALPs. In other words, does ES screening improve the employment impact of retraining, public works, wage subsidies and self-employment? The evidence suggested that use of the ES can have a positive impact on initial employment after most ALPs, but the benefit appears to disappear within two years of ES use. As previously indicated, the primary service provided by the ES was job placement. Services for job counseling and assessment were very limited.

IV. CONCLUSIONS AND IMPLICATIONS FOR POLICY

The Objective in Evaluating Active Labor Programs:

4.1 The primary objective of the Study was to determine if ALPs have significant positive net impacts, that is, if program participants have significantly better reemployment success than others with similar characteristics who did not participate in an ALP. All five ALPs evaluated were found to have a significant positive net impact for some population subgroups, a general finding supported by the 1996

OECD review that concluded that there are ALPs that work for most groups of individuals. However, the Study found the impact was not significant for some sub-groups, and for some ALPs the impact was negative. This final chapter summarizes the general trends in findings on the various outcome measures across demographic and regional subgroups.

Overall Implications of the Study

4.2 The findings of the Study have three broad implications:

- *Evaluating Programs.* Net impact evaluations should be used to supplement ongoing performance management systems that track gross outcome indicators of programs. Middle income countries can implement both performance management, and quasi-experimental design evaluation programs without large investments, when compared to the potential savings and more effective use of ALP program resources. These findings are similar to those from the previously referenced OECD and ILO reviews which conclude that the impact of ALPs varies widely between and within countries and that it is difficult to make generalizations.

- *Targeting Programs.* The general trends and variations, identified in ALP impacts across subgroups of participants emphasize the value of carefully targeting existing programs to selected participants to maximize the social benefits resulting from the public money expended. Both the ILO and OECD reviews reached the same general conclusion.

- *Designing Programs.* A review of the design of programs, and their impact, provides evidence that program design is a significant factor in determining net impact. The ILO review reached similar conclusions and emphasized the need to examine the administration, management, and institutional context of the delivery of ALPs.

Implementation of Evaluation Programs

4.3 *Administration of evaluation programs:* Agencies administering ALPs can benefit by

becoming directly involved in evaluating the impact of programs. The Study was carried out under the leadership of the national employment services of the countries involved, with support from related Ministries of Labor. This approach worked well. It helped provide access to necessary information, and increased the likelihood that the results would be used in future deliberations on program design. Furthermore, the involvement of agency staff helped to develop the internal expertise needed to replicate and extend the investigation.

4.4 It could be argued that, to ensure objectivity, such evaluation studies should be carried out completely by bodies independent of the administrating agency. In two of the countries in this study, the administration of surveys was carried out by an independent private contractor, to help ensure the validity of the results. The benefits of using this approach must be weighed against its increased cost (i.e., direct data collection costs were increased threefold), and the idea that the use of outside contractors to conduct surveys may have a negative impact on building the capacity of ALP implementing agencies to undertake evaluations, and their ability and willingness to understand and use the results to refine program operations. The OECD and ILO reviews emphasized the need make evaluations more rigorous, address net impact questions, and the OECD suggested a need to test alternate models for quasi-experimental design analysis.

4.5 *Performance Monitoring:* The development of performance indicator monitoring systems, automation of employment service records, and delineation of specific budgets to finance the direct costs of follow up surveys of participants and comparison groups are important precursors for implementation of quasi-experimental design studies. In the countries where performance indicator systems had been developed, there was much greater readiness for undertaking the net impact evaluation.

4.6 *Technical Assistance:* The use of some outside technical assistance, international or local, to assist with study design, sample selection, and data analysis will normally be necessary as these skills are not commonly available within agencies administering ALPs.

Net Impact of Training Programs

4.7 Training can have a positive impact on employment and earnings. Short-term training is as effective as long-term training. Individual "voucher type" training may be more effective than group training, if service providers are available. Private and public providers can deliver effective training, but private and enterprise provision may be more effective. This general conclusion is supported by the ILO and OECD reviews. With respect to subgroups, the findings indicate that training can have a positive net impact for: both for males and females, but the impact may be higher for females; for young and middle aged workers, as opposed to older workers; both short and long-term unemployed, but the impact may be greater for the short-term unemployed; and those with primary and secondary education, as opposed to those with post-secondary qualifications. The finding for women is similar to the finding by the OECD, but the finding on lower educational qualifications is not supported by OECD reviews. Finally, it should also be noted that the unit costs of retraining are the second lowest of the five ALPs studied, and considerably cheaper than public service employment, wage subsidy, and self-employment programs.

Net Impact of Public Service/Temporary Community Employment Programs

4.8 Public service employment (PSE) has no impact, or a significantly negative impact, on post-program employment and earnings, and should not be considered as a active labor program which assists in labor redeployment. PSEs should be looked to primarily as a targeted income support program. This conclusion is supported by other reviews of PSE programs (Fretwell and Wilson, 1999) and by the OECD review of ALPs. However, it is important to recognize the social value of PSE projects which may include helping to maintain political stability in times of high unemployment and rebuilding of public infrastructure. If PSE programs are going to be implemented, strong consideration should be given to having them operated by private sector employers, because evidence from the Study indicates that this approach appears to result in more positive post-program employment impacts than when the programs are implemented by public agencies. PSEs are increasingly being used as work-tests to eliminate individuals, who may have had hidden employment from other income support

programs. If this is to be done with equity, access concerns must be addressed and the program must be carefully designed since the findings in two countries indicate that PSE participants actually used significantly more unemployment benefits than non-participants.

Net Impact of Wage Subsidy Programs

4.9 Wage subsidy program impacts on employment and wages are mixed; there is more impact on employment than wages, and the impact varies considerably based on program design. Shorter programs appear to have more impact on employment than longer ones and impact tends to dissipate over time as the benefit to employers is removed. Programs tend to benefit all age groups, which is somewhat different than the OECD finding which noted a lack of impact on youth. Females and individuals with lower educational qualifications tend to benefit the most. The finding on females was supported by the OECD review. The problem of *dead-weight* and the issue of *displacement* must be addressed during program design and several approaches can be used including: checking for prior related layoffs at enterprises involved in the program, following up to see that employers retain participants, and requiring payback of benefits if the participant is not retained.

Net Impact of Self-employment Programs

4.10 Self-employment programs have a generally positive impact on employment, but mixed impact on earnings, perhaps because of the personal investment required to get a small business underway or under reporting of self-employment income. Program content was quite different between countries (e.g., micro-credit, extended income support, technical assistance) which suggests that there are alternative schemes that should be considered, given the fact that micro-credit schemes tend to be more expensive than other approaches. The programs appear to benefit males and females equally, and middle aged and older workers more consistently than youth. The findings are slightly different from the OECD review that concluded that the impact was most significant for males under 40. The programs significantly and equally benefited participants with primary and secondary education levels, and more so than those with higher education. This conclusion is somewhat different than findings of other studies,

including the OECD review, that indicate that participants with higher levels of education do better than less educated participants in self-employment programs. The difference may partially reflect a high demand for small business services in the countries studied.

Net Impact of Employment Service Programs

4.11 The study findings on the impact of employment services were not conclusive due to problems encountered in defining clear participants and a comparison group sample. However, the Study provided general support for the idea that employment services, in this case primarily placement services, can have a significant impact on helping individuals find initial employment, but there was no evidence of impact on wages. The conclusion that job search assistance is generally effective is supported by the findings of the ILO and OECD reviews. There was also a positive linkage between employment services and selected active labor in several countries programs (e.g., employment services improved the impact of retraining). Given the fact that employment services have the lowest unit cost of any ALP, it would appear that continued support of this program may be warranted for its own sake, as well as because of its linkage with other ALPs. The ILO review also concluded that combining measures into integrated packages can increase their effectiveness.

Impact on Special Target Groups

4.12 The previous paragraphs present Study conclusions by "program." An alternate approach to analysis and design of ALP policy could be based on identification of programs to meet the needs of selected target groups (e.g., women, youth, the unemployed with low levels of basic education) who may be particularly at risk during periods of economic restructuring. The Study concluded that:

- Women benefit as much, or more, than men from participation in most ALPs. This general conclusion was supported by the OECD review. The Study found that women benefited more than men from participation in wage subsidy and retraining programs, and equally with men in self-employment programs.

- Youth appear to benefit from training and wage subsidy programs. These conclusions are somewhat different from OECD findings which indicate a lack of impact of wage subsidy and training programs for youth.

- Low skilled/educated workers, who are often among the first and most adversely affected by economic restructuring, benefited from retraining, wage subsidy, and self-employment. The latter finding differs from the conclusion of the OECD review, but may be explained by the emerging opportunities for small business in the countries in the Study. The Study findings indicate that public service employment programs had a negative impact on this group of unemployed, in-fact more negative than for other groups.

Net Benefit Cost Analysis

4.13 One aim of the study was to generate net benefit estimates for ALPs. This proved to be a difficult objective to meet, but should be retained as an important element of such studies. If this activity is integrated into similar studies in the future more resources need to be allocated. In this study some cost data was available (e.g., unit costs of programs, additional months of unemployment benefits), as was some benefit data (e.g., savings on income support, wage gains), but other critical data (e.g., additional months of income support for non-participants, increased productivity data on participants, value of goods and services produced by public service employment programs) were not available. In addition, there is some question, based on results of similar studies (Benus, 1996) whether the wage data generated by follow-up surveys provides a good indication of long-term impact. An attempt was made, in the Poland country study, to generate net benefit estimates for the national labor office, the national government, and all society. The findings provide some guidance as to methodology for further studies, but the generally negative results should not be regarded as conclusive, given the data problems encountered. The ILO and OECD reviews did not address this topic in-depth or present net cost benefit data for programs studied.

Further Questions for Research

4.14 The study pointed to some areas where refinements are needed, if this type of evaluation is to be replicated:

- Future studies should attempt to identify a comparison group sample outside the individuals registered at the employment service, perhaps by using the household survey sample frame (when one exists). This may strengthen the comparison group sample, and is required if analysis of employment service participants is to be accomplished.

- Further research needs to be done to evaluate the impact of employment services, and the sample should be of sufficient size to evaluate sub-categories of services.

- Net cost benefit analyses need to be completed, however this will require considerable additional effort to define and evaluate additional data sources.

- Sequential follow-up surveys should be considered as an adjunct to the Study, and further research activities of this type, to evaluate longer term program impact, including wage impact; both the OECD and ILO reviews reached the same conclusion. The OECD review made the point that some programs may work better after they have been running for longer periods, and individuals may not reap the benefits for some time. Some data from the Hungary study tends to support this conclusion with regard to the impact of retraining programs.

Implications for Bank and Financing of Active Labor Policies and Programs

4.15 The conclusions of the Study have several key implications for Bank technical assistance and lending for development and implementation of labor policies and programs, including ALPs. Contacts with Borrowers should always include discussions of methods of evaluating the gross and net impact of employment policies on short and long-term employment. Bank lending for institutional development, and or/service delivery, should include development of capabilities to design and maintain evaluation

systems for labor programs. Discussions with Borrowers regarding implementation of ALPs should highlight the fact that there is a strong possibility that the programs will have a low employment and earnings impact unless they are carefully designed and targeted. Bank staff should provide background papers on the design and evaluation of ALPs, including this Study as well as ILO and OECD reviews, to Borrowers, and encourage them to visit neighboring countries which have completed rigorous evaluations of ALPs in order to improve design of their programs and related evaluation systems.

REFERENCES

Benus, Jacob M., Terry R. Johnson, Michelle Wood, Neelima Grover and Theodore Shen (1995), *Self-employment Programs: A New Reemployment Strategy*, Unemployment Insurance Occasional Paper 95-4, Washington: U.S. Department of Labor, Employment and Training Administration.

Benus, Jacob M. (1998a), Evaluating the Effectiveness of Active Labor Programs in THE CZECH REPUBLIC, Bethesda: Abt Associates, Inc., May.

_____ (1998b), Evaluating the Effectiveness of Active Labor Programs in TURKEY, Bethesda: Abt Associates, Inc., May.

Fay, Robert G (1996), *Enhancing the Effectiveness of Active Labor Market Policies: Evidence from Program Evaluations in OECD Countries*, Labor Market and Social Policy Occasional Papers No. 18, OECD, Paris.

Fretwell, David and Susan Goldberg (1993), *Developing Effective Employment Services*, World Bank Discussion Papers No. 208, Washington D.C., 1995.

Fretwell, David and Sandra Wilson (1996), *Public Service Employment - A Review of Programs in Selected OECD Countries and Transition Economies*. World Bank Discussion Papers, Washington D.C., 1999.

Meager, Nigel and Ceri Evans (1998), *The Evaluation of Active Labor Market Measures for the Long-Term Unemployed*, Employment and Training Papers No. 16, ILO, Geneva.

O'Leary, Christopher J. (1995), "Performance Indicators: A Management Tool for Active Labor Programs in Hungary and Poland." *International Labor Review*, Volume 134, Number 6.

_____ (1997), "A Net Impact Analysis of Active Labor Programs in Hungary," *The Economics of Transition*, Volume 5, Number 2, November.

_____ (1998a), *Evaluating the Effectiveness of Active Labor Programs in POLAND*, Kalamazoo: W.E. Upjohn Institute for Employment Research, April.

(1998b), *Evaluating the Effectiveness of Active Labor Programs in HUNGARY*, Kalamazoo: W.E. Upjohn Institute for Employment Research, May.

Annex I: Details of Project Design

Three of the four countries participating in the Study; the Czech Republic, Hungary, and Poland were former centrally planned economies. These countries have experienced high levels of unemployment due to economic restructuring, including significant downsizing of the state production sector with parallel increases in employment in the service sector and private firms. During this period they rapidly implemented a number of passive income support programs (unemployment benefits, means tested social assistance) and parallel active labor programs, These three countries have recently joined the OECD. Turkey, the fourth country in the study, is a long-time member of the OECD. Turkey, in comparison to the other countries, has a larger proportion of the workforce in the rural agricultural sector, and is experiencing considerable rural to urban migration that has significant implications for unemployment. Turkey is currently considering legislation to implement an unemployment benefit, upgrade employment services, and to broaden the types of ALPs available to the unemployed.

An overall Project Steering Committee (PSC) was established consisting of representatives from the World Bank, the four participating countries, the external financing agencies, and the two technical assistance contractors. The PSC agreed that the programs to be studied in each country should include the most widely used active labor programs in each country. To the extent possible similar programs across countries were selected to provide a basis for cross-country comparison.

At the inception of the study it was envisioned that the design and implementation phases would be done under a general term of reference, agreed to by the PSC, and that financing would be sought for the entire study at one time. However, after PSC discussions in 1995, it was agreed that it would be difficult to integrate and finance these two phases in one overall Terms of Reference. As a result a detailed design phase was completed first. This was reviewed and finalized within the PSC in June 1996. Next financing was arranged and implementation was completed in 1996 and 1997. The overall study, from initial discussions to in-country dissemination of results, took approximately 24 months.[7]

[7] The Terms of Reference for the Study is available from the study coordinator at the World Bank.

The study built upon previous work undertaken in Poland and Hungary, supported by World Bank projects, which facilitated the development of national performance indicator evaluation systems for ALPs. These systems provide a systematic method of identifying gross ALP program impacts, including defining performance indicators and participant follow-up procedures. The national performance indicator systems in Poland and Hungary were developed as part of the World Bank financed Poland Employment Promotion and Services Project, and the Hungary Human Resources Project[8]. Such data and procedures provide a useful foundation for the more complex quasi-experimental design evaluation as reported in this paper.[9]

It was also was agreed that results from the study would be disseminated nationally and internationally via conferences, workshops, and written documentation. Each country held a national dissemination by late 1997, where an in-depth summary of the design, implementation and results for each country was presented by the technical assistance contractors and country representatives. An international seminar, financed by the World Bank Economic Development Institute and hosted by the Turkish Employment Agency, was held in June, 1998. Over one hundred representatives from some 17 countries and international organizations attended the Conference.

[8] More details on these projects can be obtained from the World Bank by reviewing the Project Documents including the Staff Appraisal Report, and by reviewing a summary paper produced by the technical assistance contractor, W.E. Upjohn Institute for Employment Research (O'Leary, Christopher J., "Performance Indicators: A Management Tool for Active Labor Programs in Hungary and Poland." *International Labor Review*, Vol. 134, No. 6, 1995).

[9] A small scale net impact evaluation of retraining and public service employment in Hungary was conducted in 1992 and 1993 using a quasi-experimental design (O'Leary, Christopher J., "A Net Impact Analysis of Active Labor Programs in Hungary," *The Economics of Transition*, Vol. 5, No. 2., 1997).

Annex II: Czech Republic

	Units	1992	1993	1994	1995	1996
Population and Demographic Data						
Total Population (mid-year)	Thousands	10319.1	10329.9	10333.6	10327.3	10315.2
Share of children (0-17)	%	25.6	24.9	24.2	23.5	22.7
Life expectancy						
male	Yrs	68.5	69.3	69.5	70.0	70.4
female	Yrs	76.1	76.4	76.6	76.9	77.3
Labor Markets						
Labor force	Thousands	5061.8	4986.2	4819.6	5042.1	5062.2
Participation rate		78.2	79.6	80.4	79.3	
male	%	--	82.8	83.0	98.3	
female	%	--	77.0	77.9	61.5	
Unemployment rate (registered)	%	3.1	3.0	3.3	3.0	3.1
by education level						
primary or less	%		8.8	8.1	10.3	11.2
secondary general	%		2.7	2.6	2.0	2.1
secondary vocational	%		3.6	3.7	3.4	3.5
higher	%	--	1.9	1.4	1.1	0.6
by age						
youth	%	--	6.4	6.7	6.3	
prime age	%	--	3.4	3.4	3.0	
older (within 5 yrs of retirement)	%	--	1.9	1.7	2.2	
Economic Context						
Real GDP growth rate	%	-6.4	-0.9	2.6	5.0	4.8
GDP per capita	US$	2709.9	3023.9	3487.4	4567.0	5048.0
CPI Inflation (annual average)	%	11.1	20.8	10.0	9.1	8.8
Average monthly industrial wage	US$	169.9	202.2	239.6	304.5	309.7

Sources: TransMONEE Database, UNICEF-ICDC Florence; Labor Market Database (Q2 survey data); MultiQuery Database

Annex II: Hungary

	Units	1992	1993	1994	1995	1996
Population and Demographic Data						
Total Population (mid-year)	Thousands	10323.7	10293.6	10261.3	10229.0	10193.4
Share of children (0-17)	%	24.5	23.9	23.3	22.8	--
Life expectancy						
Male	Yrs	64.6	64.5	64.8	65.3	66.1
Female	Yrs	73.7	73.8	74.2	74.5	74.7
Labor Markets						
Labor force	Thousands	5468.9	5030.0	4770.6	4500.3	4074.9
Participation rate	%	72.9	70.1	67.9	73.0	
male	%	75.1	72.6	72.3	--	
female	%	68.9	66.1	63.3	--	
Unemployment rate (registered)	%	12.3	12.1	10.4	10.4	10.5
by education level						
primary or less	%	14.3	17.4	16.0	16.0	16.3
secondary general	%	6.7	8.6	7.8	6.8	7.2
secondary vocational	%	11.3	14.5	12.8	12.2	12.7
higher	%	2.3	2.9	2.9	2.7	3.1
by age						
youth	%	18.7	22.5	20.5	19.8	
prime age	%	8.6	10.8	9.5	9.0	
older (within 5 yrs of retirement)	%	8.3	10.4	5.8	5.5	
Economic Context						
Real GDP growth rate	%	-3.1	-0.6	3.0	1.5	0.2
GDP per capita	US$	3603.9	3743.5	4054.6	4266.3	4249.0
CPI Inflation (annual average)	%	23.0	22.5	18.8	28.2	23.6
Average monthly industrial wage	US$	279.6	300.0	317.9	314.1	313.9

Sources: TransMONEE Database, UNICEF-ICDC Florence; Labor Market Database (Q2 survey data); MultiQuery Database

Annex II: Poland

	Units	1992	1993	1994	1995	1996
Demographics						
Total Population (mid-year)	Thousands	38365.0	38459.0	38544.0	38588.0	38618.0
Share of children (0-17)	%	29.3	28.9	28.4	27.9	27.3
Life expectancy						
Male	yrs.	66.7	67.4	67.5	67.6	68.1
Female	yrs.	75.7	76.0	76.1	76.4	76.6
Labor Markets						
Labor force	Thousands	17520.2	17650.8	17496	17710.4	
Participation rate		59.2	59.0	70.7	61.7	
male	%	61.6	75.6	75.2	--	
female	%	56.6	65.7	64.8	--	
Unemployment rate (registered)	%	13.6	16.4	16.0	14.9	13.6
by education level						
primary or less	%	12.0	14.0	15.0	13.8	13.7
secondary general	%	13.8	13.0	13.1	12.0	11.8
secondary vocational	%	15.5	17.4	17.4	15.7	15.4
higher	%	6.6	6.0	5.2	4.1	3.6
by age						
youth	%	26.1	29.8	30.8	33.7	
prime age	%	12.2	13.0	13.2	13.6	
older (within 5 yrs of retirement)	%	1.7	2.0	1.8	8.4	
Economic Context						
Real GDP growth rate	%	2.6	3.8	5.2	7.1	6.0
GDP per capita	US$	2198.0	2232.3	2401.9	3057.4	3495.5
CPI Inflation (annual average)	%	43.0	35.3	32.2	27.8	19.9
Average monthly industrial wage	US$	218.5	234.3	264.8	328.0	373.4

Sources: TransMONEE Database, UNICEF-ICDC Florence; Labor Market Database (Q2 survey data); MultiQuery Database

Annex II: Turkey

	Units	1992	1993	1994	1995	1996
Population and Demographic Data						
Total Population (mid-year)	thousands	58,401,001	59,491,000	60,575,998	61,643,997	66,697,000
Life expectancy						
Male	yrs	70	70	71	66	66
Female	yrs	65	65	65	70	71
Labor Markets						
Labor force	thousands	21504	21469	22158	22674	22919
Participation rate	%	52	50	51	51	50
male	%	73	71	72	71	70
female	%	31	30	30	31	30
Unemployment rate (survey)	%	8.1	7.8	8.1	6.9	6.0
by education level		8.1	7.8	8.1	6.9	6.0
primary	%	7	7	7	6	--
Junior high school	%	12	12	12	10	--
voc/tech junior high school	%	10	10	15	12	--
high school	%	17	16	16	14	--
voc/tech high school	%	14	12	14	14	--
University & other post-secondary	%	8	7	7	6	--
by age						
Economic Context						
GDP growth rate	%	6.4	7.8	-5.7	7.8	
GDP per capita	US$	2,682	2,981	2,173	2,727	
Inflation (annual)	%					
Average monthly industrial wage						

Annex III Details of Sample Selection

Site and Sample selection for Hungary

The sample for analysis in Hungary was drawn from randomly selected samples in a strategically selected group of ten counties which, taken together, comprised nearly two-thirds of the nation's population: Budapest (the capital city), Baranya, Bekes, Borsod, Csongrad, Fejer, Hajdu-Bihar, Pest, Szabolcs, and Vas.[10] In 1996, these counties spanned the full range of economic conditions. Three counties enjoyed an unemployment rate of below 8 per cent, three suffered unemployment rates in excess of 15 per cent, and four had moderate rates of unemployment. Some had experienced steady labor market improvement since the peak of national unemployment reached in early 1993, while others had stagnated. Compared with the country as a whole, these counties had a somewhat smaller proportion of employment in agriculture, a higher population density, a lower unemployment rate, and higher mean monthly wages.

The surveys were administered by experts from the National Labor Center. Surveys were conducted in March and April 1997 in house-to-house visits by officials of local labor offices during their off-work hours.[11] Program participant groups were drawn from those completing their participation in the programs during the second quarter of 1996. There was random sampling from this outflow of participants where sample sizes were large enough, with random draws made according to birth date. As regards the small number of participants in self-employment, an attempt was made to contact all those who had participated in this program during the first three quarters of 1996. The comparison group was randomly selected, using birth dates, from the inflow to the register in the ten counties during the second quarter of 1995. That was judged to be about the period during which most people drawn for the participant samples had themselves registered as unemployed. The following

[10] Provincial divisions in Hungary are called counties and in Poland *voivods*.

[11] Some interviews were conducted during regular visits by the unemployed to labour offices. Such a process means estimates of the impact of ALPs on re-employment rates may be biased downwards since the unemployed are more likely to visit labour offices, and the employed are less likely to be available at home during house-to-house visits.

characteristics were used to compute regression adjusted net impact estimates for active labor programs in Hungary: prior average monthly wage, age, gender, education, prior labor market status, special job finding difficulty indicator variable, indicator variables for occupation wanted, spouse present indicator, spouse employed indicator, number of children under 6, number of children over 6 plus other dependents, net monthly household earnings, county of residence indicator. Note, some net impact estimates also controlled for use of placement services of the local labor office through an interaction model

The comparison group and the samples of ALP participants in Hungary are statistically significantly different on several demographic characteristics. [12]By contrast with the comparison group, participants in the individual retraining and group retraining samples included more women, and were younger and better educated; participants in the PSE sample were predominantly male, younger, and less educated; those in the wage subsidy sample were somewhat better educated; and those in the self-employment sample were predominantly male, closer on average to prime working age, and better educated. The following characteristics were used to draw a matched-pair comparison group for each active labor program participant sample: age, education, gender, months of work experience, date of registration as unemployed, and local labor office where registered as unemployed.

The substantial differences in sample composition suggest that there was non-random assignment of participants to particular ALPs. This means that estimates of net impact must be computed while controlling for systematic selection bias. In this article, correction in estimation is limited to adjustments based on observable characteristics.[13] The estimation methodology used and the use of a comparison group purge the net impact estimates of the effects of any creaming practiced by program administrators.

[12] In Hungary, the survey response rate among ALP participants was 81.4 per cent, while that for the comparison group was 75.6 per cent.

[13] The estimates presented in this article were all computed using an ordinary least squares regression model, which controls for observable characteristics and for use of particular ES assistance.

Site and Sample selection for Poland

The data needed to evaluate ALPs in Poland were gathered through surveys of randomly selected participant samples and strategically selected comparison samples in a group of eight *voivods*: Gorzow, Katowice, Konin, Krakow, Lublin, Olsztyn, Poznan, and Radom. Though these locations were chosen partly because of similarities in information processing, they none the less span the full range of labour market experience in Poland during the transition to a market economy.[14] Of the eight *voivods* surveyed, four are among Poland's most populous: Katowice, Krakow, Lublin, and Poznan. The eight encompass over a quarter of the population of Poland, including several large cities, and therefore yield an above-average population density. These areas also have much lower unemployment rates, somewhat higher wages, and a smaller share of agriculture than the country as a whole.

Surveys were conducted in 80 local areas between 15 February and 15 April 1997. The questionnaires were administered by experts from the *voivod* labour offices and interviews were conducted by officials from local labour offices. Some interviews were carried out during regular visits to labour offices by individuals who had previously been selected; others were carried out during house-to-house visits. The overall response rate was 92.6 per cent.

The sampling frame for participants in retraining, public works, and intervention works was entry into an ALP during the course of 1995. Random sampling of participants was done by birth date. Since a longer period is required to assess the effects of self-employment assistance, receipt of a loan during 1993 and 1994 was taken as the sampling frame. The small numbers involved meant that instead of random sampling from self-employment participants, an attempt was made to contact the whole population of recipients of assistance. For other programs, sample sizes for each *voivod* were set to be proportional to the *voivod* share of participation in programs. Once the participant samples had been selected, the observable exogenous characteristics of the selected groups were examined. The comparison group samples were drawn from the population of registered unemployed by matching

[14] A dozen different local labour office computer systems were in use around Poland at the time of the survey. Two different systems were involved in the eight *voivods* surveyed.

persons in each of the ALP participant samples to the most similar person from the unemployment register of the same local labour office. Separate comparison group samples for each program were selected from among those who had registered as unemployed within the same time period and had never participated in an ALP.

By contrast with a random sample of unemployed, participants in the retraining sample were predominantly female, younger, better educated, less likely to be in a blue-collar occupation, and more likely to be in long-term unemployment; participants in the public works were predominantly male, younger, and less educated; those in the intervention works sample were predominantly female, younger, less likely to be in a blue-collar occupation, and more likely to be in long-term unemployment; and those in self-employment were predominantly male, more likely to be of prime working age and to have received a vocational education, and slightly less likely to be in long-term unemployment.

Site and Sample Selection in Turkey

Five sites, which maintained computerized records that were linked electronically to the national office: **Istanbul, Ankara, Izmir, Adana, and** Bursa were selected for the study. These five sites represented the five largest cities in Turkey, containing a significant proportion of Turkey's population. Furthermore, these five sites represent more than half of the training courses and more than half of the individuals trained by the Employment Guarantee Training Program (EGTP). Istanbul has, by far, the most trainees -- nearly seven thousand. Izmir, with over four thousand, has the next largest number of trainees. Together, Istanbul and Izmir represent over half (54 percent) of the total number of program participants in 1995 (10,976 out of a national total of 20,037).

Based on data availability and other considerations, we selected the above five sites for our study. Specifically, the availability of electronic data files substantially reduces the cost of sample selection and enhanced the ability to select a representative sample of participants and non-participants. The five sites do not include any rural areas, therefore impact estimates may only be reflective of the impact ALPs in urban areas. Furthermore, the results are also significantly influenced by two large urban areas: Istanbul and Izmir.

The sample size was specified to be of sufficient size to be able to derive accurate impact estimates of the Employment Guarantee Training Program. This requirement, in conjunction with a budgetary constraint, resulted in a sample size of approximately 2,000 participants and 2,000 comparison group members. Participants were selected from among those who enrolled in training courses; comparison group members were selected from among IIBK registrants who did not participate in the EGTP. To assess the net impact of the training program, the Study included a dummy variable for participation status in these regressions (i.e., P=1 if in the participant group and P=0 if in the comparison group). In addition, to help isolate the impact of the training program, the Study included a number of other independent variables in the regressions. All the regression equations reported in this report include the following dummy variables: male; 20-24 years old, 25 years old and older; education (middle), education (secondary), education (other); Istanbul, Ankara, Izmir, and Adana; family size (4), family size (5), family size (6 or more); long term unemployed; prior employment status (regular), prior employment status (student/housewife), prior employment status (unemployed); and IIBK helped in job placement.

Given the small number of 1995 EGTP participants in the three smaller sites (Ankara, 329; Bursa, 224; and Adana, 115) the sample included all 668 program participants in these sites. In the larger sites (Istanbul and Izmir), the Study selected a random sample of the program participants in 1995. The size of these samples in these cities was proportionate to their number of trainees. Thus, in Istanbul the Study selected approximately 800; in Izmir, approximately 600.

For a variety of reasons the IICK believed it was preferable to have an outside contractor, rather than the local employment office, implement the follow-up surveys. As a result, a private survey firm, Strateji|Mori, Inc., to was contracted to refine, translate and implement the survey. The original intention was to conduct in-person interviews only. However, to conserve the survey budget, it was necessary to utilize both telephone and in-person interviews. Approximately one-third of the surveys

[2] All the regression equations reported in this report include the following dummy variables: male; 20-24 years old, 25 years old and older; education (middle), education (secondary), education (other); Istanbul, Ankara,Izmir, and Adana; family size (4), family size (5), family size (6 or more); long term unemployed; prior employment status (regular), prior employment status (student/housewife), prior employment status (unemployed); and IIBK helped in job placement.

were conducted by telephone; the remainder were conducted in-person.

The overall response rate for the entire sample was 39 percent. For the participant group, the response rate was 43 percent, slightly higher than for the non-participant group (37 percent). This pattern of higher response rate for the participant group than for the non-participant group was to be expected since the participant group had more contact with the IIBK and, thus, their address information was better.

There was little pattern in the response rate across sites. In some sites, for example, the response rate is higher for the participant group; in other sites, the response rate is higher for the non-participant group. Ankara had the highest response rate for both participant and non-participant groups. Overall, Ankara's response rate was 54 per cent.

Site and Sample Selection in the Czech Republic

The Czech Republic consists of 8 administrative regions and 76 counties (or districts). Based on criteria developed by the researchers and the Ministry of Labor and Social Affairs (MOLSA), 20 districts were selected to represent the nation. The selected districts represented a wide range of unemployment rates. For example, the selected sites included areas of relatively high unemployment like Most (7.6%) and Karviná (6.6%) and areas with hardly any unemployment like Jind ich v Hradec (0.8%) and Praha-západ (0.5%). In addition, the MOLSA selected districts that were spread throughout the country and districts with different types of economic production activity (e.g., industrial, agricultural, etc).

For each program, the evaluation sample was selected from among all individuals who registered at the employment service during 1994 and enrolled in an ALP in 1994 or 1995. The participant sample was thus a random sample of 1994-95 ALP participants. The sample selected in each site was in proportion to the number of program participants in the district. Thus, the analysis sample is reflective of all ALPs participants in the Czech Republic. To assess the net impact of the each ALP, the Study included a dummy variable for participation status in these regressions (i.e., P=1 if in the participant group and P=0 if in the comparison group). In addition, to help isolate the impact of the

training program, the Study included a number of other independent variables in the regressions. All the regression equations reported in this report include the following dummy variables: male; 25-40 years old, 41 years old and older; education (middle), education (secondary), education (university); town size less than 5,000, town size (5,000-10,000), town size 10,001-50,000, town size 50,001-100,000; family size (3), family size (4), family size (5 or more); long term unemployed; prior employment status (regular), prior employment status (student); and PES helped in job placement.

The sample size was specified to be of sufficient size to be able to derive accurate impact estimates of the four ALPs to be evaluated. It was estimated that there was a need to select 1,000 participants and 1,000 comparison group members for each of the four ALPs: (1) Socially Purposeful Jobs, (2) Publicly Useful Jobs, (3) Program for School Leavers, and (4) Retraining. Unforeseen events, however, significantly altered the sample selection and sample size plans as described in the following paragraphs.

For a variety of reasons, the researchers felt it was preferable to have an outside contractor, rather than the local employment office, implement the follow-up surveys. As a result, the follow-up was contracted with a private survey firm, Statistical Consultation and Computing (SC&C), to refine, translate and implement the survey. For cost reasons, however, it was necessary to use local employment service officials as interviewers. Thus, SC&C hired a number of local employment service staff to serve as their interview staff. To minimize the potentially biasing effects discussed above, the interviews were conducted in-person and away from the local employment office.

The Ministry of Labor and Social Affairs ruled that researchers could not able to gain direct access to the addresses of registered persons. Rather, potential respondents had to be contacted by the Labor Office and be asked for written permission to have their addresses given out for research purposes. Clearly, this had a devastating impact on sample plans. The immediate affect of this new requirement were: the schedule was dramatically delayed; costs were substantially increased due to the additional mailing costs; and he response rate was dramatically reduced (since most people did not

[1] All the regression equations reported in this report include the following dummy variables: male; 25-40 years old, 41 years old and older; education (middle), education (secondary), education (university); town size less than 5,000, town size (5,000-10,000), town size 10,001-50,000, town size 50,001-100,000; family size (3), family size (4), family size (5 or more); long term unemployed; prior employment status (regular), prior employment status (student); and PES helped in job placement.

respond to the letter).

The field surveys were conducted during March-May of 1997. While the original goal was to collect approximately 8,000 surveys, unexpected events in the field led to a final sample of 4471, slightly more than half the desired sample size (56%).

A total of 24,973 Labor Office registrants were contacted by mail for permission to be interviewed. In some cases, two mailings were required. Of these nearly twenty five thousand potential respondents, only 4,537 (18 percent) agreed to have their addresses provided to the survey firm. This low response rate was not surprising given the suspicion of government that still exists in many countries. Among those who provided their permission, the response rate was very high. The overall response rate, however, was 18 percent.

Annex IV Details of Statistical Analysis

The main appeal of program evaluation which uses a classically designed experiment involving random assignment is that net impacts are easy to understand, and therefore more influential for the purpose of guiding policy. If random assignment is achieved, modeling of behavior and complex econometric methods are not needed to obtain estimates of the net impact of a program. With large samples randomly assigned to treatment and control groups, observable and unobservable characteristics of the two groups should not differ on average, so that any difference in outcomes may be attributed to the program. Program impact may be measured as the simple difference between the means for treatment and control group members on outcomes of interest.

When there is non-random assignment to either the ALP participant group or the comparison group from the population of unemployed job-seekers, then statistical methods of correction must be used to offset the selection bias in order to estimate the net impact of ALPs.[15] Recent surveys of microeconomic evaluations of ALPs conducted by Fay (1996) for OECD member countries and by Meager and Evans (1998) for a selected group of countries emphasize the importance of accounting for *deadweight loss* and *displacement* effects when measuring the impact of the program. With a mixed bag of findings which reveal that the net impact of different ALPs varies widely from one population subgroup to another, the authors of both surveys argued that *targeting* of services is crucial to maximizing the social dividend from public expenditure on employment programs.[16]

---

[15] Such methods are sometimes called *quasi-experimental* because they attempt to mimic statistically the ideal of a true experiment based on random trials (Fay, 1996). Program impact reported in this article was estimated according to models such as the following:

$$y_i = a_0 + b_1 ALP_i + b_2 ES_i + b_3 ALP_i * ES_i + C X_i + u_i,$$

[16] That is for the following reasons. When an unemployed person participates in an ALP which does not improve his/her chance of re-employment, there is a *deadweight loss* to society for the expenditure incurred. If a program manager practices creaming in selecting participants for ALPs such that the people supported would have secured employment without the assistance, then a deadweight loss also results. When an ALP participant gains re-employment at the direct expense of an otherwise similar job-seeker, then *displacement* has occurred. When an employer, either government or private, receives a subsidy to hire a worker who would otherwise have been hired anyway, then *substitution* of ALP financing for other intended spending has occurred.

It is crucial to account for displacement and substitution effects when assessing the net social benefits of public programs. However, these factors are irrelevant at the individual level and very difficult to measure at the social level. The investigation summarized here focused on the net impact of ALPs, and the design using a comparison group automatically accounted for possible deadweight loss by comparing ALP participants with otherwise similar non-participants. A subgroup analysis of net impact provides a basis for targeting ALPs.

This study applied three ways of measuring program impacts: (a) unadjusted impacts, a simple difference between the participant group and the overall comparison group on outcomes of interest; (b) matched pair analysis, comparing means for participant groups with matched comparison groups, and (c) regression adjusted net impact analysis, estimating program impacts using regression models to adjust for observable differences between participant and comparison groups. Since participant and comparison samples were not created by a random assignment experiment, unadjusted differences yield gross impact rather than net impact estimate. These estimates are useful in comparison to net impacts for investigating the practice of creaming in program assignment. In this Study, net impact estimates were computed by matched pairs and regression adjustment methods. That is sample selection corrections were based entirely on observable characteristics.

Unadjusted Impact Estimates

To make ideas precise, estimation procedures can be stated algebraically. Gross program impacts may be computed as the simple difference between means of the samples of program participants and control group members on outcome measures of interest, or:

(1) $E(y_p) - E(y_c)$,

where E is the expectation operator yielding means of the random variables, y is an outcome of interest, and the index p denotes the sample of program participants while c denotes the comparison sample.

Tests of significance are done using t-statistics.

The result of the computation stated in equation (1) is equivalent to the slope coefficient estimated by ordinary least squares (OLS) applied to a simple bi-variate regression model. That is, program impacts can be estimated by running the OLS model:

(2) $y_i = a_0 + a_1 P_i + u_i$,

on a pooled sample of comparison group members and program participants, where y is the outcome of interest, a_1 is the impact of the program on the outcome for the ALP participants, a_0 is the mean value of the outcome for comparison group members, P is a dummy variable with a value of 1 for active labor program (ALP) participants and 0 otherwise, u_i is a normally distributed mean zero error term, and i is an index denoting individuals in either the participant or comparison group samples. Tests for significance of program impacts are simply t-tests on the parameter a_1.

Impact Estimates Using a Matched Pairs Comparison Group

When participant group and comparison group members differ significantly in terms of observable characteristics, it would not be surprising to observe different labor market success across program participant and comparison groups even in the absence of ALPs. To put the assessment of ALPs on an even footing, a separate comparison group for each sample of ALP participants may be formed using a matched pairs methodology. Matched pairs comparison groups were formed by comparing persons in the ALP participant samples with those in the full comparison group using the standardized Mahalanobis distance measure:

(3) $d_{pc} = Sum_k(Z_{pk} - Z_{ck})^2$

where, the index p represents observations in an ALP participant sample and the index c represents observations from the comparison group, the index k runs over the n exogenous characteristics on

which the observations are matched, and Z represents the standardized value of a characteristic where the mean and standard deviation of the characteristic is computed on the pooled sample of the comparison group sampling frame and the participants in the relevant ALP.

Using this distance measure, separate matched pairs comparison groups were selected for each ALP. The person with the smallest d_{pc} from the full comparison group sampling frame was selected for inclusion in the matched pairs comparison group, with ties being resolved randomly and each person in the ALP sample being compared to all those in the full comparison group sampling frame.[17]

After forming the matched pairs comparison groups, program impact estimates were computed using a simple difference of means, with significance of impacts being judged by t-tests. It should be noted that because a single observation from the comparison sample may be chosen more than once for the synthetic comparison group, the estimated standard error, computed in the usual way, for this group will be reduced. The t-statistics for the matched pairs analysis may therefore be overstated since they depend on a lower bound estimate of the standard error.

Regression Adjusted Impact Estimates

Multivariate regression analysis is a natural method for assessing the net impact of program participation on labor market success when observable characteristics of participant and comparison group members are dramatically different. This method involves a simple extension of equation (2). In such cases, estimation of the model:

$$(4) \qquad y_i = a_0 + a_1 P_i + b_1 X_{1i} + b_2 X_{2i} + ... + b_n X_{ni} + u_i,$$

by OLS on the pooled sample yields net program impact estimates.[18] In equation (4) y is the outcome

[17] That is, sampling was done with replacement.

[18] Since the main dependent variable of interest--in a normal job--is binary, the regression model predicts the probability of reemployment. The OLS estimation is a linear probability model, which may yield biased estimates. OLS estimates may be biased since the range of variation in the dependent variable is constrained to the zero-one

of interest, a_0 is the mean value of the outcome for comparison group members evaluated at the mean of all observable characteristics included in the regression, P is a dummy variable with a value of 1 for program participation and 0 otherwise, a_1 is the impact of the program on the outcome for the program participants evaluated at the mean of all observable characteristics, X_1 to X_n are observable characteristics measured as deviations from their mean values, u_i is a normally distributed mean zero error term, and i is an index denoting individuals in either the participant or comparison group samples.[19]

This method yields net program impacts adjusted for observable characteristics.[20] The estimates are called net because the comparison and program participant groups are statistically adjusted so as to remove heterogeneity across the samples. That is, the only remaining factor contributing to a difference in the outcome measure is exposure to the program treatment. The estimation methodology nets out all other observable factors affecting the outcome.

Subgroup Net Impact Estimation Methodology

For each separate ALP, subgroup treatment impacts were simultaneously estimated in a single regression model. The specification employed allows the treatment response for each subgroup to be estimated controlling for the influence of other subgroup characteristics. For example, the model allows estimation of treatment impacts associated with being female controlling for the fact that females are more likely to have more formal education and less likely to work in a blue collar occupation.

Suppressing subscripts and using matrix notation, the regression equation used to estimate subgroup net impact estimates can be written:

(5) $Y = a + PB + GC + GPD' + u$

---

interval. Bias is usually most severe when the bulk of probability clusters at one or other extreme of the zero-one interval reemployment probabilities for the ALP and comparison groups generally range from about 40 to 60 percent, the limited range of the dependent variable is not a likely source of severe bias in estimating parameters by OLS.

[19] The regression model is a statement of an analysis of covariance methodology, where X_1 to X_n are the covariates.

[20] The next procedure to adjust for differences across samples is to account for differences in unobservable characteristics. The technique, involves applying the methods of Heckman is problematic because instruments are usually not available to explain participation independent of reemployment success.

where Y is the outcome measure, a is the intercept, B, C, and D, are conformable parameter vectors, P is the indicator of participation in an ALP, G is the matrix of dummy variables which code for membership in a subgroup, and u is a mean zero normally distributed random error term. Equation (5) specifies a complete one-way interaction model. It allows simultaneous estimation of all subgroup treatment impacts, but imposes linear restrictions on the estimates. Treatment impacts for a particular subgroup are computed as the sum of the parameter estimate on the product of the subgroup dummy variable and the treatment indicator plus the sum of parameter estimates on the product of subgroup dummy variables and the treatment indicator multiplied by their respective population shares. In each computation, parameter estimates for the complement to the subgroup of interest are omitted.

The subgroup impact estimates may be considered to be regression adjusted in the sense that each subgroup impact is estimated while simultaneously allowing impacts to vary across other subgroups considered.

Methodology for Estimation of Program Components

To estimate the impact of separate features of an ALP on outcomes of interest, new program variables are defined from the single program variable P_i such that the vectors for the new variables add up to the vector for the old variable. For example, if P_i has a value of 1 if participated in an ALP and 0 otherwise, to examine the separate impacts of the ALP operated by public and private enterprises on outcomes of interest we may define $P_{1i} = 1$ if participated in an ALP operated by a public enterprise and 0 otherwise, and $P_{2i} = 1$ if participated in an ALP operated by a private enterprise and 0 otherwise. Therefore $P_i = P_{1i} + P_{2i}$, and the separate impacts of the ALP run by public and private enterprises on outcomes of interest can be estimated by OLS regression applied to a simple model like:

(6) $y_i = b_0 + b_1P_{1i} + b_2P_{2i} + u_i$.