MEMORANDUM

No 15/2002

Do individual programme effects exceed the costs? Norwegian evidence on long run effects of labour market training

By
Oddbjørn Raaum, Hege Torp and Tao Zhang

ISSN: 0801-1117

Department of Economics
University of Oslo

This series is published by the
University of Oslo
Department of Economics

P. O.Box 1095 Blindern
N-0317 OSLO Norway
Telephone: + 47 22855127
Fax: + 47 22855035
Internet: http://www.oekonomi.uio.no/
e-mail: econdep@econ.uio.no

In co-operation with
The Frisch Centre for Economic Research

Gaustadalleén 21
N-0371 OSLO Norway
Telephone: +47 22 95 88 20
Fax: +47 22 95 88 25
Internet: http://www.frisch.uio.no/
e-mail: frisch@frisch.uio.no

List of the last 10 Memoranda:

No 14	Oddbjørn Raaum, Hege Torp and Tao Zhang Business cycles and the impact of labour market programmes. 52 pp.
No 13	Geir B. Asheim, Anne Wenche Emblem and Tore Nilssen Deductibles in Health Insurances: Pay or Pain? 15 pp.
No 12	Oddbjørn Raaum and Knut Røed Do Business Cycle Conditions at the Time of Labour Market Entry Affect Future Unemployment?. 22 pp.
No 11	Halvor Mehlum and Karl Ove Moene Battlefields and Marketplaces. 12 pp.
No 10	Halvor Mehlum, Karl Ove Moene and Ragnar Torvik Plunder & Protections Inc. 14 pp.
No 9	Knut Røed, Peter Jensen and Anna Thoursie Unemployment Duration, Incentives and Institutions - A Micro-Econometric Analysis Based on Scandinavian Data 27 pp.
No 8	Hilde Bojer The Time Cost of Children and Equivalent Full Incomes 11 pp.
No 7	Rolf Aaberge Characterization and Measurement of Duration Dependence in Hazard Rates Models. 26 pp.
No 6	Knut Røed and Tao Zhang: The Duration and Outcome of Unemployment Spells - The role of Economic Incentives. 32 pp.
No 5	Morten Søberg: A laboratory stress-test of bid, double and offer auctions. 34 pp.
No 4	Jon Strand: Environmental Kuzents curves:Empirical relationships between environmental quality and economic development. 21 pp.

A complete list of this memo-series is available in a PDF® format at:
http://www.oekonomi.uio.no/memo/

Do individual programme effects exceed the costs? Norwegian evidence on long run effects of labour market training

by

Oddbjørn Raaum
The Ragnar Frisch Centre for Economic Research, University of Oslo

Hege Torp
Institute for Social Research, Oslo

Tao Zhang
The Ragnar Frisch Centre for Economic Research, University of Oslo

Oslo, July 2002

Abstract
Individual long run effects of a labour market training programme targeted at unemployed adults are evaluated by comparing mean post-training earnings for matched samples of participants and non-participants. Average training effects on the trained are positive and persistent over the post-training period of 5 years. Participants without recent work experience, prior to the training, gain less. For participants with recent work experience the present value of the 5 years accumulated earnings effect exceeds the direct costs of the training.

Keywords: Training unemployed, causal effects, matching estimators, accumulated long run effects.
JEL classification: C14, J64, and J68.

*) Corresponding author; hege.torp@isaf.no.

Acknowledgements: The research is funded by the Research Council of Norway (grant no 124583/510 and no 124613/510), the Ragnar Frisch Centre for Economic Research and the Institute for Social Research. Comments from Karl-Gustav Löfgren, University of Umeå, Ragnar Nymoen, University of Oslo, and Erling Holmøy, Statistics Norway on an early draft of this paper are gratefully acknowledged.

1. Introduction

Microeconomic studies of active labour market programmes typically address: Are future labour market prospects - measured for instance by employment status or earnings - improved by taking part in a programme today? Is the outcome for the participants better than what would have prevailed without the programme? In macroeconomic studies of the total effects, the general equilibrium effects of the active labour market policy, externalities and impacts for others than the participants are also included.

The international evaluation literature is extensive and growing, especially in the United States, but also in Europe, including the Nordic countries. Most of the studies are in the microeconomic tradition, focusing on the average treatment effects for the treated in a specific programme. While micro studies need to convincingly estimate the counterfactual outcome for the treated, the problems for the macroeconomic studies are associated with simultaneity and two-way causality (labour market policy is endogenous) and a small number of observations (national macroeconomic time series and international cross sectional data).

Fay (1996) and Martin and Grubb (2001) review the OECD countries' experiences, Swedish experiences are previously summarised by Björklund (1990, 1993) and recently by Calmfors, Forslund and Hemström (2002), while Danish labour market policy is surveyed by Jensen et al. (1993) and Westergård-Nielsen (2001). Various aspects of the US studies and literature has been reviewed several times, see for instance Barnow (1987), LaLonde (1995), and Heckman, LaLonde and Smith (1999).

According to the reviews the findings are mixed. No consensus about the impact of the active labour market programmes on individual success has emerged from the large number of evaluations in recent years. The same applies to the macroeconomic consequences. The content and the organisation of the programmes, the target groups and recruitment procedures as well as the economic environment at the time of the evaluation differ across the studies. There is also a large variety in evaluation design and estimating techniques. Thus there is no surprise that the results diverge. The mixed results may also reflect a lack of suitable data as well as robust estimation methods.

The reviews cited above (and our own reading) also disclose that the *evidence on long run effects* of the active labour market programmes is scarce. When long run follow-up data are available, gains (if any) do not always persist. Most rigorous microeconomic evaluations provide, however, only short-term effects. Very few studies go beyond two years after the programme.[1]

[1] As pointed out by Grubb and Martin (2001), one to two years of post-training outcomes may well be a too short period for a full assessment of the private and social returns to public investments in active measures.

Labour market programmes - and especially training programmes - should be considered as public investments in human capital. This implies that long run effects, both individual and social benefits and costs, should be taken into account when social returns of the programmes are discussed. The scarcity of long run follow-up data implies that the literature gives few examples of *studies where costs and benefits of programmes are compared.* This is especially the case for Europe, while there are more examples from the US; see Heckman et al. (1999) for a review. The ideal cost-benefit analysis of a labour market programme is, however, demanding.

First, all relevant effects should be included, effects both for the participants and for non-participants. This may include impacts on the well being of the participants, possibly reduced social conflict and e.g. criminality through a reduction in unemployment. Second, information on long run effects is needed, preferably by means of actual data. In practice, however, long run effects often rely on projections based on observed short or medium run effects. Third, outcome measures and cost indicators should all be transferable to comparable units and measured by their social value.

Our study is less ambitious. In line with most micro-studies of programme effects we focus on the *average treatment effect on the treated* (ATET). The benefits of the programme are measured by the accumulated pre-tax earnings effects for the participants. The costs of the programme are measured by the direct operation costs evaluated at market prices, corrected for net marginal costs of public funding.

Thus, as we have no consistent system for estimating the social value of increased employment and the resources used to produce the programme, taking market imperfections, such as restricted competition and taxes, into consideration, we use market prices and gross earnings. To simplify the analyses further the value of time during periods of unemployment and programme participation is set to zero. We also ignore transfers between participants and public sector, such as unemployment benefits and taxes. General equilibrium effects are not addressed, implicitly assuming that earnings of non-participants are unaffected by the programme.

The object of the present study is the Norwegian labour market training programme (L*MT programme)* which is the largest programme in Norway targeted at unemployed adults, offering classroom training in a large number of subjects, mainly vocational, but also some general subjects. The courses last typically 5 - 20 weeks. Similar programmes are found in many other countries. The data at hand are non-experimental, covering all participants in LMT 1992-1993 and all unemployed potential participants for the same period. For these populations we have information on labour market status and annual earnings 1992-1997 as well as a large number of human capital indicators from various administrative registers. Annual earnings measure the out-

come over a post-training period of 4 years and 5 years, for participants in 1993 and 1992 respectively.

It is well known that non-experimental evaluation methods may give *biased estimates* of the impact of programmes. The conventional evaluation bias comprises the bias due to selection on unobservables as well as bias due to non-overlapping supports of the explanatory variables in the treatment sample and the comparison sample (mismatching) and different distributions of these variables within the two samples (misweighting). Evaluation methods based on *matching techniques* may reduce the conventional measure of bias, as far as selection on observables is concerned.

A commonly used conditioning set is the probability of being in the participant group versus in the comparison group. Provided that the outcome is independent of participation conditional on this probability the matching estimator is unbiased. This is the *conditional independence assumption*, CIA. Participation in a specific programme is, however, not the outcome of a simple binary choice, or of a selection process with only two (mutually exclusive) outcomes. First, the target group is often offered alternative programmes. Second, those not participating in programmes do also have more than one option – not only unemployment, but possibly also options as employment, education, retirement etc. Thus, matching the samples to be compared on *all* these propensities would make the CIA more plausible.

In this paper we use probability scores matching estimators to assess the average impact of the LMT programme by comparing participants (the treatment group) with unemployed non-participants (the comparison or no-treatment group). The two samples are matched by probabilities of (a) taking part in the programme to be evaluated, (b) taking part in other alternative programmes, and (c) leaving the unemployment register, all as alternatives to staying unemployed. To make each group of participants homogenous we conduct separate analyses for those starting LMT at about the same time, i.e. in winter or in autumn each year. This gives us 4 cohorts of participants. We also cut the data by gender and unemployment benefit entitlement, giving a total of 16 subsamples for which we estimate separate training effects.

The analyses show that the overall impact of LMT on annual earnings is *positive*. There is, however, a considerable variation. The annual effects vary over the post-training period, within as well as between the 16 sub-samples. As a general trend effects are increasing over time.

First year effects are negative for most sub-samples. Second and third year effects are in most cases positive and significant, in statistical as well as in economic terms. After four and five years earnings of participants *with recent labour market experience* at the time when the training started are significantly higher than among the (matched) non-participants, and this holds for

both men and women. Among *labour market entrants* effects are more mixed and in many cases not significantly different from zero.

The average individual gain is estimated for each subsample by the accumulated discounted earnings differences over the training year (earnings forgone) and the post-training period of 4 and 5 years. These accumulated returns are then compared with the average direct costs of providing the training, based on average monthly operating costs per participant corrected for the marginal cost of public funds.

The calculations show that participants with recent labour market experience have a substantial accumulated gain in earnings. For women, the estimated gain exceeds the direct costs of training. Actually, even the lower end of the confidence interval for the estimated five-year effect exceeds the costs. The accumulated effect for male participants with recent labour market experience is lower. For a five-year post-training period the effect is very close to the direct costs, while it is lower than the cost when the post-training period is restricted to four years.

The average accumulated gain for labour market entrants is much lower than for experienced participants. For most subsamples of entrants, the accumulated return is not significantly different from zero and thus less than the costs.

Limited information on post-programme outcomes causes practical problems in most of evaluations of long run effects. Even with five years of post-training earnings, we tend to underestimate the benefits of the programme when long run effects are positive. When we project future earnings gain and expand the post-training period from four to seven years the accumulated effects increase for all groups. For experienced workers the accumulated effects clearly exceed the direct costs. The benefit-cost differential becomes economically significant for both men and women. The accumulated effects remain considerably lower for labour market entrants but the effects approach the direct costs of providing the training. Actually, the average effect for men is very close to costs. For women it is lower than the costs, even after seven years.

The remainder of the paper is organised as follows. In section 2 we first discuss possible explanations for why training may affect individual outcomes, both in the short and long run, and then we discuss some recent evaluations of long term effects of labour market programmes in Sweden and US. In section 3 we present the matching estimation model for individual effects. Estimation of social costs and benefits is briefly discussed in section 4. Next we present the LMT programme and the data, and in section 6 the matching procedure and the outcome of the matching. Section 7 presents the results of the evaluation, i.e. the average annual earnings gain for the participants (ATET) for each year in the post-training period. Accumulated long run effects are presented in section 8, and compared with estimated direct costs of operating the pro-

gramme. We also present some results based on projected future earnings gain for an extended post-training period of seven years. Section 9 concludes the paper.

2. Short and long run effects of ALMPs

Individual effects of active labour market programmes are commonly measured as impact on labour market status at a given point in time (e.g. employed or not) or by an outcome flow during a certain time period (e.g. annual earnings or duration of non-employment). Short run effects are naturally defined as those measured "early" in the post-programme period. By long run effects one can either mean effects on outcomes some years ahead or accumulated effects over what is defined as the relevant post-programme period. We will label the latter two *long run* and *accumulated* (long run) effects, respectively.

When effects vary over the post-programme period, evaluation studies face a number of problems. The short run effects are not necessarily representative and they may contain "transitory" components. Then, when the success (or failure) of a programme is evaluated by labour market status at one point in time, the robustness of the estimate is likely to be low. Time-varying programme effects also imply that information about short run effects is insufficient to calculate the accumulated long run effect. This suggests that a longer follow up period will be necessary to provide robust estimates. The need for a follow-up period beyond the typical first year or two years is then crucially dependent on how the effect changes as time passes.

Imagine a simple search framework where programme participation only affects the job offer arrival rate. Compare two unemployed individuals. One starts training, while the other continues to search for jobs. Arrival rates for job offers are likely to drop while on training because of less time available for search. If the training increases qualifications, the post-training job arrival rate is likely to be improved by participation. Assuming jobs do not terminate, the expected earnings profile over time will differ between the two. If the positive shift in the post-training arrival rate is sufficiently large, relative to the negative lock-in effect while on training, the earnings of participants will eventually exceed that of non-participants. Accumulated earnings will typically be lower among participants if measured early in the post-training period and then possibly overtaking the earnings' of the non-participants as the positive training effects dominate the foregone job opportunities. As time goes by and the distance to the training period increases, we would expect convergence in period-earnings between participants and participants. Thus, the length of the post-training period as well as the weighting of gains and losses across time periods (e.g. the discount rate) may turn out to be crucial.

In general, participation in a programme like LMT may affect future employment and earnings for the participants through various channels. First, successful training helps trainees to accumulate human capital that is relevant to potential employers. Increased human capital may have a positive effect on wages as well as the probability of employment. However, if training increases the reservation wage, this may have the opposite effect on the employment probability. Secondly, as training represents a meaningful activity to most participants, it may help to prevent social isolation and mental problems during a period of non-employment. This may in turn enhance job search efficiency and reduce the probability that unemployed workers drop out of the labour force. Thirdly, LMT may represent a signal about unobserved characteristics like motivation and effort, which correlates with productivity. Potential employers may consider a personal unemployment record which include LMT to be better than a record with only open unemployment. This "signalling effect" of LMT is crucially dependent on the reputation of the programme. Programmes associated with long-term or low-qualified unemployed may give a negative signal to employers. Finally, training has an alternative cost as time available to ordinary job search activities is reduced. Various empirical studies (national and international) show that labour programme participants have very low transition rates to ordinary employment during the programme period; see e.g. Røed and Zhang (1999).

Taking part in LMT is also a potential stepping stone to further, ordinary education and thus a delayed labour market entry. Actually, the purpose of some courses in the evaluated programme is to qualify the participants for ordinary education at secondary and high school level, or at college level. In such cases we may observe negative short run effects on annual earnings, while long term effects should be positive.

To sum up, there are several possible sources of persistent differences in post training period-earnings between participants and non-participants. Program participation may (i) prevent labour force withdrawal (discouraged worker effect) and (ii) increase productivity, shift the wage (offer) distribution, and reduce future job separations. An increase in job arrival rates may also (iii) raise the reservation wages, which will contribute to negative short run effects. Finally, participation may also provide positive or negative (iv) stigma effects.

Empirical results - US evaluations

Most evaluations of labour market programmes focus on short run effects, typically employment probabilities or annual earnings one or two years after the programme. In such cases cost benefit analyses of the programmes would have to be based on extrapolations of short-term effects. When period-effects are volatile, extrapolating from short or medium run effects to predict long

run effects can easily be misleading. This is discussed by Couch (1992) and Friedlander and Burtless (1995) and recently also demonstrated by Hotz, Imbens and Klerman (2000).

Stanley, Katz and Kreuger (1998) summarise quantitative evidence from the evaluation literature on the impacts of various US labour market programmes and services on outcomes like employment, earnings and educational achievement. They refer very few analyses including long term effects, except Friedlander and Hamilton (1993) and Friedlander and Burtless (1995). Positive effects on employment and thus on annual earnings are reported in many studies. However, "the gains generally fade out after about five years".[2] Grubb (1999) draws similar conclusions in a survey of training for youth in OECD, and shortage of evidence on long-run effects of ALMPs is one of the main findings of Martin and Grubb (2001) in their review of OECD countries' experiences with active labour market policies.

Couch (1992) presents analyses based on post training earnings records for participants in the National Supported Work (NSW) experiment, which operated from 1975 to 1979. The original analyses of the NSW by Hollister, Kemper and Maynard (1984) were based on self-reported earnings and earnings records 1972-77. This early analyses showed that training in NSW led to sizeable earnings increase for former AFDC recipients in the first post-programme year 1979, but not for youth. To obtain likely long-term effects for the cost benefit analyses Kemper, Long and Thornton (1984) assumed that the first year effect would decay with 3 to 17 per cent per year.

Couch (1992) shows that estimates based on earnings records give positive impact of training in NSW every year in the post training period 1979-1986 for adult AFDC recipients (but not for youth). The estimated effect is significant at either 5 or 10 per cent from 1982 through 1986, increasing from 1979 till 1984 and then decreasing. Thus the first year effect did not decay as assumed by Kemper et al. (1984). The average cumulative increase in real earnings (not discounted) over the post training period is US$ 2,728 – which is slightly more than the average cost of training, US$ 2,674. Thus within this time perspective the internal rate of return is close to zero. Any positive impact on earnings after 1986 will increase the internal rate.

Friedlander and Burtless (1995) analyse long term effects of four experimental welfare-to-work programmes to test their effectiveness: Three Work Incentive programmes (WIN 1982-87) and one Saturation Work Initiative Model (SWIM 1985-88). They use information covering a

[2] In their survey Stanley et al. (1998) find that programmes encouraging additional job search for dislocated workers appear to work well. When it comes to training programmes for the same target group the impact have not been well researched and available results are mixed. Government run programmes are generally not found to be effective. Classroom training in community colleges has been found to have significant positive effects in comparison studies, but may vary considerably based on the courses taken. Disadvantaged adults (especially single mothers seeking to leave welfare) seem to respond well to training programmes. Voluntary training programmes typically produce substantial earnings gains and high benefit-cost-ratio. Studies of long-term effects are few. However, evaluations of the JTPA show that earnings gains were still growing at the end of the thirty months evaluation period.

longer follow-up period than in the original evaluations (20-22 quarters or almost five years) and find that earnings impacts are more than double compared with those observed over the original two-year evaluation period.[3]

For all four programmes the five-year accumulated gains in enrolees' earnings (not discounted) are several times larger than the rather modest programme cost (difference between average cost per experimental and average cost per control). Across the four programmes, accumulated earnings impact is larger for the more costly programmes.[4]

The time profile in earnings differences for experimental-controls shows that the programme impacts remain strong for 3-4 years and declines thereafter. Friedlander and Burtless (1995) comment that it "... seems likely that additions to total impacts after five years will be much smaller than in preceding years ..." They are however reluctant to interpret it as an "impact decay". Given the probable importance of control-group catch-up – possibly assisted by receipt of programme services after the end of the formal program, they describe the narrowing difference as "convergence".

Bloom et al. (1997) present key findings on costs and benefits from the national JTPA study (Job Training Partnership Act title II-A). They compare incremental benefits and costs from three perspectives: programme enrolees, other persons (rest of society) and society as a whole. The main benefit is increased enrolee earnings. The main cost is the employment and training services received by the enrolees.[5]

It turns out that within the 30 months observation period, benefits exceed costs per enrolees for the society as whole for adult men and adult women, but not for male and female youths. The average earnings gain for enrolees is almost US$ 1,700 and 1,400 for adult males and females respectively. Corrected for other costs and benefits for the enrolees as well as for the rest of society, the net benefits for the society as a whole are estimated to at average US$ 524 and 512 respectively.

---

[3] Increased earnings in the post-training period are mainly due to reduced unemployment and longer working hours. The programmes (except one) were, however, not successful in helping the participants in finding better-paying jobs. As the embargo against controls to participate in programmes only was effective in the original two-year follow-up period some of the controls may have participated in year three, four and five. Thus the impact estimates for these years are interpreted as lower bounds of the long-term effects.

[4] Increased employment also means that the enrolees leave the welfare programme earlier than the non-enrolees do. Thus accumulated welfare payments are reduced, but usually not more than earnings are increased. In two of the four programmes welfare savings are large enough to pay back the cost of the programmes. In addition there are savings due to reduced payment of other benefits as well as increased income taxes from enrolees.

[5] With three different perspectives they are able to measure redistribution effects. For instance, reduction in welfare benefits during the program period is a loss (a cost) for the enrolees, but a benefit for others (the taxpayers). Same applies for increased taxes due to increased earnings: this is a cost for the enrolees and a benefit for the rest of the taxpayers. However, as long as the cost of taxation is not considered (as is not done by Bloom et al.) such transfers between groups will not affect the net outcome for the society as a whole.

Among youth even the enrolees have negative net benefits of JTPA during the first 30 months after assignment. For young females the average welfare benefit reduction is larger than the average earnings gain; for young males the earnings gain is negative. Adding the training costs and other costs and benefits gives an average net cost for the society as a whole of US$ 2,900 and 1,200 respectively.

Heckman and Smith (1997) elaborate the cost benefit analyses of the JTPA programme by including welfare costs of raising government tax revenue to finance the programme, by discounting future benefits and by assuming an extended benefit period. Estimates of costs and benefits are essentially the same as presented by Bloom et al. (1997) (but here drawn from Orr et al. 1995). However, Heckman and Smith (1997), break up the time after assignment into 6 months periods. All costs are assumed to be paid in the first period. Benefits, measured by the difference in mean earnings between treatments and controls, are assumed to be received every 6 months period. Present values of benefits minus cost estimates are presented for alternative assumptions on the discount rate (0.000 and 0.025) and welfare cost of taxes (0.0 and 0.5, i.e. it costs an extra US$ 0.5 to raise US$ 1.0).

The analyses clearly illustrate that by discounting future benefits and especially by including welfare costs of raising government tax revenues the returns to the programme are substantially reduced. For instance, including welfare cost of taxes of 0.5 turns the net return for adult women from positive to negative.

The net return is also sensitive to the assumed (and actual) duration of benefits. Heckman and Smith (1997) illustrate this as they assume the benefit period to be 7 years (84 months, or 14 periods) and use the observed average benefits for period 4 and 5 (18-24 months and 25-30 months) to predict the benefits in future periods. In the case of male youth, where average earnings gain is negative, things get worse by extending the benefit period: The negative net revenue increases (gets more negative). In all other cases, where earnings gain are positive – but not always large enough in the 30 months period to net out the programme costs – extending the benefit period from 30 to 84 months helps. Net revenues that were positive for 30 months increase, and net revenues that were negative approach zero.

Similar illustrations of the results' sensitivity with respect to future effects and costs of public funding are found in Heckman, LaLonde and Smith (1999) who re-examine the net social return (including earnings gain, reduced transfers, programme operating costs, forgone earnings etc) of some US experimental labour market programmes. The internal rate of return (IRR) is calculated for alternative assumptions about duration of the earnings impact: from 3 years to

indefinite, the discount rate: from 0.0 to 0.1, and deadweight loss associated with the taxes that finance the programmes: from 0.0 to 1.0.

Hotz, Imbens and Klerman (2000) re-examines some welfare-to-work programmes in California called GAIN, Greater Avenues to INndependence. In a study of these programmes conducted by Manpower Demonstration Research Corporation (MDRC) in the late 1980s and early 1990s, welfare recipients in six California counties were randomly assigned to either a treatment group that was to receive the services offered by the county, or to a control group to which the services were denied (Riccio and Friedlander 1992).[6]

Over the first three years after randomisation the largest effects are found for the Riverside County's programme. Annual employment rates and annual earnings of heads of single adult households (usually a female) participating in the programme were (at average) 39 per cent and 63 per cent respectively higher compared with non-participants (Riccio et al. 1994). Different from the other counties covered by the study, Riverside emphasised a tightly focused job search programme (Job Club) while rather few were offered human capital development programmes. In Los Angeles, where the majority of the participants were offered basic skill education programmes, the average effects over the first three years on annual employment rates and annual earnings of single heads (females) were 7 per cent and 0 per cent respectively (both insignificant). The Riverside programmes make good scores on other outcome variables as well (welfare participation), and for other target groups (two parent households).

Hotz et al. (2000) point at several reasons why the findings from MDRCs evaluation of Riverside's GAIN programme do not necessarily imply that the *work-first* strategy (at average) is more effective than the *human-capital* strategy for increasing the self-sufficiency of participants. First, the estimates of average impacts are based on a post programme period of three years, which is not extremely short, but still not long enough to assess long term effects of the various programmes.[7] Second, short run estimates will typically understate the relative effectiveness of human capital development programmes compared with programmes that emphasise early labour force entry - simply because the first kind of programmes takes longer time to complete. Third, the design of the GAIN evaluation implies that estimated average effects are not comparable between counties. Programme effects may be heterogeneous across individuals (with different human capital) and across locations (with different economic and labour market conditions). As the population at risk (welfare recipients) and thus also the mix of randomly assigned participants

[6] In the experiment the counties were given considerable discretion in the types of welfare recipients to be selected to the random evaluation (all applicants or only long-term recipients) as well as in the way they designed (and mixed) the programmes. Thus the MDRC study was an evaluation of six separate programmes, each with its own distinct population and random-assignment design.

[7] Five years estimates are presented by Freedman et al. (1996).

vary from county to county it is not straightforward to extrapolate the estimated average effects from one site to other.

In their paper Hotz et al. (2000) extend the follow-up period from three to nine years and estimate average effects for three distinct post programme periods: 1-3 years, 4-6 years and 7-9 years after randomisation. They find that for most outcome variables the stronger average impacts for the Riverside County's *work first* programme tend to shrink, whereas the weaker average impacts for the *human capital* programmes in for instance Los Angeles tend to remain constant or even grow over time. Thus they conclude that simple extrapolations of results from early periods to later periods are not warranted. The relation of short-term and long-term results appears to vary with the content of the programme.

The conclusions are, however, bungled by the fact that the embargo against controls to participate in GAIN programmes was lifted at minimum three and maximum five years after the randomisation. In general estimated long-term average effects are expected to decline as members of the control groups receive training. The long-term estimates are also affected by the fact that the mix of programmes in some counties did change after some years.[8]

European evaluations

It seems to be only a handful of European studies that contain long run effects of active labour market programmes (Fay 1996, Martin and Grubb 2001). In Sweden, for instance, the surveys by Björklund (1990, 1993) and by Calmfors, Forslund and Hemström (2002) indicate that very few microeconomic evaluations cover post-treatment periods of more than 2 years, and the number of studies involving cost-benefit-analyses are less.

One example of the latter is Axelsson (1992) who evaluates a Swedish labour market training programme (similar to the programme evaluated in this paper) by comparing annual earnings (before taxes) for a sample of participants and a sample of unemployed non-participants in 1981 (analysed sample, n=2,129). The analyses are based on non-experimental data (with information from registers and questionnaires) within the framework of a log-linear earnings model, correcting for selection by Heckman's λ (two-step estimation), as well as by difference in differences. The overall impact is estimated to be positive and significant, and the second

[8] Hotz et al. (2000) also develop methods to allow for comparison of average effects of programmes implemented in different locations and with different mix of participant populations. These comparisons are based on information on individual characteristics, regression analyses and difference-in-differences estimators. These analyses are less relevant for our paper. However, the results strengthen the conclusions from the simple comparison of short and long term effects: Adjusting for a rich set of pre-randomisation variables reduces and in most cases eliminates the differences in average outcomes between controls in various counties. The adjusted differences in average outcomes for participants (and the adjusted difference-in-differences) are thus interpreted as differences in causal impacts. As Ricco et al. (1994), Hotz et al. (2000) find substantial positive effects of the Riverside programme compared with the programmes of the other counties, including Los Angeles in the first three years of post-randomisation. The differ-

year effect turns out to be larger than the first year effect: about Sek 9,000 and 7,000 respectively (i.e about US$ 860 and 670 or Euro 980 and 760).

In the cost benefit analyses the estimated impact on earnings is assumed to last till retirement but is discounted by four per cent (i.e. the impact is cut by half after 17 years) in addition to the overall discount rate of three per cent. To find the benefits for the society the expected private benefits are corrected according to payroll taxes (about 38-42 per cent of the earnings before taxes), and value added taxes (about 18-19 per cent of the value included the VAT). It turns out that social value of the earning gains is approximately 1.7 times the private value. The main costs are income forgone for the enrolees during the training period corrected for payroll taxes and value added taxes, and programme costs. According to the CBA the average net impact of the programme is positive for the society as whole, varying from Sek 14,700 per participant to Sek 123,000 depending on alternative assumptions.

In a more recent study on Swedish data, Sianesi (2002) investigates the effectiveness of Swedish labour market programmes during the 1990s. About 116,000 individuals aged 18-55 who became unemployed for the first time during 1994 are followed - by register data - until the end of November 1999, yielding a post-training period of up to 5 years. The study looks at the programme-benefit system in its entirety by lumping all kind programmes into one "treatment". The labour market programmes are aimed at equipping job seekers with skills to improve their labour market opportunities. However, at the same time, in Sweden, participation in programmes allows renewing eligibility for quite generous unemployment compensation and thus reinforcing the work disincentive associated the unemployment benefit system.

Sianesi (2002) estimates the effects of the programmes on various outcome variables by matching participants with non-participants – conditional on unemployment duration, assuming selection on observables. Thus the estimated impact relates to how those joining a programme at some point in time - perform on average compared to the hypothetical state were they would have waited longer, i.e. continue job-searching in open unemployment, assisted by the public employment services (PES). The analyses show that unemployed who enter a programme early (compared to later or never) have a significantly higher probability of being employed – from six months after joining the programme till at least five years. At the same time, joining a programme prolongs the current unemployment spell by a couple of months and increases the probability of being in benefit-compensated unemployment over time, of participation in further programmes over time as well as of being in the chain of programmes and spells of compensated unemployment.

---

ences taper off, however. In some cases the differences turn significantly negative in years 4-6 and 7-9 after randomi-

When it comes to the probability of being employed, the impact is negative the first 5 months after joining the programme, but then it turns positive. After 12 months the difference in employment probability between participants and non-participants is about 5 per cent and then quite stable over the rest of the post-training period. For both groups the employment probability is increasing; for participants from about 0.25 to 0.45 over the last four years of the post-training period.

In a companion paper, Sianesi (2001) applies a multiple-treatment-matching framework to evaluate the differential performance of six main types of Swedish labour market programmes. This study covers 30,600 adults 25-54 who became unemployed for the first time during 1994 and who were eligible for unemployment benefits. The sample is followed by register data until the end of November 1999, i.e. a post-training period of maximum 5 years. The differential performance of the six programmes – and the non-treatment state (waiting longer in open unemployment and searching for a job) - is assessed in relation to employment rates over time and the probability to be in a compensated unemployment spell. The analyses confirm the results from Sianesi (2002): On average people who is in a programme (any programme) at a given moment subsequently enjoy higher employment rates than if they had postponed participation. The most successful programme in terms of preceding employment is clearly *employment subsidies*, not surprisingly as this is an arrangement based on a job promise by the programme employer – after completion of the programme. The employment probability is 40 percentage points higher about 7 months after entering the programme - compared with waiting. The impact drops over time and is about 20 percentage points 60 months after entering the programme.

One of the six programmes evaluated is *labour market training*, a programme very similar to the Norwegian LMT-programme. Compared with waiting, participation in LMT is found to have a positive, significant effect on employment, increasing from about 5 percentage points 12 months after entering the programme to almost 20 percentage points 60 months after entering. Compared with the other five programmes LMT is the least effective when it comes to employment rates over time.

Sianesi (2001) also presents cost information, i.e. average monthly cost of each of the six programmes. Employment subsidy is the least expensive programme, while LMT is the most costly programme. Unfortunately, there is no information on the average duration of various programmes and the success criterion (employment probability) is difficult to translate into a benefit measure that can be compared with cost information. Consequently, no formal cost-benefit analyses are made.

---

sation.

3. Estimating individual returns

There are various concepts of causal effects, even for a specific and well-defined treatment and for a given outcome. First, the treatment in question needs to be contrasted to *an alternative* treatment or to non-treatment. Second, we have to specify for *whom* we are evaluating the impact: The average effect for a specific group (all or some of the participants, or the non-participants), the effect for a random member of the eligible population, the effect for the marginal participant (who is the one to be affected by a marginal change in the programme), or the whole distribution of effects for eligible individuals.

Denote Y_1 as the given outcome at the relevant point in time conditional on the specific treatment of interest, and denote Y_0 as the outcome conditional on non-treatment, or the alternative treatment. In our case, Y_1 and Y_0 is the sum of future (discounted) earnings in the case of training and no training, respectively. Note that we measure the earnings flow from the start of the training period onwards, taking into account the alternative costs (or "lock-in" effects) of participation.

Defining the impact as the difference between these two, we get $(Y_1 - Y_0.)$ Thus the causal impact of the treatment does not only rely on the specification of the treatment to be evaluated. The definition of the non-treatment status is just as important.

For each person only one outcome is observed. Thus whether we want to estimate the expected impact for *any* potential participant, for those *not* participating, or for those who *do* participate, we need to estimate or simulate the counterfactual outcome.

Assume we have cross-sectional data. Let D=1 for those in the treatment group and let D=0 for those in the non-treatment group. Let X be a vector of observed characteristics. Assume the outcome Y depends on X and D, as well as an unobserved error term U:

(1a) D=1: $Y_1 = a_1X + U_1$

(1b) D=0: $Y_0 = a_0X + U_0$

In this study we restrict ourselves to consider the mean impact for participants or *average (expected) treatment effect for the treated* (ATET).[9] The ATET is the expected difference between Y_1 and Y_0, conditional on D=1, given by

[9] Other parameters of interest are for instance the average (expected) treatment effect for a person drawn randomly from the eligible population or the expected effect for a person drawn randomly from the combined sample of participants and non-participants. In addition it is of interest to assess the whole distribution of effects: What fraction

(2) $\Delta(X) = E(Y_1 - Y_0 \mid X, D=1) = E(Y_1 \mid X, D=1) - E(Y_0 \mid X, D=1)$

To identify this parameter we have to predict Y_0, because this is not observable for D=1. Given model (1) the effect $\Delta(X)$ defined by (2) is a mix of structural effects $\{a_1 X - a_0 X\}$ and error terms $E(U_1 - U_0 \mid X, D=1)$.

There are many methods of constructing the unobserved counterfactual $E(Y_0 \mid X, D=1)$. One common method is to use the outcomes of non-participants (or participants in the alternative treatment) as a proxy, i.e. $E(Y_0 \mid X, D=0)$. However, comparing participants and non-participants for instance in a standard regression analyses, i.e. comparing the expectations $E(Y_1|X, D=1)$ and $E(Y_0|X, D=0)$, we may get a biased estimate of $\Delta(X)$. This *selection bias* is given by

(3) $B(X) = E(Y_0 \mid X, D=1) - E(Y_0 \mid X, D=0)$

B(X) is rigorously defined only for values of X common to D=1 and D=0. Conditional on this X the bias rigorously defined is due to genuine differences in the distributions of the error terms (unobserved differences).

The *conventional evaluation bias* (LaLonde 1986) defined by $B=E(Y_0|D=1) - E(Y_0|D=0)$ is analogous to selection bias B(X) given by (3) but does not condition on X. Heckman, Ichimura, Smith and Todd (1998) show that the conventional evaluation bias comprises the selection bias rigorously defined as well as bias due to non-overlapping supports of X in the two samples (mismatching) and different distributions of X within the two samples (misweighting). Heckman, Ichimura and Todd (1997) demonstrate that, in the JTPA study, bias due to selection on unobservables is empirically less important than selection due to lack of matching on X for the samples of participants and non-participants.

The idea of matching is to re-establish some of the features characterising experimental data when we actually use non-experimental data. By matching we construct samples of participants and non-participants to ensure that they meet certain conditions related to independence between the outcome (or the effect to be evaluated) and treatment status. The brief presentation to follow leans heavily on Heckman, Ichimura and Todd (1997, 1998).

---

of the participants benefits from the treatment, and what is the effect for those in the left-hand-side tail of the outcome distribution?

Assume that the outcomes (Y_0, Y_1) and the treatment status D are statistical independent conditional on X. (This X-vector may be the same or another than the X-vector in the outcome model.) Thus

(4) $(Y_0, Y_1) \perp\!\!\!\perp D \mid X$

This is equivalent to Prob(D=1| Y_0, Y_1, X) = Prob(D=1|X), which rules out the Roy model of self-selection. In addition, assume that

(5) $0 < P(X) = \text{Prob}(D=1|X) < 1$

By (5) we exclude cases of P(X)=1 and P(X)=0, i.e. persons with X-values that ensure they will always or never receive treatment. Such persons are not possible to match with persons from the other group. According to Rosenbaum and Rubin (1983) condition (4) is the *ignorability condition* for D, while together with (5) it constitutes the *strong* ignorability condition.

Conditions (4) and (5) are, however, stronger than what is necessary to estimate ATET. To identify $E(Y_0|X,D=1)$ it is sufficient to assume

(4') $Y_0 \perp\!\!\!\perp D \mid X$

(5') $P(X) < 1$

(4') is called the conditional independence assumption (CIA). This does not rule out the dependence of D and Y_1. To get an unbiased estimate of ATET it is sufficient with the even weaker assumption: $E(Y_0|X,D=1) = E(Y_0|X,D=0)$

Assume the X-variables that meet the conditions (4') and (5') are identified. Thus by *matching* the two subsamples on these variables we eliminate the bias in the $\Delta(X)$ estimator given by (2), but only the bias due to observables.[10] Provided that the CIA holds, we have B(X) = 0 for the matched samples. If CIA does not hold, other estimation methods may eliminate selection on unobservables. Difference-in-differences will for instance eliminate selection on person specific, time-invariant unobservables.

[10] Matching here means pairing each programme participant with one (or several) non-participants, selected from the population of non-participants (without or with replacement). The pairs are constructed on the bases of identity or similarity in the X variables. The mean impact of the treatment on treated is then estimated by the mean differences in the outcomes of the matched samples.

When the set of matching variables affecting participation status is large, simple stratification matching is hard to handle. Rosenbaum and Rubin (1983) show that if CIA holds, matching the two samples on the *propensity score* P(X) is sufficient to secure unbiased estimates. They show that (for random variables D, Y and X) when Y_0 is independent of D conditional on X, Y_0 is also independent D, conditional on P(X)= Prob(D=1|X). If the propensity score is smaller than one, then $E(Y_0| D=1, P(X)) = E(Y_0| D=0, P(X))$. Thus, if P(X) is known or if it can be parametrically (or semi-parametrically) estimated, we may match the two samples on the univariate propensity score.

Heckman, Ichimura and Todd (1997, 1998) further develop the propensity score matching methods, see also Heckman, Ichimura, Smith and Todd (1998), Imbens (2000) and Lechner (2001a). Empirical implementations of the various estimators are found in some of the same papers as well as in Deheija and Wahba (1998, 1999), Brodaty, Crepon and Fougere (2001), Smith and Todd (2002), Larsson (2000) and Lechner (2001b).

Although increasingly popular, the propensity score matching technique is not necessarily an easy way to obtain non-biased estimates using non-experimental data. For instance, Smith and Todd (2002) find little support for claims by e.g. Deheija and Wahba (1998, 1999), about the effectiveness of these estimators as a method for controlling for selectivity bias. They find that various cross-sectional matching estimators are highly sensitive to the choice of sub-sample and to the variables used to estimate the propensity scores. Smith and Todd (2002) find that difference-in-differences matching estimators may perform better. As an explanation they point at possible problems with the data, for instance that the features (4) and (5) mentioned above, are not achieved.

Of special interest for our study is the extension of the method from a conventional two-state framework to allow for the case with *multiple mutually exclusive states*, developed by Imbens (2000) and Lechner (2001a). Lechner (2001a) presents a *matching protocol*, suggesting a specific algorithm – with some variants - in four steps for estimating the treatment effects. As pointed by Lechner (2001a) this algorithm minimises the bias, but does not give asymptotically efficient estimators since the trade-off between bias and variance is not addressed. See Hirano, Imbens and Ridder (2000) for a discussion of efficiency of estimators based on propensity score matching.

More sophisticated and computer intensive matching estimators - that also control for *un*-observables - are discussed by Heckman, Ichimura and Todd (1998). In the present paper we match the samples on observables only presuming that CIA holds. However, participants and non-participants are sorted by gender, recent labour market experience (e.g. unemployment benefit entitlement) and training period. In each group, non-participants are matched on three pro-

pensity scores within the framework of a multiple choice model – estimated with an extensive set of explanatory variables.

4. Estimating net social returns

The information needed to perform the ideal social welfare analysis of a training programme is paramount. Consider a relatively minor policy change from A to B. Even if distributional concerns are ignored and the social gain is measured by the change in total output, an evaluation of the social return to switching from A to B requires knowledge that is hard to acquire. As Heckman and Smith (1997) discuss in detail, three pieces of information are needed.

First, the gain for those who switch from non-participation to participation because of B rather than A must be identified. This is a "local average treatment effect" (LATE) of Imbens and Angrist (1994). *Second*, B instead of A may involve different outcomes for groups not affected by the change, i.e. the "always takers" and the "never takers". Assume that the policy change is increased program capacity. Then, the "always participants" may be affected by the quality of the training, while the "never participants" can be affected by displacement effects arising from a more intense competition for jobs. *Third*, the evaluator needs to specify the changes in social costs. Again, if the policy change is to expand programme capacity, information about the cost structure is required.

Our study is less ambitious. We compare costs and benefits in a much simpler way.[11] *Benefits* of LMT are measured by the *accumulated individual benefits*, estimated by the discounted sum of average effects on pre-tax annual earnings for the participants, i.e. "average treatment effects on the treated" (ATET) during the post-training period. We cannot identify a LATE because relevant instruments are not available. If the treatment effects are the same for all potential participants, this is a trivial simplification. Since we estimate ATETs separately for rather homogenous groups, one may argue that LATE is not expected to be very different.

[11] Actually our study is more in line with previous (US) studies, including the social return estimates in Heckman et al. (1999) and Heckman and Smith (1997).

(Part of) the alternative cost of training is incorporated since the earnings differential between participants and non-participants during the training period (e.g. earnings forgone) is included in the accumulated individual benefits. This implies that the value of time during periods of unemployment and programme participation – during the training and post-training period - is set to zero. We also ignore unemployment benefits and taxes during the training and post-training period, as this just represents transfers from the government to the private sector. We also ignore the cost of funding the *net change* in these transfers. General equilibrium effects are not included, implicitly assuming that earnings of non-participants are unaffected by the programme.

Costs of LMT are included only as far as the *direct costs of running and financing* the training programme are concerned. Training production requires instructors, teaching facilities and localities, and an administrative structure. These resources are assumed to have alternative use, with a value given by the market prices. The training programme evaluated is financed by the central government and organised by the local public employment service (PES) under the supervision of the Directorate of Labour and the Ministry of Labour. Most of the courses are, however, arranged by external institutions and paid for by the local PES. The costs of LMT include both courses arranged by PES and courses arranged by external institutions.

Cost of funding

Taxation represents a potential efficiency loss. Since the training programme is financed by taxes, the cost of public funding is part of the social costs of running the training programme. The marginal cost of public funds (MCF) in Norway is recently analysed by Holmøy and Strøm (2002). The study is based on Norwegian data from 1992 and the multi-sectoral growth model, MSG-6, which is used for economic forecasting by Statistics Norway. Holmøy and Strøm (2002) estimate the MCF to be about 1.20 (at average), i.e. public funding by taxes adds an extra cost (efficiency loss) amounting to 20 per cent of the project to be financed. This figure rests on various conditions concerning consumer preferences and incentive effects of taxes. The figure seems, however, to be the same whether the extra public services are financed by increased income taxes, pay-roll taxes or value added taxes. A governmental commission in 1997 (NOU 1997:27, NOU 1998:16) also recommended a marginal cost of funding at 1.20.

Public expenditures on LMT – including wages for own employees as well as services (labour costs and materials) delivered by others – will generate earnings and consumption and thus various taxes: pay-roll taxes, income taxes and value added taxes. Thus the need for extra funding is less than the cost of the programme. This kind of corrections is also analysed by Holmøy and Strøm (2002) by model simulations. It turns out that the need for funding is only about 74 per

cent of the costs or the total expenditures (at average). Again the point estimate rests on various assumptions built into the model. The figures presented by Holmøy and Strøm (2002) indicate that *net* marginal cost of funding in Norway is about 0.15 (given by 0.74 x 0.20 ≈ 0.15.)

The tax wedge

As indicated above we use *market prices* and *market wages* to evaluate both costs and benefits. Due to market imperfections, caused by restricted competition or taxes, the gross wage rate do not represent the marginal productivity or the social value of (changes in) employment. For instance, the existence of pay roll taxes and value added taxes would tend to make the social value of increased employment larger than the gross wage.

Careful analyses of the impact of such imperfections go, however, far beyond the scope of this paper. Thus we restrict ourselves to some brief comments on the tax-wedge. Assume no loss of welfare due to less time for leisure when employment increases, and disregard other taxes than pay roll taxes and value added taxes. The relevant tax wedge is then given by $q = (1+t_1)(1+t_2)$, where t_1 is the average level of pay roll taxes paid by the employer, and t_2 is the value-added taxes paid by the consumer. Thus, the gross wage rate or annual earnings should be multiplied with q to give the social benefit of increased employment. Taxes are pretty high in Norway and figures for 1998 are stipulated by Holmøy (2002) to be $t_1=0.14$ and $t_2=0.21$, and thus $q=1.38$. As an average for the period relevant for our analyses (1992-97) these figures are probably too high.

Taxes may also affect the way we measure the costs. The Diamond-Mirrlees theorem states that – under certain conditions – we should use the same prices as private agents (Diamond and Mirrlees 1971). As far as the price includes value added taxes, they should be deducted because private producers do not pay VAT, only consumers. Given the tax system in Norway in the 1990s, the infusion of value added taxes in the direct costs of LMT was probably rather weak. Thus the market price for the courses is probably a good proxy for the social cost.

Our strategy

Without a consistent model for estimating the impact of taxes and market imperfections on the social value of changes in employment (and productivity), we stick to the more common and simple strategy. The benefits of the programme will be based on estimated changes in *un*corrected annual, pre-tax earnings among participants, i.e. individual effects. This will probably underestimate the social benefit of increased employment. When comparing accumulated individual effects with the direct costs of the programme we will use costs *corrected* for the net MCF equal to 1.15. This strategy provides a conservative estimate of the net social effects of LMT.

5. The LMT programme and the data

In the training period, 1992-1993, overall unemployment was relatively high by Norwegian standards. The rate of open unemployment increased from 1.5 per cent of the labour force in 1987 and reached a peak of 5.5 per cent in 1993.[12] In the same period average number of persons engaged in ALMPs increased from 7,000 in 1987 to 57,000 in 1993. In 1993 2.5 per cent of the labour force participated in these programmes. As most ALMPs last for less than half a year, the total number of persons participating in programmes *during* one year is about twice the participation rate at a point in time.

During the post-training period from 1993 to 1997, unemployment decreased from 5.5 to 3.3 per cent. In 1998 unemployment was 2.4 per cent of the labour force. At the same time the average participation in ALMPs decreased from 57,000 to 23,000. In 1999 the number of participants was as low as 8,000.

The *Labour Market Training* programme (LMT or *AMO-kurs* in Norwegian) is by far the largest programme, covering about 40 per cent of all ALMP-participants. The aim of LMT is to preserve and improve the skills of the unemployed and thereby to enhance their employability. The programme is organised as off-the job courses, mainly targeted at unemployed adults. Moreover, a substantial number of people (re-)enter the labour market via the training programme.

In 1993 the average number of participants in LMT reached a peak of 23,000, before it started to decrease: in 1996 the average number of participants was 14,000, in 1999 only 4,500. The programme is financed by the central government and organised by the local employment service under the supervision of the Directorate of Labour and the Ministry of Labour. The courses are provided by the employment service. Technically, most of the courses are arranged by external institutions. Vocational training is dominant and a wide range of subjects and crafts are covered. Most of the courses are short, from 5 to 20 weeks. In some cases there are basic courses and follow-up courses within the same subject, with a total duration of one year or more. Our data based on training records 1991-1996 shows that average duration of participation is 20-25 weeks per year among those who participate. Thus many participate in more than one course during the year.

[12] Persons registered as full-time unemployed and searching for a job define open unemployment. These figures deviate from the statistics published by OECD, which are based on the Norwegian Labour Force Survey.

LMT is available for all job seekers and participation is voluntary.[13] Unemployed persons who refuse to accept offers of training may lose their unemployment insurance benefit. This sanction is, however, rarely carried out.

The courses are free of charge. All participants get a training allowance. Those who are entitled to unemployment insurance benefits may opt to collect their benefits, as this is more than the allowance for most of those eligible for UB.[14] The unemployment benefits compensate about 62.4 per cent of previous earnings, after tax, while the training allowance is fixed and flat rated.

In general the incentives to take part in ALMPs may depend on how programmes are related to the eligibility and exhaustion of unemployment benefits. In Norway, time spent in LMT and the allowances collected *do not* qualify the participants for unemployment insurance. According to the Norwegian system eligibility for UB depends on recent job experiences. To qualify for unemployment insurance benefits it is necessary to have some minimum earnings during the last year or the last three years.[15] As unemployed not eligible for UB get the training allowance, they sure have economic incentives to take part in LMT.

During our training period, 1992-93, the capacity of the programme and of most courses was limited. The rate of rationing at each course depends on the number of qualified applicants related to the capacity of the course. Thus the recruitment to LMT is partly a self-selection process and partly an administrative selection process. Previous studies of the training programme in Norway indicate a positive selection with regard to both observed and unobserved characteristics (see for instance Raaum and Torp 2002). When it comes to selection on observables, we believe this is taken care of in this study by the matching procedure, the separate analyses of subsamples, the use of register data, and the large number of explanatory variables. When it comes to unobservables, no practical and rigorous test is available. Without experimental data it is not possible to be conclusive neither on the existence nor on the non-existence of this kind of bias.

Data and design of study

The data are drawn from a large *Frisch Centre Database* containing individual level information from numerous administrative registers, delivered by *Statistics Norway*. Our samples are drawn from the population of all entrants (and re-entrants) in the public unemployment register during

[13] For some courses applicants have to qualify through education, previous vocational training or work experience to be eligible.

[14] In Norway membership in the public social insurance scheme is compulsory for all residents. To be entitled to UB it is necessary to register at public employment service, to search actively for a job and to accept any job offer or programme presented by the local PES. In addition annual earnings last year have to exceed some minimum requirements. At present (2002) these are Nok 64,200 for 78 weeks of unemployment benefits, and Nok 102,360 for 156 weeks of unemployment benefits.

[15] Until 1997 (i.e. in our period) earnings received during a temporary employment programme (but not in a training programme as LMT), would qualify for unemployment insurance benefits.

December 1991 – July 1993. This register contains monthly observations of unemployment, labour market programme participation by type, and unemployment benefit entitlement. From this population we select 4 cohorts of LMT participants, entering the programme in four different periods: January-February 1992 (Winter 92), August-September 1992 (Autumn 92), January-February 1993 (Winter 93), and August September 1993 (Autumn 93).

We use tax register information on annual labour earnings 1992-1997, measured by Norwegian *kroner* (Nok) in 1997-values to estimate the impact of the programme. Thus we have information on earnings in the year of training for all, and on post-training earnings in five years for those who enter LMT in 1992, and on post-training earnings in four years for those entering the programme in 1993. This enables us to estimate long run effects of the programme. See Raaum, Torp and Zhang (2002a) for more details on the data.

Participants and non-participants

LMT courses typically start in January and February and then there is another wave of courses starting after the summer holiday in August or September. The winter courses are similar to those offered in the autumn. Most courses in the autumn are completed by the end of the year, but in some cases continuation courses start early next year.

Since the post-programme success of the training is measured by annual earnings, and the time passed after having completed the training may affect the impact on earnings, it is preferable to analyse the impact of autumn and winter courses separately. Each of the four cohorts is also split by gender and unemployment benefits entitlement. This gives us a total of 16 sub-samples. We restrict ourselves to participants aged 25-50, since selection into other programmes and ordinary education, as well as labour market behaviour in general, are different for teen-agers and young adults. The upper age limit set is to avoid transitions out of the labour force due to early retirement or disability pension which become increasingly important as we include unemployed in their fifties and sixties.

To evaluate the impact of LMT we use a matched sample of unemployed non-participants as a comparison group for each of the 16 sub-samples. The participants and non-participants are defined by the same procedure across cohorts and groups. The population of potential LMT participants consists of all full-time unemployed persons registered at the end of December and July, for the winter and autumn cohorts respectively.[16] Then we consider the register status two months later, i.e. at the end of February and September, respectively. LMT participants constitute the treatment group. In order to define a suitable comparison group we divide

non-participants into three groups according to their status in the register: still unemployed (U), participating in another labour market programme (PROG), or having left the unemployment register (OUT). Those who leave enter jobs or exit from the labour force, but we cannot distinguish between the two transitions.

The comparison group is selected among those still unemployed (i.e. status U at the end of February or September). From these individuals we select non-participants who are "observationally equivalent" to the participants, as far as pre-training characteristics are concerned. The logic behind this matching, how it is implemented and the results of procedure are described in section 6 below.

In Table 1 we report the sample sizes – before matching - of the different cohorts, by group. The sum of the four columns marked U, OUT, PROG and LMT constitutes the populations at risk, defined as the members of the stock of unemployed two months before. The transition columns show how they are distributed according to LMT, PROG and OUT transitions. The U-group consists of those still unemployed. The samples of participants in LMT vary between 700 and 2,500 individuals. The last column shows the fraction of the original sample that enters

Table 1. Sample sizes (before matching) and transitions from full-time unemployment

Cohort	UB	Gen	No transition	Transition from U to			
			U	OUT	PROG	LMT	LMT-rate
			(1)	(2)	(3)	(4)	(4)/[(1)+(2)+(3)+(4)]
1 W 92	Yes	M	19,699	5,806	989	1,558	0.056
2 A 92	Yes	M	19,144	8,643	1,612	2,491	0.078
3 W 93	Yes	M	21,442	5,374	1,362	1,900	0.063
4 A 93	Yes	M	18,157	8,842	2,379	2,206	0.070
1 W 92	Yes	F	11,298	3,850	509	988	0.059
2 A 92	Yes	F	13,368	8,538	1,151	2,110	0.084
3 W 93	Yes	F	12,831	4,075	906	963	0.051
4 A 93	Yes	F	13,014	8,872	1,739	1,752	0.069
1 W 92	No	M	5,422	2,310	463	733	0.082
2 A 92	No	M	5,427	3,384	575	1,132	0.108
3 W 93	No	M	6,065	2,461	385	807	0.083
4 A 93	No	M	6,828	3,562	740	1,127	0.092
1 W 92	No	F	3,429	1,792	345	760	0.120
2 A 92	No	F	3,901	2,860	518	1,705	0.190
3 W 93	No	F	3,928	1,923	339	776	0.111
4 A 93	No	F	4,756	3,171	782	1,561	0.152
All			168,709	75,463	14,794	22,569	0,080

Notes. Cohort: W = winter, A= autumn. UB eligibility for unemployment benefits at the time of registering (Yes, No). Gender: M=male, F=female.

[16] By this sample restriction we exclude LMT participants who enter training directly from outside the unemployment register.

LMT. As we can see, unemployed who are not to entitled unemployment benefits are more likely to enter training. Among those with UB, men and women are equally likely to participate in training. For those without UB, more women than men enter LMT.

Earnings profiles of participants

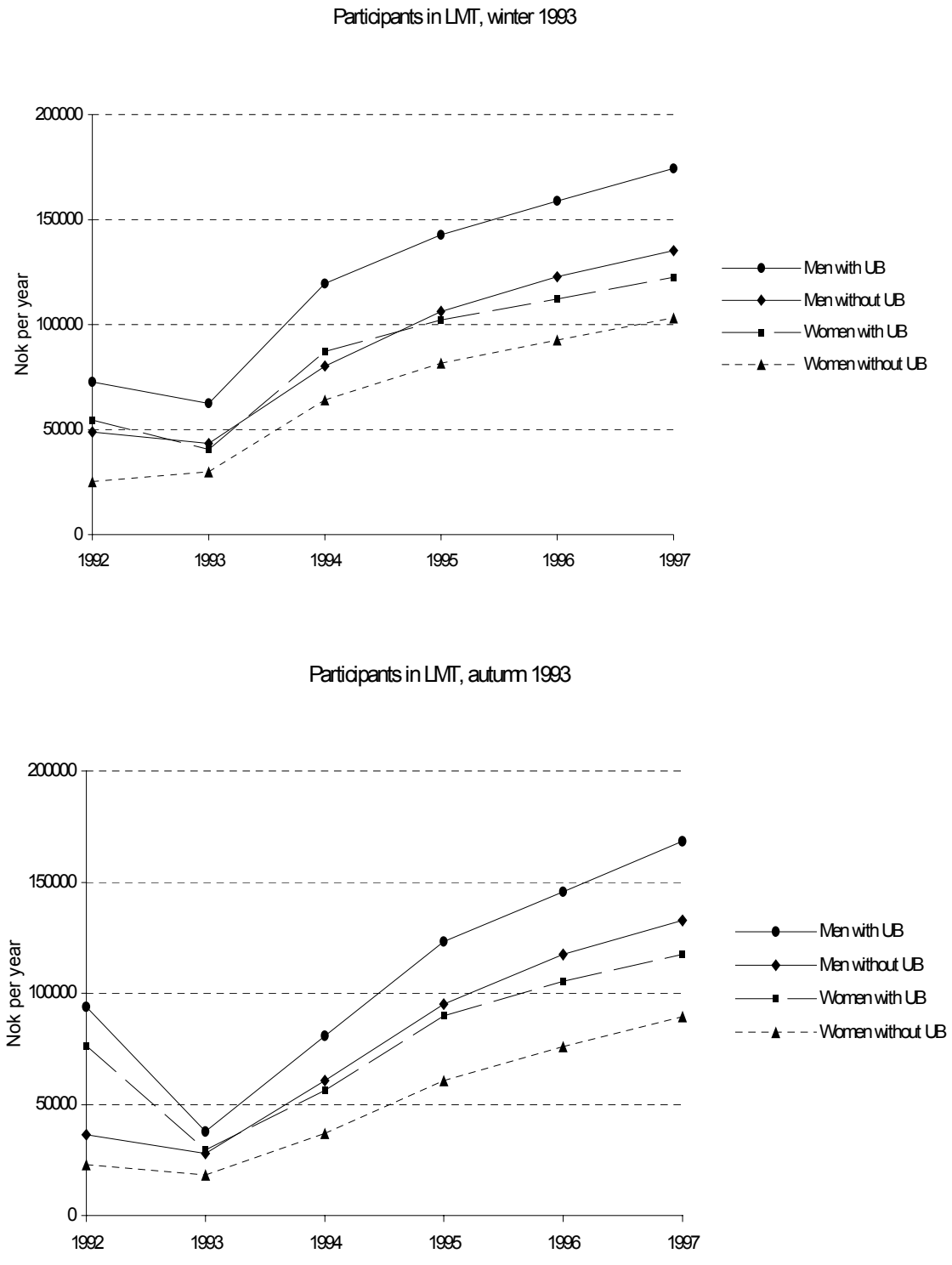
According to the Norwegian labour market authorities, the main objective of LMT is to help unemployed get stable jobs. Even if employment is the overall goal of the programme, several arguments favour the use of *post-training earnings* to measure programme effects. The first argument is *relevance*. Post-training earnings in year t (Y_t) can be decomposed into days of employment (e_t), average hours per day employed (h_t) and average wages per hour (w_t), which gives: $Y_t = e_t h_t w_t$. Here e_t measures how quickly the person enters employment as well as the stability of the job. h_t depends on opportunities, qualifications and preferences of the individual. Part-time unemployment is common among LMT applicants, indicating that many are rationed with respect to working hours. If the training effect on earnings is due to longer daily working hours, this should be considered as a success in line with (re-)employment. The hourly wage reflects productivity and the quality of the employment match. If LMT contributes to more productive employees and a better matching, these effects are obviously socially beneficial. As the Norwegian wage structure is fairly compressed, see e.g. Barth and Zweimüller (1994) and Stewart (1996), earnings mainly reflect the duration of employment. If there is a positive effect of training on earnings we do not expect wage increments to be an important explanation. Finally, cost-benefit comparisons also favour earnings as a measure to evaluate the programme effect.

In line with the objective of LMT, earnings should include wages as well as income from self-employment. Transfers, unemployment benefits, social support and training allowances ought to be excluded. Our data on earnings are collected from public tax- and wage-registers. Unemployment benefits are subtracted, implying that our earnings are very close to income from work, including earnings of self-employed (see Appendix).

Since Ashenfelter (1978), studies of programme effects have been concerned about the earnings dynamics. Participants typically experience that earnings fall prior to the training period and gradually increase in the post-programme period. Figure 1 (two panels) illustrates the mean earnings profiles of the participants during winter and autumn 1993, by gender and unemployment benefits eligibility. The "Ashenfelter-dip" is clearly experienced by those eligible for UB – although we have only one pre-training observation.[17]

[17] In a companion paper, Raaum, Torp and Zhang (2002b), this is more strikingly illustrated by information on annual earnings for a longer pre-training period, up to 4 years. As our database does not include information on earn-

Figure 1. Earnings profiles. Participants in LMT by gender and unemployment benefits entitlement. Annual earnings, Norwegian kroner, 1997-values



ings before 1992, it is not possible to produce similar profiles with pre-training earnings for cohorts 1 and 2, participating in LMT in 1992.

The earnings profiles of the groups without unemployment benefits are different, illustrating that LMT seems to be a "stepping stone" in the process of entering the labour market. One might also suspect that financial incentives (i.e. training allowances) make it economically wise to spend time on LMT during this process, even if the effects on future labour market prospects from this investment is minor. Anyhow, Figure 1 clearly motivates our split by gender and unemployment benefits eligibility when we estimate earnings effects of LMT.

Assessment of data quality

In the jungle of complicated econometric evaluation models, it is important to keep in mind one of the fundamentals in empirical research; "Good data help a lot".[18] Based on evaluations of evaluation strategies, using US data, there seems to be a consensus that some features are of particular importance. Heckman, Ichimura and Todd (1997) summarise these as follows:

(I) Participants and controls have the same distributions of unobserved attributes. (II) Participants and controls have the same distributions of observed attributes. (III) The same questionnaire is administrated to both groups, so outcomes and characteristics are measured in the same way. (IV) Participants and controls are placed in a common economic environment.

In the present study of LMT the treatment and the comparison groups are sampled from the same populations: unemployed registered at the local branch of PES, at the same time, i.e. taking care of (IV). Information on all groups is collected in the same way and from the same sources without sample attrition (administrative registers), i.e. fulfilling (III). The matching procedure described in the next section takes care of feature (II).

6. Selection on observables and matching

This section describes how the comparison group of non-participants is established and used to simulate the counterfactual outcome of LMT participants, i.e. $E(Y_0 \mid X, D=1)$. There are various matching techniques and matching estimators used in the evaluation lit-

[18] This has indeed been stressed by e.g. Heckman and his colleagues in numerous contributions over the last ten years.

erature. In this study we apply a variant of traditional pair-wise nearest-neighbour-matching in the case of a multinomial choice model, inspired by the *matching protocol* suggested by Lechner (2001a, 2001b).

We start out with the population of all full-time unemployed, registered at time *t,* and eligible for the programme to be evaluated. Each member of the population has several options, here specified as four mutually exclusive states: to remain unemployed (U), to take part in the programme to be evaluated (LMT), to take part in another programme (PROG), or to leave the unemployment register (OUT). We specify the complete choice problem in one multinomial logit model (the structural form approach) to estimate the observed state at time $t+dt$, assumed to depend on observed individual characteristics.

The estimated parameters from the multinomial logit model are used to predict the probabilities of LMT, PROG and OUT for each individual in the subsample of participants (LMT=1) as well as for all those still potential participants when the programme starts, i.e. unemployed non-participants (U=1). As we have 4 cohorts split by gender and unemployment benefit status, we have to estimate and predict probabilities for 16 separate samples.

The transition probabilities are estimated as functions of a large number of individual characteristics, separately for each of the 16 sub-samples. From the unemployment register we collect individual information from the pre-training period on

- labour market program participation,
- unemployment experience
- previous occupation and
- unemployment benefits entitlement.

In addition we have register information on

- age, gender, material status and number of children,
- educational attainment,
- work experience (by yearly pension, proportional to annual earnings),
- immigrant status, and
- school enrolment during the previous six months.

The county of residence at the time when training starts is used to control for variations in local labour market conditions and supply of labour market programmes. All these variables are used in modelling the transitions from unemployment. See Appendix for definitions and detailed information.

To eliminate as much as possible of the bias due to selection we select unemployed non-participants with the same predicted structure of transition probabilities as those in the treatment group.

The first step in the matching procedure is to exclude observations outside the common support, i.e. we exclude observations from the sample of participants with estimated probabilities that are larger than the maximum value of the same probabilities in the comparison group. Similar we exclude observations from the LMT-sample with estimated probabilities that are smaller than the minimum value of the same probabilities in the comparison group. Then we use the same procedure to exclude observations in the comparison group with estimated probabilities outside the range of the probabilities in the LMT-sample.[19] This defines the common support samples.

Next we pick one observation from the sample of participants and search through the comparison group to find the closest match based on the three estimated probabilities. In this process we use the *Mahalanobis metric* as a measure of distance with the inverse covariance matrix from the original gross sample as weights; see Rubin (1979). This is repeated until all observations of participants are matched. In line with the matching protocol suggested by Lechner (2001a, 2001b) we sample from the comparison group *with replacement*.

This procedure has the potential problem that a small number of observations from the comparison group is used heavily although other, very similar observations are available – leading to an unnecessary inflation of variance. Repeated use of observations turns out to be negligible in our case. Only 2,815 observations in the comparison group are used more than once, and 655 are used more than twice. That is only 1.7 and 0.4 per cent of the common support sample of non-participants (n=166,564), and 12.5 and 2.9 per cent of the matched sample of non-participants (n=22,500). Thus we have chosen to neglect the fact of repeated use in estimating the standard errors of the estimators.

What explains participation?

The observables used to estimate the propensity scores are defined in the Appendix of this paper. Mean values of selected explanatory variables for subsamples from cohort Autumn 1992 are presented in Table A1, Appendix. In a separate working paper, Raaum, Torp and Zhang (2002a), we report the estimates of the multinomial logit model for selected cohorts, women and men,

[19] We compare one probability at the time: First we accept all observations from the comparison group with estimated values Prob(LMT=1| X) within the range of estimated values of Prob(LMT=1|X) for the participant group. Next we accept all observations from the treatment group with estimated values Prob(LMT=1| X) within the range of estimated values of Prob(LMT=1|X) for the comparison group. Then we proceed with similar comparisons of estimated values Prob(OUT=1|X) and Prob(PROG=1|X) for both samples.

with and without unemployment benefits. The estimations show that various explanatory variables have some influence on the transitions from unemployment to the three other states. The partial impact of most variables differs, however, across subsamples.

When it comes to the relative probability of LMT, there are some common results of interest. First of all, the probability is higher for those who participated in LMT the previous quarter as well (*LMT8* = 1) – ceteris paribus. This parameter is significantly positive for most of the 16 subsamples. Next, those with a pretty long unemployment record (11 months or more, *Months11 – Moths1923*) are less probable to participate in LMT (relative to U). We also find that in a majority of the subsamples the relative probability of LMT is larger for immigrants than others. This may mirror the fact that LMT includes special courses target at unemployed immigrants.

The partial impact of age seems to be of little importance (ceteris paribus), even if those aged 46-50 years (*Age5*) are less apt to participate in LMT for some subsamples. Education is somewhat more important, as low education (10 years or less) and unknown education (*Educ1, Educ2 and Educ6*) is negatively correlated with the relative probability of LMT.

There are quite large regional differences. For many of the subsamples the relative probability of LMT is larger in the northern counties of Norway (Finmark, Troms, and Nordland) than in the southern and central parts. This illustrates the importance of comparing participants and non-participants from the same region if we are to eliminate the misweighting on observables affecting the propensity scores.

Matching results

The success of the matching procedure can be assessed in different ways. First, we compare the distributions of the predicted probabilities among (i) the participants, (ii) all unemployed non-participants (potential members of the comparison group) and (iii) the unemployed non-participants picked by the matching procedure. As illustrated in Table 1 the number of observations in the original sample of unemployed non-participants is much larger than the number of observations in the original sample of participants. This holds for all 16 subsamples. This simplifies the matching. The common support criteria (based on the predicted probabilities) leaves out rather few observations. Across all cohorts and subsamples only 69 of the original sample of 20,569 participants in LMT (0.3 per cent) do not meet the common support criteria. Corresponding figures for unemployed non-participants are 2,145 of 168,709 (1.3 per cent).

Figures 2 and 3 (each with 12 panels) present plots of the predicted probabilities of Prob(LMT=1), Prob(OUT=1), Prob(PROG=1) for two cohorts, winter and autumn 1993. The

plots are estimates of kernel densities on the predicted probabilities.[20] The first panel in Figure 2 presents the predicted values of Prob(LMT=1) for men with unemployment benefits, cohort Winter 1993. The second panel is for men without unemployment benefits. The next two panels present the predicted values for Prob(OUT=1) for the same two subsamples, and then similar panels for Prob(PROG=1). The last six panels in Figure 2 present similar predicted probabilities for the two subsamples of women, cohort Winter 1993. In each panel there are three lines. The line marked with *squares* is for all unemployed non-participants (the original sample). The two other lines are for the matched samples, non-participants marked with *triangles* and participants marked with *circles*.

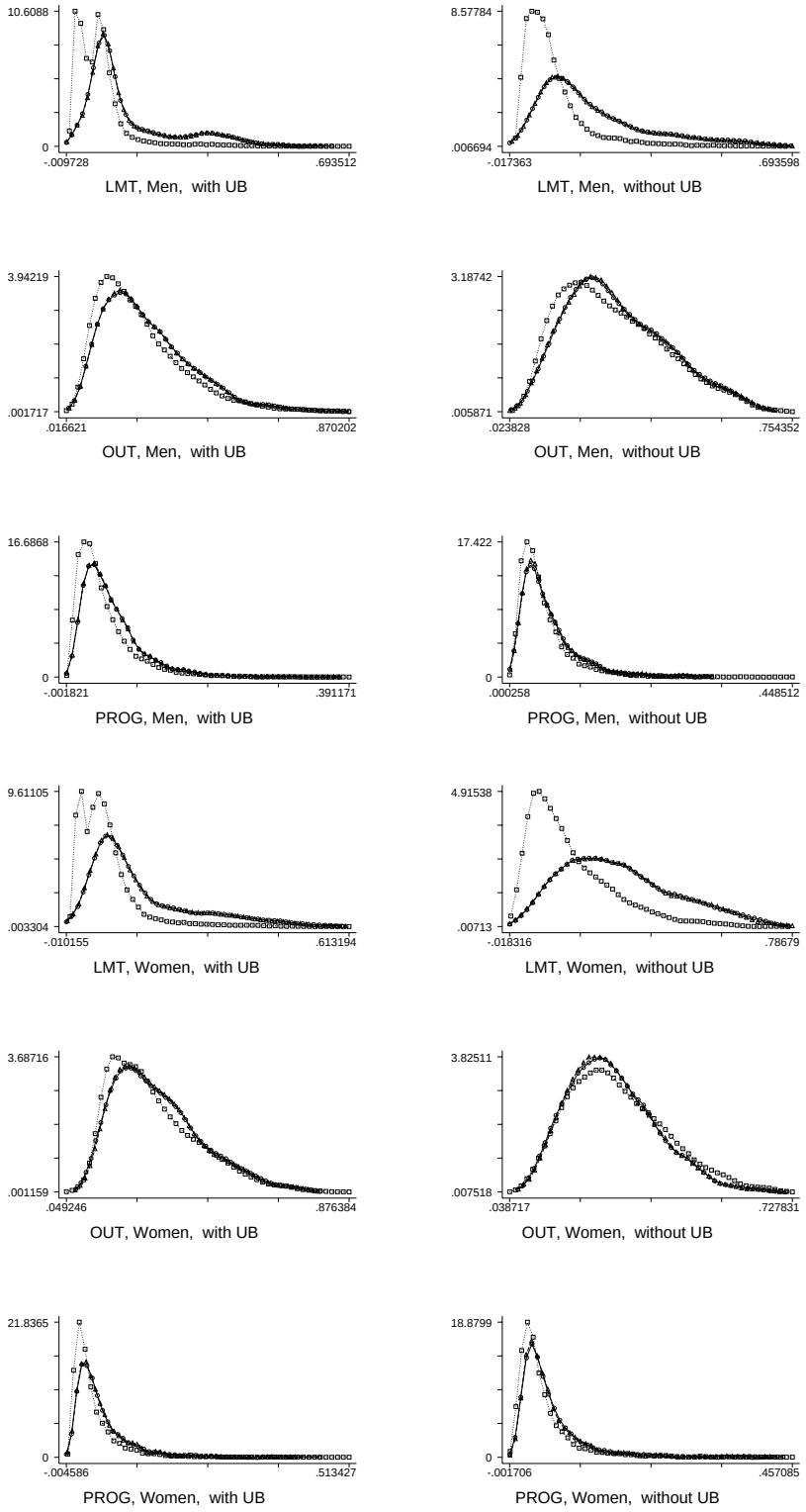
For the predicted values of Prob(LMT=1) the line with *squares* (i.e. for all unemployed) is (in general) to the left of the two other lines. The thicker left-side tail indicates more people (among all unemployed) with a low probability of LMT=1. When it comes to the predicted values of Prob(OUT=1) and Prob(PROG=1) the difference between the three lines is not as large. As can be seen the two lines for the matched samples are pretty close for all three outcomes. The closer the lines, the more successful is the matching with respect to the propensity scores. Figure 3 present similar plots of the predicted probabilities for all subsamples in cohort autumn 1993.

A complete description of mean propensity scores is displayed in Table A2, Appendix. Generally the average predicted probabilities are very close for participants and the matched non-participants. Comparing the mean values of the predicted probability of taking part in LMT, Prob(LMT=1), for participants and non-participants we typically find differences less than 0.1 per cent. Similar small differences are found for the predicted Prob(PROG=1) and Prob(OUT=1.)

The matching outcome can also be assessed by studying differences in pre-training variables. While Table A1-1 (Appendix) presents mean values *before* matching for the autumn 1992 cohort, Table A1-2 (Appendix) presents mean values *after* matching. By comparing the two we see that the matching process has considerable impact on the distribution of the characteristics of the comparison group. While participants are younger, more educated, have a shorter unemployment record, are more likely to have spent the pre-training period in schools or LMT and have a larger fraction of immigrants, these observed differences are more or less removed by the matching.

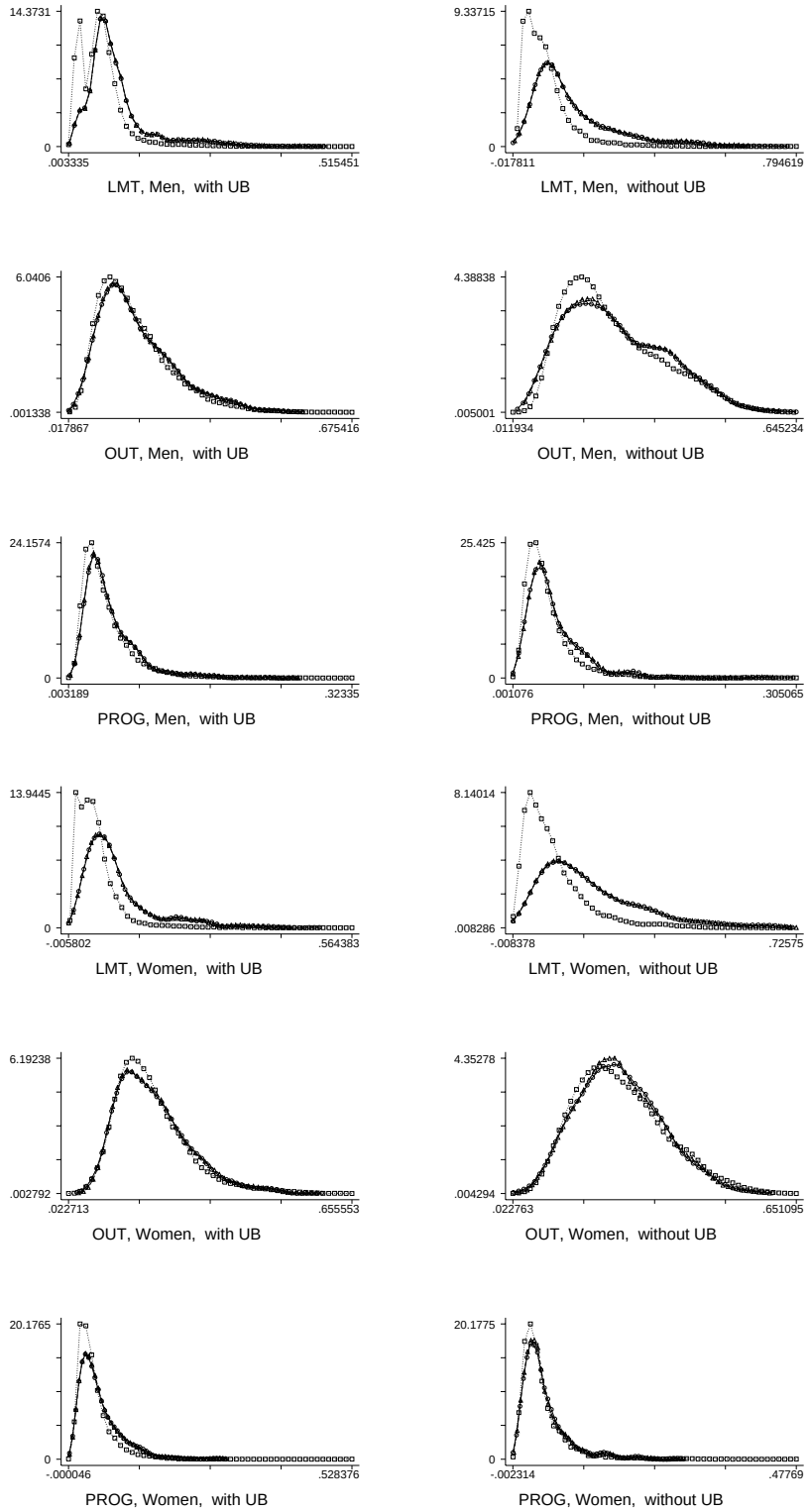
[20] Plots are estimates of *Epanevhnikov Kernel* densities on predicted probabilities P_i(I=LMT,OUT). Bandwidth is estimated by $h=0.9m/(n^{1/5})$, where m=min(sqrt(variance(p_i), interquartilerange(p_i)). The densities are estimated with *STATA,* see "Reference Manual, [R] kdensity" (2001), *Stata Statistical Software, Release 7.0*, StataCorp.

Figure 2. Predicted probability distributions, cohort winter 1993. All subsamples



Explanation:
——0—— LMT participants, --- Δ --- Matched non-participants, ...□.....All unemployed

Figure 3. Predicted probability distributions, cohort autumn 1993. All subsamples



Explanation:
——0—— LMT participants, --- Δ --- Matched non-participants, ...□.....All unemployed

Crossing over and substitution

Participation in LMT is defined according to training status in the unemployment register by the end of February (Winter) and September (Autumn). Most courses start in the beginning of the term, but some start later. Non-participants are not excluded by the programme regulations, nor by our matching procedure, to start training later, either in the same term or in the next terms. If members of the comparison group start in LMT the same term, they are characterised as *cross-overs*. If they enrol during the following terms, we label it *substitution*.

As reported in Table 2, the average fraction of cross-overs is modest; about 6 to 9 per cent of the non-participants enrolled during (i.e. later in) the training term. Between 12 and 17 per cent of the non-participants turned up as participants during the following terms. Participants, however, are much more likely to be enrolled in the following two terms, see Table 2. This is partly due to long-lasting courses stretching into the next term and partly due to follow-up courses. There is now strong indication of inter-temporal substitution in the sense that participation is just delayed for a substantial fraction of the non-participants.

Table 2. Participation in LMT. Rates of cross-over and substitution by gender and unemployment benefits eligibility. Average across all cohorts, winter and autumn, 1992 and 1993

	Cross over	Substitution ("delay")	
Participation period:	Same term	Next term	Two terms later
Males with UB			
Participants	1	0.5129	0.2521
Non-participants	0.0729	0.1425	0.1294
Females with UB			
Participants	1	0.5391	0.2827
Non-participants	0.0870	0.1703	0.1573
Males without UB			
Participants	1	0.5376	0.2220
Non-participants	0.0672	0.1248	0.1184
Females without UB			
Participants	1	0.6052	0.2825
Non-participants	0.0889	0.1693	0.1566

7. Average annual earnings effects

The effects on annual earnings are estimated group wise (i.e. by gender, cohort and unemployment benefits entitlement) and separately for each of the post-training years. All effects are *average training effect on the trained*, simply defined as the mean earnings differential between participants and matched non-participants. Figures 4 and 5 present the estimated effects from year 0 to year 5 (or 4). Table A3 in Appendix displays exact numbers and standard errors.

The effect in year 0 is an estimate of foregone earnings during the training year. In most of the 16 sub-samples, the loss is moderate and varies between zero and about Nok 7,000 (i.e. about US$ 790 or Euro 900). On average, the training period loss is larger for men than for women. The loss is statistically significant for 6 of the 16 subsamples. In one case the training year effect is significantly positive.

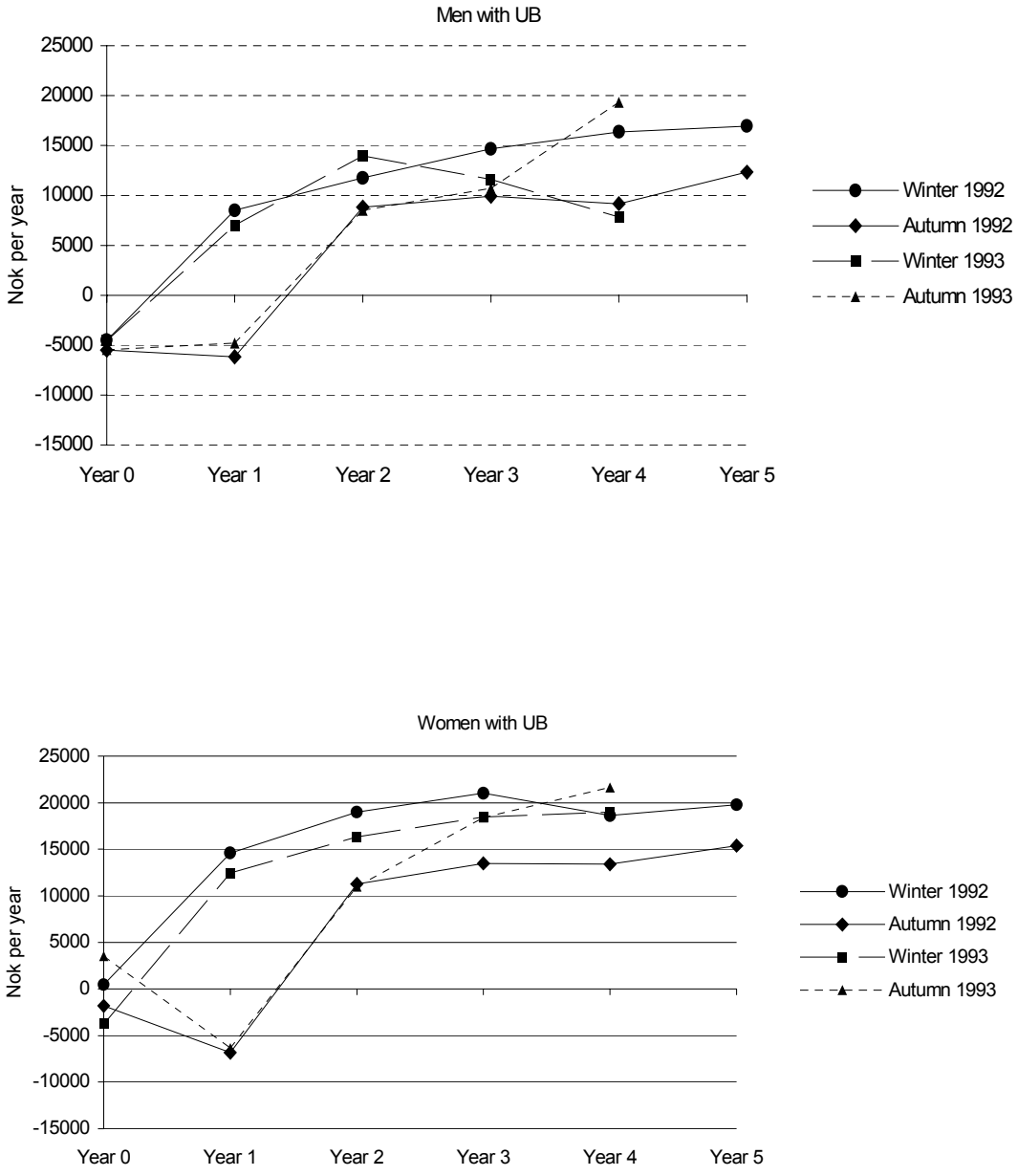
When it comes to post-training effects, there is a considerable variation in annual effects from Nok 0 to Nok 20,000 (about US$ 2,250 or Euro 2,550). The effect varies within, as well as between, groups. The effects are typically increasing over time. Short run (= first year) effects are negative for 6 of 8 autumn sub-samples and positive for 5 of 8 winter sub-samples (see Appendix, Table A3). The negative effects for the autumn sub-samples are likely to reflect the short time span between the training and the outcome periods, amplified by the continuation of training into the next calendar year among autumn course participants. Medium term (= second and third year) effects are positive and statistically significant in 22 of the 32 sub-samples. Consequently, most of the negative short run effects for the autumn courses turn positive as we look further into the post-training period. For winter courses effects are significantly different from zero, in statistical as well as in economic terms, in 11 of 16 cases. The variation across groups, within a given training year, is reduced compared to the short run effects.

The positive training effects are indeed long lasting. Even after four and five years, earnings of participants *entitled to UB*, i.e. those with recent labour market experience, are significantly higher than among the non-participants. As for the short- and medium run effects, the effects are more mixed among participants *not entitled* to UB, i.e. labour market entrants. Long run effects (fourth and fifth year) are positive and statistically significant in 7 of 12 cases. In most cases the estimated average effects increase over time.

Our training effects are based on non-experimental data, without any embargo against the "controls". Members of the comparison group may participate in LMT and other ALMPs during the post-training period (as may members of the treatment group). Without any embargo, members of the comparison group are supposed to catch-up with the treatment group. From this perspective the estimated effects may be interpreted as *lower bounds* of the true effects. As illustrated in Table 2 only about 8 per cent of the comparison group participate in LMT during the training period. Thus the cross-over argument is rather weak. However, also seen from Table 2, participants tend to be *more* inclined to participate in LMT than the non-participants during the post-training period. More than 50 percent of the participants take part in LMT at least in two succeeding terms. This may imply that *short run* effects on earnings of LMT are possibly *under*estimated. Secondly, the estimated *long run* effects should not be interpreted as the effect of taking

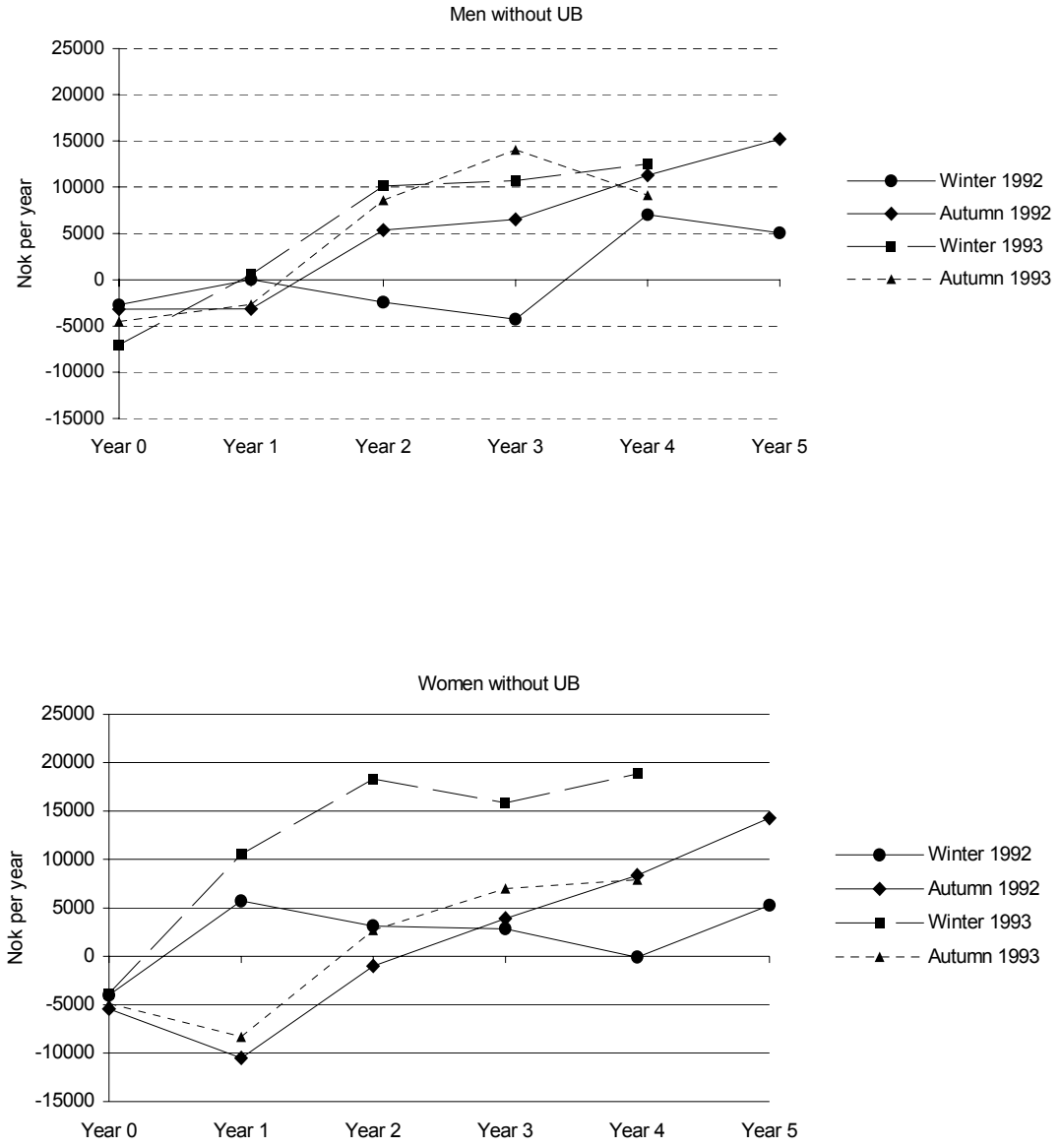
part in LMT for just one term, but more generally as the average effect of participating in LMT of average duration.

Figure 4. Short run and long run effects of LMT on annual earnings by cohort, men and women with unemployment benefits entitlement. Norwegian kroner, 1997-values



Note: Year 0 = training year. Nok 100 = Euro 12.8 = US$ 11.2.

Figure 5. Short run and long run effects of LMT on annual earnings by cohort, men and women without unemployment benefits entitlement. Norwegian kroner, 1997-values



Note: Year 0 = training year. Nok 100 = Euro 12.8 = US$ 11.2.

Irrespective of these modifications, some may find the estimated training effects suspiciously large, given that the average length of the training period is not more than 4-7 months during a calendar year. As mentioned, the annual gains vary a great deal across years and groups and the standard errors are rather large. In year 3, 4 and 5 of the post training period the estimated aver-

age gain represents, however, as much as 10-15 per cent of the earnings without training for some groups. This is pretty much compared with the percentage increase in earnings from an additional year of formal schooling. In Norway this wage premium is estimated by Barth and Røed (1999) and Raaum and Aabø (1999) to be about seven per cent.

However, it is important to keep in mind that, unlike the standard marginal effects of schooling, our estimated training effects do *not* measure the impact on *permanent earnings* per unit of labour (hourly wages). Rather, the main factors behind earnings development in this transition period will likely be the time it takes to become employed and the stability of the new (and future) job(s). We expect the effects of LMT on permanent earnings to be considerably lower, as the *really* long run differences between participants and non-participants will consist primarily of any wage effects of training, rather than differences in hours worked due to an earlier return to work in the post-training period. In interpreting the estimated impacts, it is also important to keep in mind the width of the confidence intervals.

Positive selection bias into LMT is another possible explanation for the rather large earnings gains. We are not extremely successful in modelling the selection into training and critical observers may argue that unobserved characteristics which determine participation is likely to be correlated with earnings potential, violating the CIA.

Ideally we would like to have an internal comparison group of rejected applicants to measure the counterfactual outcome for participants, as recommended by Raaum and Torp (2002). In this previous study of the effects of LMT the argument goes as follows: "Our data indicate that training programmes attract applicants with better employment prospects than the average unemployed. This kind of self-selection, e.g. on post-training variables, is hard to identify and correct for". However, the magnitude of the bias is not very large. Moreover, the conclusion cited above is based on a stock-sampled comparison group with far less care – than in the present study - in creating an appropriate sample of non-participation. We believe the previous warnings about external comparison groups do not necessarily undermine the strategy in this paper.

When comparing short and long run effects we should also have in mind that the *business cycle* may influence the estimated impact of ALMPs. During a slump in the labour market when no employer open new jobs, improvement of skills through training will hardly help any unemployed back to work. During a boom, when employers are posting vacancies and job opportunities are improving, unemployed are more likely to gain from participation in training and other programmes. However, with full employment, when there is a lack of labour, all unemployed may get a job, both trained and untrained.

The post-training period of the present study, 1994-97, is a period of steadily decreasing unemployment and steadily increasing employment in Norway. Thus, the fortunate business cycle may be part of the explanation why the estimated annual effects are quite are large and why long run effects are just as – or even larger than – short run effects.

In a companion paper analysing short run effects of LMT over the whole business cycle 1992-1997, Raaum, Torp and Zhang (2002b) demonstrate that the training effect is larger when job opportunities in the national and local labour market are favourable.

8. Accumulated returns compared with direct training costs

While participants have lower earnings than the matched non-participants do during the training year, the single year differentials turn in favour of the participants during the post-training period; see Figures 4 and 5. The earnings loss while on training is substantially lower than the earnings gains for most groups. We measure the individual gain by means of the accumulated discounted earnings over the training and post-training periods and compare with the direct costs of providing the training. To assess the sensitivity with respect to valuation of future earnings, we calculate the returns for three alternative discount rates: 0.03, 0.05 and 0.07.

Estimates of the *direct operating costs* per participant are based on the average monthly cost of putting up the LMT courses in 1992 and 1993 respectively (measured by 1997 Nok), as reported by the *Directorate of Labour*. Average monthly costs are multiplied by the mean duration of training defined as the number of consecutive months in LMT programme, by cohort, gender and UB entitlement to get average costs per participant.[21] Then we correct the direct training cost with the *net* marginal cost of public funds, estimated by Holmøy and Strøm (2002) to be 1.15 in Norway (based on data from 1992). See Appendix, Table A5 for more details.

Table 3 displays the accumulated effects of LMT for each training cohort, by gender and UB entitlement. Earnings for five post-training years are available for 1993 participants only.[22] We also report accumulated effects as an average over the four cohorts by gender and UB entitlement (average 1992-93).

Focusing on these average effects, ignoring the heterogeneity across cohorts, we find that *participants with recent labour market experience* (i.e. with UB entitlement) had a substantial accumulated gain in earnings. For women, the estimated gain exceeds the direct costs of training in most cases.

[21] Duration of training at the individual level is recorded in our database. Information on monthly direct costs according to characteristics of the participants like gender and UB entitlement etc. is not available.
[22] In Table 4 we present the accumulated effects for an expanded post-training period where we extrapolate the annual effects to get estimates beyond the observation period.

Table 3. Accumulated private returns of LMT. Present value of annual average earnings gain (standard errors)

			4 post training years			5 post training years			Direct costs
Discount rate:			0.03	0.05	0.07	0.03	0.05	0.07	MCF 1.15
Cohort	UB	Gen							
1 W92	Yes	M	**42,829**	**40,424**	**38,201**	**57,452**	**53,706**	**50,288**	28,850
			(11,933)	(11,464)	(11,031)	(14,361)	(13,637)	(12,975)	
2 A92	Yes	M	14,083	12,770	11,563	**24,709**	**22,422**	**20,345**	40,220
			(7,868)	(7,510)	(7,180)	(10,039)	(9,484)	(8,978)	
3 W93	Yes	M	**33,052**	**31,314**	**29,698**				29,773
			(11,330)	(10,854)	(10,416)				
4 A93	Yes	M	**24,873**	**22,839**	**20,974**				38,235
			(9,372)	(8,923)	(8,509)				
Average 1992-93	*Yes*	*M*	*26,914*	*25,098*	*23,424*	*37,310*	*34,461*	*31,868*	35,081
			(4,937)	(4,720)	(4,521)	(8,287)	(7,847)	(7,445)	
1 W92	Yes	F	**68,279**	**65,032**	**62,023**	**85,336**	**80,525**	**76,121**	29,378
			(11,842)	(11,308)	(10,815)	(13,895)	(13,175)	(12,518)	
2 A92	Yes	F	**26,420**	**24,573**	**22,874**	**39,676**	**36,614**	**33,831**	43,685
			(6,605)	(6,287)	(5,994)	(8,134)	(7,685)	(7,277)	
3 W93	Yes	F	**57,501**	**54,493**	**51,709**				31,582
			(11,805)	(11,281)	(10,797)				
4 A93	Yes	F	**36,759**	**34,117**	**31,691**				41,000
			(7,324)	(6,973)	(6,649)				
Average 1992-93	*Yes*	*F*	*41,800*	*39,283*	*36,962*	*54,238*	*50,618*	*47,318*	38,441
			(4,301)	(4,100)	(3,915)	(7,094)	(6,712)	(6,364)	
1 W92	No	M	-2,591	-2,750	-2,888	1,788	1,228	732	30,564
			(19,017)	(18,230)	(17,502)	(21,793)	(20,650)	(19,609)	
2 A92	No	M	14,846	13,634	12,523	27,957	25,544	23,361	40,764
			(11,862)	(11,349)	(10,875)	(14,686)	(13,887)	(13,160)	
3 W93	No	M	23,979	22,236	20,627				31,200
			(18,184)	(17,383)	(16,644)				
4 A93	No	M	21,959	20,382	18,927				39,461
			(12,274)	(11,708)	(11,187)				
Average 1992-93	*No*	*M*	*15,532*	*14,302*	*13,171*	*17,672*	*15,987*	*14,467*	36,373
			(7,358)	(7,036)	(6,739)	(12,362)	(11,701)	(11,100)	
1 W92	No	F	7,001	6,644	6,307	11,553	10,779	10,070	31,204
			(12,436)	(11,903)	(11,411)	(17,774)	(16,847)	(16,001)	
2 A92	No	F	-5,524	-6,046	-6,512	6,774	5,125	3,653	46,289
			(7,240)	(6,918)	(6,621)	(9,694)	(9,154)	(8,663)	
3 W93	No	F	**54,948**	**52,048**	**49,364**				35,155
			(14,167)	(13,568)	(13,015)				
4 A93	No	F	3,054	2,231	1,482				44,361
			(7,879)	(7,511)	(7,173)				
Average 1992-93	*No*	*F*	*9,019*	*8,041*	*7,145*	*8,247*	*6,868*	*5,631*	41,483
			(4,721)	(4,511)	(4,318)	(8,660)	(8,190)	(7,762)	

Notes. 1) Cohorts. 2) UB eligibility. 3) Gender: M=male, F=female. 4) Present values of accumulated effects (incl. earnings forgone in training year) significant at 5 per cent in **bold.** 5) Standard errors are estimated by bootstrapping, 200 replications. 6) Direct costs: Average costs for LMT courses of average duration. Market prices corrected for MCF. See Appendix, Table A5. 7) Average 1992-93 indicates weighted averages of present values and costs, weighted according to # of participants in each cohort. All Nok 1997-values (Nok 100 = Euro 12.8 = US$ 11.2).

Actually, even the lower end of the confidence interval for the estimated five-year effect exceeds the costs, assuming a reasonable discount factor of 0.05. The accumulated effect for men with UB entitlement is lower. While the male five-year effect is very close to the direct costs, the average effect is lower when post-training period is restricted to four years.

There is substantial variation across cohorts. Table 3 clearly shows that the accumulated effects are higher for winter than for autumn courses, largely explained by the negative first year effects of autumn courses discussed in section 7. Typically, the accumulated effects exceed the direct costs for winter courses while the opposite is true for autumn courses.

The average effects for *labour market entrants*, i.e. participants *without* UB entitlement, are much lower than for experienced participants and not significantly different from zero. The accumulated average effects are (insignificantly) higher for men than for women. The direct costs are two times higher than the effects for men and the difference is even larger for women, since the effects are close to zero. Again, there is substantial cohort heterogeneity. Most of the cohort-specific effects are close to zero and some four-year effects are even negative. However, the precision of the estimates is low and the null-effect hypothesis cannot be rejected, except for women in the winter 1993 cohort for which the effect is significantly positive.

Projected future earnings gain

The limited information on post-programme outcomes causes practical problems in most of evaluations of long run effects. Even with five years of post-training earnings, we tend to underestimate the benefits of the programme when long run effects are positive. The asymmetry in timing of costs and benefits is the core of the problem. While (most) costs are realised during the training period, benefits may prevail through a long post-training period and are thus less easy to observe and measure. Therefore, earlier studies of social returns like Heckman and Smith (1997) and Heckman et al. (1999) extrapolate, using the most recent effects (i.e. the effects in the end of the observation period) to predict effects beyond the observation period. This strategy is reasonable unless the period-effects are highly volatile. In our case, the effects during the end of the observation period are fairly stable (or slightly increasing), see Appendix, Table A3.

In Table 4 we extend the post-programme period and report accumulated effects for different alternatives, using a discount rate of 0.05. The first column (4 years) simply repeats the average accumulated effect over the first four years from Table 3. Figures in the next three columns (5 years, 6 years and 7 years) are partly based on projected future earnings. For more details, see Appendix, Table A4.

As we expand the post-training period from four to seven years the accumulated effects increase for all groups. The accumulated earnings gain more than doubles for participants with recent labour market experience. For labour market entrants the accumulated earnings gain increases even more, because the time profile of their annual earnings gain is steeper.

Table 4. Accumulated effects of LMT for expanded post-training periods based on projected future earnings gain

Cohort	UB	Gen	Accumulated earnings effects (discounted at 5 per cent), after 4 years	5 years	6 years	7 years	Direct costs
All	Yes	M	25,098	25,090	35,535	54,771	35,081
All	Yes	F	39,283	41,398	55,563	81,692	38,441
All	No	M	14,302	14,311	23,252	39,151	36,373
All	No	F	8,041	8,038	16,758	31,568	41,483
Projected future earnings			None	Year 5 for the 1993 cohorts	Year 5 for the 1993 cohorts, year 6 for all	Year 5 for 1993, year 6 and 7 for all	

Note: Earnings gain for the non-observed years are projected by means of the most recent two-year average. The 5th year effects for 1993 cohorts are estimated by the average of the 3rd and 4th year effects etc. See Appendix, Tables A3 and A4. Direct costs of operating LMT are measured by market prices corrected for MCF=1.15. See Appendix, Table A4.

For workers with recent labour market experience the accumulated earnings gain over 7 years clearly exceeds the direct cost. The benefit-cost differential becomes economically significant for both men and women. The accumulated effect remains considerably lower for labour market entrants, but it approaches the direct costs of providing the training as we expand the post-training period. Actually, the average effect for men is very close to costs. For women it is lower than the costs, even after seven years.

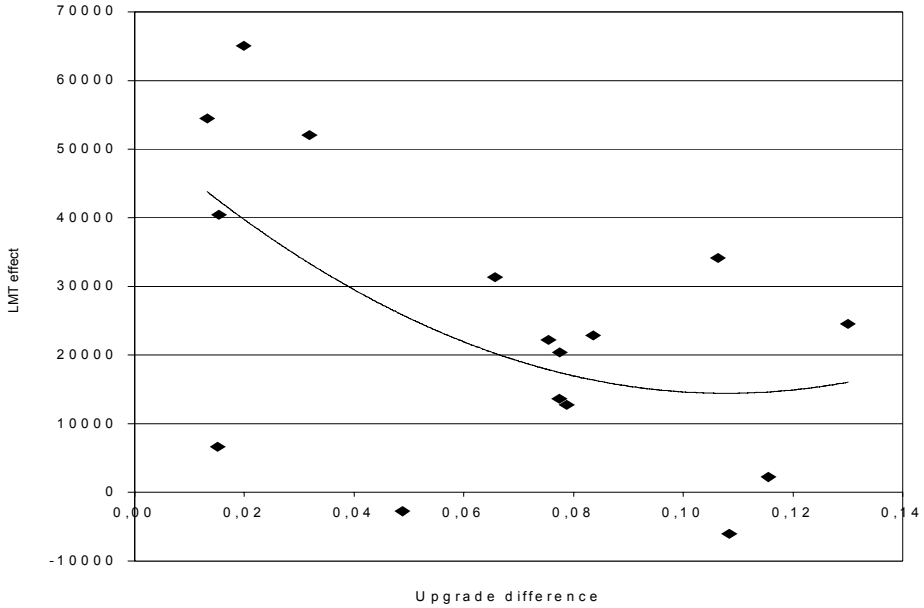
Focusing on efficiency, one policy implication from Table 4 seems obvious. Even if distributional concerns may indicate the opposite, priority should be given to unemployed with labour market experience when distributing scarce training slots. Differences in financial incentives may explain why participants without UB entitlement, on average, gain less in terms of future earnings. The training allowance provides an *economic incentive to participate*, even for unemployed who have a minor advantage in terms of improved post-training outcomes. As most participants with UB entitlement receive the same benefits whether they participate or not, they do not have the same economic incentives. This may explain why average training effects are higher for experienced participants.

The interplay between labour market training and participation in *ordinary education* does, however, represent another explanation for why labour market entrants who participate in LMT

gain less in terms of earnings during the first post-training years. Especially among workers without UB entitlement, LMT may operate as a stepping-stone into ordinary education, which will take place during what we define as the post-training period. A success criterion based on accumulated annual earnings will tend to disguise a potentially important effect of training, operating through acquisition of formal educational qualifications. If LMT qualifies or motivates the participants to invest more in education, earnings will drop during this investment period, and our criterion will indicate failure rather than success.

Our data sources enable us to identify whether participants and non-participants upgrade their formal qualifications during the training and post-training period. Some of the LMT courses are sufficiently long to give this upgrading, but the observed upgrading will also contain completion of more formal education. Table A6 in Appendix reveals: (i) Between 23 and 37 per cent of the participants in LMT upgrade their highest educational attainment within the period 1993-1999. (ii) The difference between participants and non-participants varies between 1.5 and 13.0 per cent across group. (iii) Women are more likely to upgrade than men are, and (iv) female participants in autumn courses have the highest rate of upgrading compared with non-participants.

Figure 6. Accumulated four-year effects versus difference in upgrading of highest formal educational attainment during 1993-1999 (the post-training period). Plot and polynomial regression for 16 subsamples



Notes: Accumulated long run training effects are taken from Table 3, first column (4 years present value, discount rate 0.05). Differences in upgrading of highest formal educational attainment are taken from Table A6, Appendix. The regression line is a second order polynomial based on the 16 observations.

When it comes to labour market experiences, here indicated by UB entitlement, (v) the difference in upgrading between participants and the comparison group is about the same for those with and those without UB entitlement.

In Figure 6 we plot the difference in fraction with upgraded education between participants and non-participants education against the accumulated four-year effect (discount rate equal to 0.05) for each of the 16 subsamples. As can be seen, the accumulated earnings effect is negatively correlated with the upgrading difference. Consequently, the small and even insignificant accumulated earnings effects for some groups seem to partly reflect a higher enrolment into ordinary education.

9. Conclusions

Individual long run effects of a labour market training (LMT) programme targeted at unemployed adults are evaluated by comparing mean post-training earnings for matched samples of participants and non-participants. Our Norwegian data cover all LMT participants as well as all unemployed individuals during 1992 and 1993. The matching procedure selects unemployed non-participants who have the closest set of predicted probabilities from a multinomial choice model where training, participation in other programmes, exiting the unemployment register or remaining unemployed are alternative transitions from the state of unemployment. The average treatment effect on the treated (ATET) is defined as the accumulated (discounted) earnings differential between participants and non-participants, measured over the training year and the post-training period of 4-5 years. The average direct cost of providing the training is based on average monthly operating costs per participant corrected with the marginal cost of public funds, vary from Nok 29,000 to Nok 46,000, depending on the duration of training.

The training effects (ATET) for participants with recent labour market experience during the pre-training years are substantial, at average varying from Nok 25,000 to Nok 55,000 - depending on duration of the post-training period and the discount rate. The accumulated effects for labour market entrants are much lower and in most cases not significantly different from zero. For participants with recent working experience, the average female ATET exceeds the direct cost of training, while the average effect for males is close to the direct costs of training. The training effects for both female and male labour market entrants are clearly lower than the costs.

Since *annual* training effects are positive after 4 or 5 years, the accumulated ATETs are sensitive to the number of post-training outcome years covered by the data. When we extrapolate the annual earnings gain and expand the post-training period from four to seven years, the accumulated effects increase substantially for all groups. For participants with recent works experi-

ence, the accumulated effects clearly exceed the direct costs and the difference turns economically significant for both men and women. Accumulated effects remain considerably lower for labour market entrants but the effects approach the direct costs of providing the training. Actually, the average effect for men is very close to costs. For women it remains lower than the costs, even after seven years.

Lessons to be drawn from this study are related to actual labour policies in (countries like) Norway as well as to more general evaluation issues. At the general level, it illustrates the need for long run post-programme data. Since substantial productive resources are needed to implement ALMP, evaluations that compare benefits in terms of individual effects and social costs are highly valuable to policy makers.

In our view, previous evaluation studies have been hampered by short-run data, and outcome measures that are difficult to transform into cost-comparable units. There are, however, problems associated with the earnings-based success criterion as well. As far as programme participation qualifies or motivates the participants to invest more in education, earnings will drop during the investment period. Enrolment in ordinary education will be interpreted as failure rather than success in the short run. Since any positive earnings effects are delayed, this further emphasise that post-programme earnings need to be recorded over several years.

As far as the Norwegian ALMPs are concerned, we show that labour market training participation does raise the post-training earnings of unemployed with recent work experience. For women the accumulated effect over a period of 5 years exceeds the direct operating costs corrected for marginal costs of public funds. The same is true for men if we expand the post-training period to six or seven years. For labour market entrants, however, the effects are significantly lower. Since the annual effects are increasing as the post-training period expands, the accumulated effects are highly sensitive to the duration of the post-training period. On the basis of the first four years of post-training outcomes, the policy implication would certainly be to target LMT at unemployed with recent work experience.

A possible explanation for the less encouraging results for labour market entrants is that they receive a training allowance and thus have an incentive to participate, even if the training has a minor impact on their future labour market opportunities. The increasing annual effects for this group and a conceivable interaction between LMT and ordinary education, postponing the earnings effect, call for further investigation into this issue.

References

AM (2000), Effects of Danish Employability Enhancement programmes. www.am.dk/english/publications/effects/eodeep.pdf. Copenhagen: Danish Ministry of Labour.

Ashenfelter, Orley (1978), "Estimating the Effect of Training programs on Earnings", *Review of Economics and Statistics* 60, pp. 47-57.

Axelsson, Roger (1992) *Arbetsmarknadsutbildningens privat- och samhällsekonomiska effekter*, EFA-Report No 25. Stockholm: Arbetsmarknadsdepartementet (Ministry of Labour).

Barth, Erling and Marianne Røed (1999), "The return to human capital in Norway: A review of the literature". In: Asplund, R., Pereira, P. (Eds.), Returns to human capital in Europe: A literature review. Taloustieto Oy, ETLA-Helsinki.

Barth, Erling and Josef Zweimüller (1994), "Bargaining structure, wage determination, and wage dispersion in 6 OECD countries." *Kyklos* 46, 81-93.

Barnow, Burt S. (1987), "The impact of CETA programs on Earnings: A review of the Literature." *Journal of Human Resources* XXII, pp. 157-193.

Björklund, Anders (1990), "Evaluation of Swedish Labor Market Policy." *Finnish Economic Paper* 3. pp. 3-13.

Björklund, Anders (1993), "The Swedish experince." In Jensen, K. and P.K. Madsen (eds) *Measureing labour market measures: Evaluating the effects of active labour market policies*. Preceedings from the Danish precidency conference. Copenhagen: Ministry of Labour.

Bloom, Howard S., Larry L. Orr, Stephen H. Bell, George Cave, Fred Doolittle, Winston Lin and Johannes M. Bos (1997), "The benefits and costs of JTPA Title II-A programs. Key findings from the national Job Training Partnership Act Study." *Journal of Human Resources*, 32:549-576.

Brodaty, Thomas, Bruno Crepon and Denis Fougere (2001), "Using matrching estimators to evaluate alterrnative youth employment programs: Evidence from France, 1986-1988." In M. Lechner and F. Pfeiffer (eds) *Econometric evaluation of Labour market Policies*. Heidelberg: Physica.

Calmfors, Lars, Anders Forslund and Maria Hemström (2002), *Does active labour market policy work? Lessons from the Swedish experiences*. Working paper 2002:4. Uppsala: IFAU – office of labour market policy evaluation (to be published in *Swedish Economic Policy Review*).

Couch, Kenneth (1992), "New evidence on long term effects of employment and training programs." *Journal of Labor Economics*, 10, pp 380-388.

Deheija, Rajeev and Sadek Wahba (1998), "Propensity score marching methods for nonexperimental causal studies." *NBER Working Paper No 6829*. Cambridge, MA.

Deheija, Rajeev and Sadek Wahba (1999), "Causal effects in nonexperimental studies: Reevaluating the evaluations of training programs." *Journal of the American Statistical Association,* 94, pp 1053-1062.

Diamond Peter E. and James A. Mirrlees (1971), "Optimal taxation and public production I: Production efficiency – and – II: Tax rules." American Economic Review, 61:261-278

Fay, Robert G. (1996), "Enhancing the effectiveness of active labour market policies: evidence from programme evaluations in OECD countries." *Labour market and social policy occasional papers no 18*. Paris: OECD.

Freedman, Stephen, Daniel Friedlander, Winston Lin and Amanda Schweder (1996), *The GAIN evaluation: Five-year impacts on employment, earnings and AFDC receipt*. Working paper 96:1. Manpower Demonstration Research Corporation.

Friedlander, Daniel and Gary Burtless (1995), *Five years after: The long term effects of Welfare-to-work programs*. New York: Russell Sage.

Friedlander, Daniel and Gayle Hamilton (1993), *SWIM: A five-year follow-up study.* Manpower Demonstration Research Corporation.

Grubb, W. Norton (1999), "Lessons from education and training for youth: five precepts, in preparing youth for the 21st century: the transition from education to the labour market." Paris: OECD.

Heckman, James J., Hidehiko Ichimura, Jeffrey Smith and Petra Todd (1998), "Characterizing selection bias using experimental data." *Econometrica,* 66, pp.1017-1089.

Heckman, James J., Hidehiko Ichimura and Petra Todd (1997), "Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme." *Review of Economic Studies,* 64, pp. 605-654.

Heckman, James J., Hidehiko Ichimura and Petra Todd (1998), "Matching as an econometric evaluation estimator." *Review of Economic Studies,* 65, pp. 261-294.

Heckman, James J. and Jeffrey Smith (1997), "Evaluating the welfare state." In Strøm, Steinar (ed): *Econometrics and Economic Theory in the 20th Century. The Ragnar Frisch Centennial Symposium.* Cambridge: Econometric Society Monographs No. 31,

Heckman, James J., Robert LaLonde and Jeffrey Smith (1999), "The Economics and Econometrics of Active Labor Market Programs." In O. Ashenfelter and D. Cards, (eds.), *Handbook of Labor Economics, Volume III,* Amsterdam: North-Holland Elsevier.

Hirano, Keisuke, Guido W. Imbens and Geert Ridder (2000), "Eficient estimation of average treatment effects using the estimated propensity score." *NBER Technical Working Paper No 251.* Cambridge, MA.

Hollister, Robinson, Peter Kemper and Rebecca Maynard (1984), *The national supported work demonstration.* Madison: University of Wisconsin Press.

Holmøy, Erling (2002), "Hva koster tidligere pensjonering samfunnet?" (Social costs of early retirement) *Økonomiske analyser 2/2002:41-47.* Oslo: Statistics Norway.

Holmøy, Erling and Birger Strøm (2002), *The social cost of public spending: A CGE assessment of resource costs and costs of public funds in Norway.* Discussion Paper Oslo: Statistics Norway.

Hotz, V. Joseph, Guido W. Imbens and Jacob A. Klerman (2000), "The long-term gains from GAIN: A re-analysis of the impacts of the Californian GAIN program." *NBER Working Paper No 8007.* Cambridge, MA.

Imbens, Guido W. (2000), "The role of the propensity score in estimation dose-response functions." *Biometrica* 87:706-710.

Imbens Guido W., and Joshua D. Angrist (1994), "Identification and estimation of local average treatment effects." *Econometrica* 62:467-476.

Jensen, Peter, Peder J. Pedersen, Nina Smith and Niels Westergård-Nielsen (1993), "The effects of labour market training on wages and unemployment: Some Danish results." In Bunzel et al. (eds.), *Panel Data and Labour Market Dynamics.* Amsterdam: North-Holland Elsevier.

Kemper, Peter, David A. Long and Craig Thornton (1984), "A benefit-cost analysis of the Supported Work Experiment." In: Hollister, R., et al. (ed), *The national supported work demonstration,* pp 239-285. Madison: University of Wisconsin Press.

LaLonde, Robert (1986), "Evaluating the Econometric Evaluations of Training Programs with Experimental Data." *American Economic Review* 76, pp. 604-620.

Larsson, Laura (2000), Evaluation of Swedish youth labour market programmes. *Working Paper 2000:1. Uppsala: The Office of Labour Market Policy Evaluation.*

Lechner, Michael (2001a), "Identification and estimation of causal effects of multiple treatments under the conditional independence assumption." In M. Lechner and F. Pfeiffer (eds) *Econometric evaluation of Labour market Policies,* 43-58. Heidelberg: Physica.

Lechner, Michael (2001b), "Programme heterogeneity and propensity score matching: An application to the evaluation of active labour market policies." *Discussion Paper*. Swiss Institute for international economics and applied economic research. University of St.Gallen.

Martin, John P., and David Grubb (2001), *What works for whome: a review of OECD countries' experiences with active labour market policies*. Working Paper 2001:14. Stockholm: IFAU Office of Labour Market Policy Evaluation.

NOU 1997:27 (1997), Nyttekostnadsanalyser (Cost benefit analyses). Oslo: Ministry of Finance.

NOU 1998:16 (1998), Nyttekostnadsanalyser. Veiledning i bruk av lønnsomhetsvurderinger i offentlig sektor (Cost benefit analyses. Guidelines for public sector evaluations). Oslo: Ministry of Finance.

Orr, Larry L., Howard S. Bloom, Stephen H. Bell, Winston Lin, George Cave, Fred Doolittle (1995), *The national JTPA study: Impacts, benefits, and costs of Title II-A*. Bethesda, MD: Abt Associates.

Riccio, James and Daniel Friedlander (1992), *GAIN: Program strategies, participation patterns, and first-year effects in six counties*. Manpower Demonstration Research Corporation.

Riccio, James Daniel Friedlander and S. Freedman (1994), *GAIN: Benefits, costs and three-year impacts of a welfare-to-work program*. Manpower Demonstration Research Corporation.

Rosenbaum, Paul R. and Donald B. Rubin (1983), "The central role of the propensity score in observational studies of causal effects." *Biometrica*, 70:41-55.

Rubin, Donald B. (1979), "Using multivariate matched sampling and regression adjustment to control bias in observational studies." *Journal of American Statistical Association,* 74:318-328.

Røed, Knut and Tao Zhang (1999), "Unemployment Duration in a Non-Stationary Macroeconomic Environment", Memorandum No 14, Department of Economics, University of Oslo

Raaum, Oddbjørn and Hege Torp (2002), "Labour market training in Norway – effect on earnings." *Labour Economics* 9 (2002), 207-247.

Raaum, Oddbjørn, Hege Torp and Tao Zhang (2002a) "Labour Market Training Participation in Norway 1991-1996: Construction of treatment and non-treatment groups by matching on observables", Working Paper No 2 :2002. Oslo: Ragnar Frisch Centre for Economic Research.

Raaum, Oddbjørn, Hege Torp and Tao Zhang (2002b), "Business cycles and the impact of labour market programmes. Memorandum No 14 (2002). Department of Economics, University of Oslo.

Raaum, Oddbjørn and Tom Erik Aabø (1999), *The effect of schooling on earnings: The role of family background studied by a large sample of Norwegian twins*. Memorandum No.16 (1999), Department of Economics, University of Oslo.

Sianesi, Barbara (2001), "Differential effects of Swedish active labour market programmes for unemployed adults during the 1990s" Working Paper W01/25. London: The Institute for Fiscal Studies.

Sianesi, Barbara (2002), "An evaluation of the Swedish system of active labour market programmes in the 1990s" Working Paper W02/01. London: The Institute for Fiscal Studies.

Smith, Jeffrey and Petra Todd (2002). "Is matching the answer to LaLonde's critique of nonexperimental methods?" *Journal of Econometrics*, forthcoming.

Stanley, Marcus, Lawrence Katz and Alan Kreuger (1998), "Developing skills" *Working Paper*. Cambridge, MA: Harvard University.

Stewart, Mark (1996), "Nordic wage differential: A view from outside." In: Wadensjö, E. (Ed.), *The Nordic Labour Markets in the 1990s* (Part 1). North-Holland Elsevier, Amsterdam.

Westergård-Nielsen, Niels (2001), *Danish labour Market Policy: Is it worth it?* Working paper 01-10. Aarhus: Centre for Labour Market and Social Research.

Appendix

1. Dependent variable: Annual earnings

Our measure of annual earnings is based on information from the tax register, a measure of income called "*Income Qualifying for Pensions*" (PI). It includes wages, sickness payment, and earnings of self-employed. Unemployment benefits are included in PI, but not in our measure of annual earnings. Training allowances are not included, but earnings from participation in other labour market programme than training are difficult to sort out and are therefore included.

2. Explanatory variables used for estimation of the propensity scores

(a) Married (dummy)

(b) Level of education: *Educ1 - Educ6*
6 dummies: le 9 years, 10 ys, 11-12 ys (reference), 13-16 ys, ge 17 ys, and unknown

(c) County of residence: 19 dummies
one for each county in Norway, county of Oslo as reference

(d) Age: *Age1 – Age5*
5 dummies: 25-30 ys, 31-35 ys, 36-40 ys (reference), 41-45 ys, 46-50 ys

(e) Immigrant from outside OECD: Immigrant (dummy)

(f) Unemployment history: *Month1 – Months1923*,
i.e. number of months of unemployment before *t*: 16 dummies: 0 (reference), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12,13-15, 16-18, 19-23 months

(g) Earnings history: *Earnings1-Earnings2123*,
i.e. number of years of annual earnings above B.a. before the year *T* (B.a. = Basic amount in the Norwegian Social Insurance Scheme, annually regulated, about Euro 5-6,000 in the period of interest). Earnings history serves as an indicator of aggregated experience

13 dummies: 0 (reference), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11-15, 16-20, and 21-23 years of income above B.a. (23 is maximum for T= 1991 as the Norwegian Social Insurance Scheme was established in 1967)

(h) LMT history: *LMT1 – LMT8,*
i.e. number of quarters participated in LMT during last 23 months before *t:* 8 dummies, one for each of the latest 8 quarters LMT8 means last quarter, LMT7 means the second last quarter .... LMT1 means two years ago

(i) Programme history: *PROG1 – PROG8,*
i.e. number of quarters participated in other programmes than LMT during last 23 months before *t:* 8 dummies, on for each of the latest 8 quarters PROG8 means last quarter, PROG7 means the second last quarter .... PROG1 means two years ago

(j) Occupational background: *Occup1-Occup6*
categories (based on ISCO): (1) technical, physical science, humanistic and artistic work (teachers, nurses, doctors, technicians etc), (2) administrative executive work, clerical work and sales work, (3) agriculture, forestry, fishing and related work, (4) manufacturing work, mining, quarrying, building and construction work (reference), (5) service work, transport and communication, (6) unknown

(k) Children in household: *Kid1 – Kid3*

4 dummies: 0 (reference), 1, 2 or 3 children and more, below the age of 18

(l) Previous annual earnings: AnnEarn1 – AnnEarn3

3 continuously distributed variables for last year before *t,* second last year, and next to second last year: annual earnings measured by B.a.

(m) Left ordinary education just before *t; LeftEduc* (dummy),

(n) Left ordinary high level education just before *t; LeftEducHigh*
(dummy, two highest levels of education)

Table A1-1. Selected explanatory variables used for estimation of propensity scores. Cohort: Autumn 1992, BEFORE matching. Mean values by gender, unemployment benefits entitlement, and LMT status

	Men				Women			
	With UB		No UB		With UB		No UB	
	LMT	Non-part.	LMT	Non-part.	LMT	Non-part.	LMT.	Non-part.
Married	0.534	0.515	0.506	0.567	0.265	0.273	0.231	0.308
Educ le 9 years, and unknown.	0.054	0.272	0.091	0.303	0.047	0.239	0.080	0.269
Educ 10-12 years	0.811	0.614	0.695	0.558	0.831	0.650	0.791	0.611
Educ 13 + years	0.135	0.114	0.214	0.138	0.122	0.111	0.129	0.120
Age 25-30 years	0.422	0.354	0.435	0.369	0.340	0.391	0.377	0.349
Age 31-35 years	0.211	0.211	0.246	0.230	0.233	0.205	0.270	0.225
Age 36-40 years	0.149	0.163	0.156	0.162	0.186	0.152	0.180	0.186
Age 41-50 years	0.219	0.272	0.163	0.239	0.240	0.252	0.173	0.240
Immigrant	0.091	0.094	0.383	0.209	0.041	0.050	0.196	0.107
Unemployment history - months last two years								
- 0 months	0.042	0.038	0.106	0.090	0.041	0.051	0.128	0.151
- 1-3 months	0.195	0.145	0.313	0.192	0.236	0.183	0.410	0.243
- 4-12 months	0.525	0.496	0.409	0.389	0.537	0.521	0.367	0.389
- 13 months +	0.238	0.322	0.172	0.329	0.185	0.245	0.096	0.217
Earnings history – Social Insurance Points Years								
- 0 point years	0.004	0.004	0.342	0.147	0.010	0.007	0.286	0.173
- 1-10 point years	0.474	0.410	0.361	0.436	0.613	0.621	0.585	0.613
- 11-15 point years	0.203	0.205	0.129	0.179	0.248	0.234	0.093	0.146
- 16-20 point years	0.150	0.177	0.104	0.145	0.096	0.103	0.031	0.055
- 21-23 point years	0.169	0.204	0.064	0.093	0.032	0.035	0.004	0.013
Part. In LMT last quarter	0.232	0.043	0.230	0.058	0.254	0.049	0.238	0.075
Part in programme last quarter	0.041	0.034	0.070	0.058	0.043	0.024	0.032	0.047
Occupation – based on ISCO-88(COM)								
- ISCO 2, 3 and 513	0.133	0.115	0.117	0.095	0.192	0.185	0.174	0.162
- ISCO 1, 4 and 52	0.142	0.149	0.077	0.085	0.482	0.425	0.253	0.231
- ISCO 6	0.029	0.023	0.047	0.043	0.014	0.015	0.005	0.013
- ISCO 7 and 81, 82	0.509	0.501	0.310	0.346	0.093	0.125	0.068	0.077
- ISCO 511-12, 514, 516, 9	0.161	0.193	0.128	0.162	0.197	0.227	0.192	0.219
- Unknown occupation.	0.026	0.018	0.321	0.269	0.021	0.022	0.307	0.298
No children	0.500	0.478	0.510	0.534	0.250	0.268	0.181	0.277
1 child	0.220	0.247	0.224	0.215	0.282	0.312	0.253	0.287
2 children +	0.280	0.274	0.266	0.251	0.467	0.420	0.566	0.436
Index of earnings – last three years	10.954	11.310	3.744	5.118	7.778	8.314	1.865	3.537
Just left education	0.320	0.112	0.438	0.202	0.321	0.110	0.379	0.211
(N=)	(2,491)	(19,144)	(1,132)	(5,427)	(2,110)	(13,368)	(1,705)	(3,901)

Note: All variables are binary except Index of earnings. The estimated propensity scores include other explanatory variables as well. Complete statistics for this and other cohorts are available from the authors.

Table A1-2. Selected explanatory variables used for estimation of propensity scores. Cohort: Autumn 1992, AFTER matching. Mean values by gender, unemployment benefits entitlement, and LMT status

	Men				Women			
	With UB		No UB		With UB		No UB	
	LMT	Non-part.	LMT	Non-part.	LMT	Non-part.	LMT.	Non-part.
Married	0.534	0.519	0.508	0.515	0.266	0.275	0.231	0.240
Educ le 9 years, and unknown.	0.054	0.052	0.091	0.077	0.047	0.048	0.081	0.071
Educ 10-12 years	0.811	0.811	0.695	0.712	0.831	0.833	0.791	0.804
Educ 13 + years	0.135	0.137	0.214	0.210	0.122	0.119	0.128	0.124
Age 25-30 years	0.422	0.420	0.435	0.458	0.340	0.346	0.377	0.376
Age 31-35 years	0.211	0.222	0.245	0.227	0.233	0.235	0.271	0.286
Age 36-40 years	0.148	0.135	0.157	0.174	0.186	0.174	0.179	0.158
Age 41-50 years	0.219	0.223	0.163	0.142	0.240	0.245	0.174	0.180
Immigrant	0.090	0.097	0.379	0.383	0.040	0.029	0.194	0.190
Unemployment history - months last two years								
- 0 months	0.042	0.031	0.105	0.098	0.041	0.042	0.128	0.130
- 1-3 months	0.195	0.198	0.314	0.312	0.236	0.225	0.409	0.390
- 4-12 months	0.526	0.534	0.407	0.426	0.538	0.550	0.367	0.387
- 13 months +	0.237	0.237	0.174	0.165	0.185	0.183	0.096	0.093
Earnings history – Social Insurance Points Years								
- 0 point years	0.004	0.005	0.337	0.336	0.009	0.012	0.284	0.279
- 1-10 point years	0.474	0.474	0.364	0.373	0.614	0.615	0.587	0.586
- 11-15 point years	0.203	0.217	0.130	0.129	0.249	0.238	0.094	0.095
- 16-20 point years	0.150	0.139	0.104	0.102	0.096	0.091	0.031	0.036
- 21-23 point years	0.169	0.165	0.064	0.061	0.032	0.043	0.004	0.004
Part. In LMT last quarter	0.232	0.228	0.225	0.202	0.254	0.238	0.236	0.238
Part in programme last quarter	0.041	0.040	0.070	0.074	0.043	0.045	0.032	0.038
Occupation – based on ISCO-88(COM)								
- ISCO 2, 3 and 513	0.133	0.126	0.118	0.115	0.192	0.179	0.174	0.166
- ISCO 1, 4 and 52	0.141	0.140	0.077	0.070	0.482	0.507	0.253	0.270
- ISCO 6	0.029	0.029	0.047	0.031	0.014	0.020	0.005	0.002
- ISCO 7 and 81, 82	0.509	0.516	0.312	0.338	0.093	0.078	0.068	0.067
- ISCO 511-12, 514, 516, 9	0.161	0.159	0.129	0.147	0.198	0.194	0.193	0.205
- Unknown occupation.	0.026	0.029	0.317	0.298	0.021	0.022	0.306	0.289
No children	0.500	0.496	0.511	0.509	0.250	0.256	0.180	0.187
1 child	0.220	0.219	0.224	0.217	0.283	0.283	0.253	0.245
2 children +	0.280	0.285	0.265	0.273	0.467	0.460	0.566	0.567
Index of earnings – last three years	10.962	11.094	3.774	3.751	7.788	7.574	1.871	1.833
Just left education	0.320	0.322	0.436	0.388	0.321	0.314	0.377	0.370
(N=)	(2,488)	(2,488)	(1,123)	(1,123)	(2,106)	(2,106)	(1,697)	(1,697)

Note: All variables are binary except Index of earnings. The estimated propensity scores include other explanatory variables as well. Complete statistics for this and other cohorts are available from the authors.

Table A2. Means predicted propensity scores for the matched samples by cohort, UB entitlement and gender

Cohort	Winter 1992		Autumn 1992		Winter 1993		Autumn 1993	
Subsample	Partic.	Non-part	Partic.	Non-part	Partic.	Non-part	Partic.	Non-part
Men with UB								
Pred prob of LMT	0.0986	0.0982	0.1442	0.1438	0.0908	0.0905	0.1139	0.1137
Pred prob of PROG	0.0407	0.0406	0.0608	0.0607	0.0500	0.0499	0.0830	0.0829
Pred prob of OUT	0.2062	0.2055	0.2600	0.2597	0.1777	0.1774	0.2654	0.2653
Men without UB								
Pred prob of LMT	0.1589	0.1574	0.1928	0.1912	0.1550	0.1531	0.1503	0.1493
Pred prob of PROG	0.0514	0.0512	0.0617	0.0617	0.0459	0.0456	0.0638	0.0636
Pred prob of OUT	0.2299	0.2302	0.3041	0.3034	0.2414	0.2415	0.2811	0.2813
Women with UB								
Pred prob of LMT	0.1093	0.1084	0.1514	0.1505	0.0920	0.0910	0.1150	0.1145
Pred prob of PROG	0.0379	0.0377	0.0587	0.0584	0.0576	0.0571	0.0782	0.0781
Pred prob of OUT	0.2280	0.2273	0.3179	0.3178	0.2110	0.2114	0.3239	0.3237
Women without UB								
Pred prob of LMT	0.2127	0.2099	0.2938	0.2923	0.1945	0.1921	0.2350	0.2339
Pred prob of PROG	0.0544	0.0535	0.0634	0.0629	0.0504	0.0498	0.0782	0.0781
Pred prob of OUT	0.2608	0.2626	0.2887	0.2887	0.2599	0.2600	0.2788	0.2786

Table A3. Short-term and long-term average effects of LMT on annual earnings by cohort, unemployment benefits entitlement and gender. Mean values (and standard errors). Significant values in ***bold*** *(5 per cent). Norwegian kroner, 1997-values*

Cohort	UB	Gender	Training year	1. year effect	2. year effect	3. year effect	4. year effect	5. year effect
1 W92	Yes	M	-4,484	**8,509**	**11,761**	**14,661**	**16,376**	**16,952**
			(2,395)	(3,210)	(3,515)	(3,569)	(3,778)	(4,120)
2 A92	Yes	M	**-5,481**	**-6,185**	**8,845**	**9,913**	**9,183**	**12,318**
			(1,376)	(2,051)	(2,542)	(2,703)	(2,948)	(3,086)
3 W93	Yes	M	**-4,523**	**7,014**	**13,968**	**11,608**	**7,853**	
			(2,260)	(3,085)	(3,226)	(3,482)	(3,795)	
4 A93	Yes	M	**-5,442**	**-4,810**	**8,499**	**10,707**	**19,332**	
			(1,543)	(2,346)	(2,790)	(3,095)	(3,337)	
1 W92	Yes	F	459	**14,603**	**18,970**	**21,003**	**18,617**	**19,773**
			(1,931)	(2,860)	(3,217)	(3,384)	(3,573)	(3,751)
2 A92	Yes	F	-1,801	**-6,831**	**11,269**	**13,460**	**13,409**	**15,368**
			(1,115)	(1,646)	(2,088)	(2,234)	(2,401)	(2,565)
3 W93	Yes	F	-3,760	**12,445**	**16,355**	**18,466**	**18,981**	
			(2,081)	(3,100)	(3,455)	(3,694)	(3,923)	
4 A93	Yes	F	**3,515**	**-6,349**	**10,985**	**18,419**	**21,641**	
			(1,251)	(1,958)	(2,350)	(2,581)	(2,780)	
1 W92	No	M	-2,704	38	-2,417	-4,253	7,030	5,077
			(3,342)	(4,554)	(4,934)	(5,237)	(5,660)	(5,997)
2 A92	No	M	-3,189	-3,128	5,365	6,516	**11,313**	**15,200**
			(2,455)	(3,029)	(3,770)	(3,992)	(4,552)	(4,841)
3 W93	No	M	**-7,046**	568	**10,133**	10,711	**12,517**	
			(3,224)	(4,389)	(4,858)	(5,712)	(5,773)	
4 A93	No	M	**-4,509**	-2,679	**8,592**	**14,058**	9,122	
			(2,139)	(3,195)	(3,912)	(4,405)	(4,852)	
1 W92	No	F	-4,010	5,708	3,130	2,833	-84	5,277
			(2,196)	(3,214)	(3,619)	(3,778)	(4,101)	(4,431)
2 A92	No	F	-5,409	**-10,501**	-993	3,921	**8,360**	**14,257**
			(1,157)	(1,796)	(2,213)	(2,403)	(2,564)	(2,746)
3 W93	No	F	-3,836	**10,570**	**18,332**	**15,841**	**18,848**	
			(2,314)	(3,416)	(3,749)	(3,966)	(4,284)	
4 A93	No	F	**-4,867**	**-8,320**	2,712	**6,996**	**7,923**	
			(1,370)	(1,965)	(2,400)	(2,652)	(3,002)	

Notes. 1) Cohort: W = winter, A= autumn. 2) UB eligibility. 3) Gender: M=male, F=female. 4) Estimated average effect on annual earnings. Significant values at 5 per cent in **bold** (t-value more than 1.96). Standard errors estimated by bootstrapping, 200 independent replications. (Nok 100 = Euro 12.8 = US$ 11.2)

Table A4. Projected future earnings gain. Long-term average effects of LMT on annual earnings by cohort, unemployment benefits entitlement and gender. Mean values (and standard errors). Significant values in ***bold*** *(5 per cent). Norwegian kroner, 1997-values*

Cohort	UB	Gen	4. year effects	5. year effects. Projected for 1993 cohorts	Projected 6. year effects	Projected 7. year effects	Accumulated effects. Year 0-7 Disc rate 0.05
1 W92	Yes	M	**16,376**	**16,952**	**16,664**	**16,808**	78,086
			(3,778)	(4,120)	(3,949)	(4,034)	
2 A92	Yes	M	**9,183**	**12,318**	**10,751**	**11,534**	38,641
			(2,948)	(3,086)	(3,017)	(3,051)	
3 W93	Yes	M	**7,853**	**9,730**	**8,791**	**9,261**	52,080
			(3,795)	(3,639)	(3,717)	(3,678)	
4 A93	Yes	M	**19,332**	**15,020**	**17,176**	**16,098**	58,865
			(3,337)	(3,216)	(3,277)	(3,247)	
1 W92	Yes	F	**18,617**	**19,773**	**19,195**	**19,484**	108,696
			(3,573)	(3,751)	(3,662)	(3,707)	
2 A92	Yes	F	**13,409**	**15,368**	**14,388**	**14,878**	57,924
			(2,401)	(2,565)	(2,483)	(2,524)	
3 W93	Yes	F	**18,981**	**18,723**	**18,852**	**18,787**	96,583
			(3,923)	(3,809)	(3,866)	(3,837)	
4 A93	Yes	F	**21,641**	**20,030**	**20,836**	**20,433**	86,909
			(2,780)	(2,681)	(2,730)	(2,706)	
1 W92	No	M	7,030	5,077	6,053	5,565	9,700
			(5,660)	(5,997)	(5,829)	(5,913)	
2 A92	No	M	**11,313**	**15,200**	**13,257**	**14,228**	45,548
			(4,552)	(4,841)	(4,696)	(4,769)	
3 W93	No	M	**12,517**	**11,614**	**12,065**	**11,840**	48,753
			(5,773)	(5,742)	(5,758)	(5,750)	
4 A93	No	M	9,122	**11,590**	**10,356**	**10,973**	44,989
			(4,852)	(4,629)	(4,741)	(4,685)	
1 W92	No	F	-84	5,277	2,597	3,937	15,515
			(4,101)	(4,431)	(4,266)	(4,348)	
2 A92	No	F	**8,360**	**14,257**	**11,308**	**12,783**	22,648
			(2,564)	(2,746)	(2,655)	(2,700)	
3 W93	No	F	**18,848**	**17,345**	**18,096**	**17,720**	91,735
			(4,284)	(4,125)	(4,205)	(4,165)	
4 A93	No	F	**7,923**	**7,460**	**7,691**	**7,575**	19,199
			(3,002)	(2,827)	(2,914)	(2,871)	

Notes. 1) Cohort: W = winter, A= autumn. 2) UB eligibility. 3) Gender: M=male, F=female. 4) and 5) Estimated average effect on annual earnings. 6) and 7) Projected effects for cohorts 1992: Effects for year 6 and 7 equals the present value of the average of the estimated effects for year 4 and 5 (see Table A3). Projected effects for cohorts 1993: Effects for year 5, 6 and 7 equals the present value of the average of the estimated effects for year 3 and 4 (see Table A3). Significant values at 5 per cent in **bold**. 8) Accumulated present values, training year (earnings forgone) and 7 post-training years. Discount rate 0.05. (Nok 100 = Euro 12.8 = US$ 11.2)

Table A5. Average duration of LMT courses and number of participants by cohort, UB-eligibility and gender. Average cost per month per participant, and average cost per participant. Nok current values, 1997-values, and 1997-values corrected for marginal costs of public funds

Cohort	UB	Gen	Average duration (months)	No of partici-pants	Average cost per month	Average cost per partici-pant	Average cost 1997 values	Average cost Corrected for MCF 1.15
1 W 92	Yes	M	4,41	1,553	5,510	24,314	26,843	28,850
2 A 92	Yes	M	6,15	2,488	5,510	33,897	37,422	40,220
3 W 93	Yes	M	4,64	1,896	5,539	25,673	27,701	29,773
4 A 93	Yes	M	5,95	2,204	5,539	32,970	35,575	38,235
1 W 92	Yes	F	4,49	985	5,510	24,759	27,334	29,378
2 A 92	Yes	F	6,68	2,106	5,510	36,817	40,646	43,685
3 W 93	Yes	F	4,92	962	5,539	27,234	29,385	31,582
4 A 93	Yes	F	6,38	1,750	5,539	35,355	38,148	41,000
1 W 92	No	M	4,67	729	5,510	25,759	28,438	30,564
2 A 92	No	M	6,24	1,123	5,510	34,355	37,928	40,764
3 W 93	No	M	4,86	805	5,539	26,904	29,029	31,200
4 A 93	No	M	6,14	1,123	5,539	34,028	36,716	39,461
1 W 92	No	F	4,77	753	5,510	26,299	29,034	31,204
2 A 92	No	F	7,08	1,697	5,510	39,012	43,069	46,289
3 W 93	No	F	5,47	772	5,539	30,314	32,709	35,155
4 A 93	No	F	6,91	1,554	5,539	38,253	41,275	44,361

Notes: 1) Cohort: W = winter, A= autumn. 2) UB eligibility at the time of registering: Yes/No. 3) Gender: M=male, F=female. 4) Duration of training, defined as the average number of consecutive months enrolled in the LMT programme. By cohort, UB entitlement and gender. 5) Number of observations. 6) Average variable costs per month for LMT, by year. 7) Duration times average cost. 8) Average cost per participant, Nok 1997-values. 9) Average cost per participant (Nok 1997-values) corrected for marginal costs of public funds, set to 1.15. (Nok 100 = Euro 12.8 = US$ 11.2)

Note on LMT courses:

Most of the LMT courses are arranged by external institutions. A small fraction is arranged by the LMT-Centres (owned by the Directorate of Labour). Courses arranged by the LMT-Centres are at average more costly than other courses, presumably because the Centres cover vocational training in need of more expensive machinery. Of the total training volume (persons x months) 12.2 per cent were covered by the LMT-Centres in 1992, 17.5 per cent in 1993.

In 1992 the average monthly cost of LMT per participant was Nok 5,165 for courses arranged by external institutions. In 1993 the average monthly cost was Nok 4,853. Similar figures for the LMT-Centres are Nok 8,000 (estimated) in 1992 and Nok 8,775 in 1993. By weighting the average costs with the respective fractions of the total training volume, we get average monthly costs of Nok 5,510 in 1992 and Nok 5,539 in 1993.

Table A6. Upgrading of highest formal educational attainment during 1993-1999 (the post-training period). Average rate among LMT participants and difference between participants and the comparison group by cohort, UB-eligibility and gender

	Men		Women	
Cohort	UB entitlement	No UB entitlement	UB entitlement	No UB entitlement
Rate of upgrading among participants				
All	0.261	0.273	0.287	0.325
Winter 92	0.247	0.276	0.290	0.291
Autumn 92	0.285	0.286	0.338	0.367
Winter 93	0.269	0.294	0.237	0.324
Autumn 93	0.242	0.237	0.285	0.316
Difference in upgrading: Participants - comparison group				
All	0.061	0.070	0.067	0.068
Winter 92	0.015	0.049	0.020	0.015
Autumn 92	0.079	0.077	0.130	0.108
Winter 93	0.066	0.075	0.013	0.032
Autumn 93	0.084	0.077	0.106	0.115