Economic Evaluation of the Training Opportunities Programme in New Zealand

Geoff Perry, AUT University
Tim Maloney, University of Auckland

Abstract

It is well accepted that a highly educated and well trained labour force is a precondition for sustained economic growth and development, and that the labour market outcomes for individuals are enhanced with higher levels of education and training. Recognition of these facts has influenced the types of active labour market programmes that are provided for the unemployed by governments throughout the OECD, with particular emphasis on training. In New Zealand, the Training Opportunities scheme, introduced in the early 1990's, remains today the major active labour market programme for the unemployed. This paper contributes to the literature in two ways. Firstly, the impact on male participants of being involved in Training Opportunities in the mid 1990's is evaluated. Secondly, short and medium term impacts for men are estimated using Difference-in-Differences matching, with careful attention to methodological concerns. The key findings are that while there is a short term beneficial effect for the programme as a whole, this is not consistent across all sub-groups. Further, the beneficial effect dissipates by the second year after receiving the intervention.

1. Introduction

As noted by Robinson (2000), one of the clearest findings of labour economics is that those who are more skilled and have higher qualifications generally have higher wage rates and probabilities of employment. This finding has influenced the approach by governments to active labour market policy with training becoming a major component. In New Zealand, the Training Opportunities Programme was introduced by the government in 1993 as the primary training scheme in its menu of active labour market policies intended to reduce unemployment. It remains in place as the major overall active labour market programme in 2007.

The international evaluations of training have produced inconclusive results as to the effects of these on the labour market outcomes of participants (Barnow, 1987;

LaLonde, 1995; Friedlander, *et al*., 1997; Stanley, *et al*., 1998; Heckman, LaLonde *et al*., 1999; Wunsch, 2005). However, the research indicates that results can vary over different time periods, the impacts on men tend to be lower than for women, that least advantaged groups benefit less than those that are more advantaged and that stigma and locking-in effects can contribute to negative outcomes. There has been little rigorous economic evaluation in New Zealand of active labour market programmes, including Training Opportunities. Most evaluations have been process evaluations or have evaluated the impact of an intervention only three or six months post participation.[1] This paper addresses these issues by applying a rigorous economic evaluation approach in estimating the impact of Training Opportunities in New Zealand upon male participants. The evaluation allows for impacts up to three years post-intervention. Careful attention is given to methodological aspects, with difference-in-differences, nearest-neighbour matching used to estimate these impacts.

   The paper is organised as follows. Firstly, the institutional backgrounds of both New Zealand and the Training Opportunities Programme are provided. Secondly, the 'evaluation problem' is raised, and a particular solution is proposed in this paper for mitigating its effects. Thirdly, the data used in this study are described, and the impacts of participation in Training Opportunities upon male participants are estimated and analysed. The final section provides a conclusion to this analysis.

2. Institutional Background

The implementation of Training Opportunities by the New Zealand government in 1993 was a response to the relatively high unemployment rate of 9.5 per cent overall and 10 per cent for males at the time. In this section details on unemployment in New Zealand are outlined, and the New Zealand approach to active labour market policy and the specific characteristics of the Training Opportunities Programme are described.

Unemployment in New Zealand

The unemployment pattern and trends in New Zealand in the 1980's and 1990's had been similar to that of Australia and, with some lag in time, to that of the OECD. The characteristics of the unemployed clearly indicate that the effects of unemployment are not distributed evenly across New Zealand society. (See table 1) Males were more likely to be unemployed than females, especially in the early 1990's when the proportion of the unemployed who were male peaked at nearly 60 per cent. Maori, who constitute 13 per cent of the overall population, are disproportionately represented with over 20 per cent of the male registered unemployed. Further, there is no indication that the proportion of registered unemployed who are Maori had fallen, with the proportion remaining fairly stable between 1986 and 1998. Another group affected strongly by unemployment are those with no or very low qualifications.

   Over the period between 1986 and 1997, the age composition of unemployment changed. The proportion of males who were unemployed between the ages 25 and 54 increased steadily from 37.9 per cent to 56 per cent. On the other hand, the proportion of young males registered as unemployed fell from 54.5 per cent to 36.5 per cent.

---

[1] The exception is the work by Mare (2002).

Table 1 - Characteristics of the Unemployed 1986-1997 - Males

Year	Percentage Maori	Percentage Aged 15-24	Percentage Aged 25-54	Percentage Aged 55+	Percentage Long-Term Unemployed[1]
1986	26.1	54.5	37.9	7.6	17.9
1987	23.0	51.1	43.4	5.5	23.8
1988	20.7	47.8	47.2	5.1	31.2
1989	21.1	43.9	50.2	5.9	36.1
1990	20.7	39.7	54.7	5.5	39.6
1991	19.9	40.2	54.9	4.9	46.7
1992	21.2	38.4	55.6	6.0	52.9
1993	20.9	36.8	57.2	6.0	51.6
1994	21.8	36.4	57.4	6.2	50.5
1995	21.4	37.5	56.3	6.1	43.7
1996	21.0	38.5	54.0	7.4	36.9
1997	21.5	36.5	56.0	7.5	37.3

[1] Unemployed for 26 weeks or more
Source: Statistics NZ, HLFS

New Zealand's Approach to Active Labour Market Policy

The changes in the scale and characteristics of unemployment in New Zealand have been important influences on government labour market policy. Following an emphasis on job creation and work experience programmes in the 1980's the government reoriented labour market policy in the late 1980's and 1990's. As part of the economic reform process in the 1980's the government decided to target its active labour market policies more effectively; in particular at reducing and preventing long-term unemployment and helping those labour market groups considered to be particularly at risk[2] or disadvantaged[3] (Department of Labour, 1993). Recognition that New Zealand lagged behind other OECD countries in the area of workforce skills, competence and human capital (OECD, 1996) had seen the government introduce a new approach to education and training.[4] Employment policies emphasised new training and skill development programmes that were linked to the new qualifications framework.

Active labour market policies in the 1990's, therefore, involved a set of specific

[2] To be considered at risk by the New Zealand Employment Service, and subsequently by Work and Income New Zealand, the client needed to have one of the following barriers; extensive enrolment history without interruption, a disability, domestic purposes or widows on a benefit for a year or more, community wage partners where the primary client has been on a benefit for one year or more, aged 55 or older, 16-20 year olds who had been on the benefit for 13 weeks or 16-18 year olds receiving an Independent Youth Benefit, quota refugees, Department of Correction clients, a person returning to the workforce (not registered and unemployed for four years or more) or women in non-traditional work.

[3] To be considered disadvantaged in the local labour market the client had to be less likely to gain employment than other people in the area due to their lack of one or more of the following; skills, qualifications, work experience or knowledge of the local labour market.

[4] This involved in the early 1990's the introduction of the New Zealand Qualifications Authority, the National Qualifications Framework involving nationally recognised unit standards that were transferable between training providers, the establishment of Industry Training Organisations (ITOs) and the Education and Training Support Agency (ETSA), and the bulk funding of schools (OECD, 1993).

Table 2 - Types of Active Labour Market Interventions as a Percentage of Total Spending on Active Labour Market Programmes

	1985						1990						1995						1998					
	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F
Australia[1]	26.2	2.4	16.7	0.0	28.6	26.2	33.3	22.2	14.8	0.0	0.0	29.6	28.8	17.5	7.5	3.8	27.5	15.0	36.5	9.6	7.7	3.8	13.5	28.8
Belgium	13.8	8.9	0.0	0.0	64.2	13.0	16.8	12.1	3.7	0.9	50.5	15.9	15.9	11.6	8.7	0.0	40.6	23.2	14.2	12.7	22.4	0.0	35.1	15.7
Denmark[2]	7.3	35.8	0.0	0.9	0.0	56.0	7.1	31.7	0.0	5.6	0.0	55.6	6.2	43.8	1.7	3.9	11.2	33.1	7.2	42.5	1.2	2.4	12.6	34.1
France	19.4	32.8	3.0	6.0	0.0	38.8	16.3	35.0	3.8	2.5	1.3	41.3	11.6	26.4	12.4	3.1	15.5	31.0	12.0	21.1	20.3	0.0	14.3	32.3
Germany[3]	25.9	18.5	6.2	0.0	14.8	34.6	21.6	34.3	4.9	0.0	9.8	29.4	16.1	31.5	4.9	2.1	22.4	23.1	17.5	27.0	2.4	2.4	25.4	25.4
Netherlands[4]	4.2	10.0	1.1	0.0	2.1	82.6	9.3	20.6	1.9	0.0	2.8	65.4	22.5	21.2	1.3	0.0	7.9	47.0	17.8	16.1	4.6	0.0	22.4	39.1
New Zealand	11.9	10.7	4.8	0.0	69.0	3.6	19.3	47.0	8.4	4.8	9.6	10.8	17.6	43.2	12.2	1.4	4.1	21.6	19.4	51.6	14.5	3.2	6.5	4.8
Spain[5]	26.5	5.9	23.5	11.8	29.4	2.9	17.3	9.3	18.7	25.3	16.0	13.3	12.1	37.9	12.1	4.5	7.6	25.8	10.3	14.7	30.9	4.4	11.8	27.9
Sweden	11.8	22.7	4.7	0.5	15.2	45.0	12.3	30.4	1.2	0.6	6.4	49.1	11.0	22.9	13.6	3.0	18.2	31.4	14.2	22.8	7.1	4.1	19.8	32.0
Switzerland	40.0	5.0	0.0	0.0	0.0	55.0	27.3	4.5	0.0	0.0	0.0	68.2	22.9	16.7	2.1	0.0	16.7	41.7	20.0	27.1	1.4	1.4	27.1	22.9
United Kingdom[6]	18.9	9.5	0.0	4.1	25.7	41.9	24.6	32.8	0.0	3.3	0.0	39.3	42.2	20.0	0.0	2.2	2.2	33.3	43.2	16.2	0.0	0.0	0.0	40.5
United States	25.0	42.9	3.6	0.0	3.6	25.0	32.0	36.0	0.0	0.0	4.0	28.0	36.8	21.1	0.0	0.0	5.3	36.8	35.3	23.5	0.0	0.0	5.9	35.3

[1] 1997 not 1998. [2] 1986 not 1985. [3] West Germany only in 1985. [4] 1996 not 1995. [5] 1996 not 1995. [6] 1997 not 1998.

A Employment Services – includes job placement, counselling and vocational guidance, administering unemployment benefits and referring unemployed to labour market programmes
B Training of unemployed adults
C Wage Subsidies
D Subsidies to start-up a business
E Work experience/job creation subsidies
F Other – includes youth programmes, programmes for the disadvantaged and training for the employed
Source: OECD (1992, 2000)

programmes that were aimed at increasing skill levels, maintaining workplace attachment and subsidising job search. As a result, the expenditure on training programmes in New Zealand increased from 10.7 per cent of active labour market expenditure in 1985 to 51.6 per cent in 1998 (see table 2). The use of work experience schemes declined and subsidy programmes to the private sector increased. The emphasis was in helping those who were disadvantaged, and this was often most easily identified by duration of unemployment.

Training Opportunities

In 1993, the New Zealand Government introduced the Training Opportunities Programme with the key objective of enabling the disadvantaged in the labour market to develop skills and capabilities that would assist them into further education and employment. The predecessor of the Training Opportunities Programme was the Access Programme, which was in place from 1987 to 1992. Although it was deemed to have been reasonably successful, this earlier programme was considered to have had a complex administrative structure, an inconsistent standard of qualifications across regions and providers and a cost-ineffective delivery system for people with higher levels of education (Ministry of Education, 2001).

Training Opportunities replaced Access in early 1993, and later in that same year subsumed Maori Access, which had run alongside the main Access Programme. This situation continued until 1998 when a number of changes were made. Firstly, the programme was split into two; Training Opportunities for those eighteen years of age and over and Youth Training for those who were sixteen or seventeen years of age. Secondly, the funding that had been provided through Vote Education and administered by Skill New Zealand, formerly the Education and Training Support Agency, was split into two with a percentage subsequently being funded through Vote Work and Income, with the intention of providing Work and Income New Zealand with more flexibility in accessing the programme. Training Opportunities is still the major active labour market training programme in 2007.

The original aim of Training Opportunities included targeting school leavers and long-term job seekers with no or low qualifications[5] and assisting them in gaining a recognised qualification that would help them move on to further education and eventually employment. The rationale for the programme was that 'participation in second chance education provides the opportunity to break the pattern of disadvantage' (Te Puni Kokiri, 2001). Therefore, the programme was integrated into the National Qualifications Framework, introduced in New Zealand in the early 1990's, with participants in the programme gaining unit standards from the Framework that could be built on to attain a recognised qualification, such as a National Certificate.

The eligibility criteria for the programme from 1993 to 1998 were slightly different for school leavers compared with all other potential participants. As far as school leavers were concerned, they needed to be eighteen or nineteen year olds with low qualifications who had left school in the last six months and were registered as

---

[5] No qualification was defined as having fewer than three School Certificate subjects and low qualification having no qualification higher than sixth form certificate. School Certificate was a national examination for year 11 students and sixth form certificate for year 12 students.

unemployed. Those who were not school leavers needed to have low qualifications, to have been registered as unemployed for at least twenty six weeks and be available to work at least twenty hours per week, although there was an exception for youth who needed to have been registered as unemployed for only thirteen weeks. Common to all potential participants was that they remained eligible to participate in the programme until they had earned 240 credits on the register of National Qualifications, whether inside or outside of the programme.

The programme was administered by Skill New Zealand, clients were referred by the New Zealand Employment Service and the average length of Training Opportunities courses was twenty one weeks. The training varied from foundation or generic courses that focused on developing employability skills including literacy, numeracy, English, communication, use of technology, decision making, information gathering and analysing, planning, organisation and problem solving, to vocational and industrial skill based courses and also work based options for learners who were close to be being work-ready in the view of the employment administrator. The approach for some was progressive, with those participants moving from foundation to vocational training. The training was mainly provided by New Zealand Qualification Authority registered and accredited Private Providers who were contracted by Skill New Zealand and who had to meet performance targets as agreed in their contracts. These were derived from targets set by the New Zealand government for Skill New Zealand that a participant should be in employment or further training outside of the programme two months after completing Training Opportunities and that educational outcomes include the achievement of credits from the Register of National Qualifications.

There have been several reviews and evaluations, both qualitative and quantitative, which have been used to judge the effectiveness of the programme and to refine its operation.[6] In the end, the Training Opportunities Programme has remained an important component of the current government's active labour market policy portfolio.

3. Evaluation Framework

The issue in evaluating active labour market policies is one of causal inference, where a cause is viewed as a manipulation or treatment that brings about a change in the outcome of interest compared to some baseline outcome measure (Dehejia and Wahba, 1998). This involves identifying a causal effect while controlling for confounding variables that also influence the outcome variable of interest. Since it is not possible to observe an individual simultaneously in both the treated and non-treated state, the causal effect of treatment is generally framed within a potential outcomes framework (Sianesi, 2001; and Lechner, 1999). This framework for treatment evaluation is used below to outline the evaluation problem and motivate the issue of sample selection bias. Difference-in-Differences matching is then outlined as a useful approach for mitigating both the evaluation problem and the selection bias issue.

---

[6] These include Education Training and Support Agency, N.Z. (1993), Nielsen (1999), Ministry of Education, N. Z. (2001), Te Puni Kokiri, N. Z. (2001), Mare (2002) and de Boer (2003).

The Evaluation Problem

Evaluation is concerned with how an individual's outcomes are altered as a result of an intervention; that is, as a result of participation or treatment in a programme. This impact is measured as the difference in the outcome for an individual in the non-treated state compared with the outcome in the treated state. The evaluation problem arises as it is not possible to observe an individual in both of these states. Here a potential outcomes framework allowing for heterogeneity of outcomes is utilised to address this issue

The estimator of interest in this study is the effect of participation in Training Opportunities on those who participate, known as the treatment of the treated estimator. This evaluation parameter relates to the question: 'What is the expected gain, on average, to individuals who receive treatment as compared with the counterfactual situation where they do not receive treatment?' Several methodological reviews (Heckman, LaLonde *et al.*, 1999 and Cobb-Clark and Crossley, 2002) have that this is estimated as follows:

$$E\left(\Delta_i | D_i = 1, c_{ai}\right) = E\left(Y_{1i} | D_i = 1, c_{ai}\right) - E\left(Y_{0i} | D_i = 0, c_{ai}\right) \tag{1}$$

where Δ_i is the change resulting from treatment for individual i, D_i is treatment status with 1 being with training and 0 without training, Y_{1i} and Y_{0i} are the outcomes with training and without training for individual I respectively and c_{ai} are the set of available explanatory variables.

The evaluation problem, therefore, is a missing data problem since $E\left(Y_{0i} | D_i = 1, c_{ai}\right)$ is unobserved and the researcher is required to identify the relevant counterfactual, $E\left(Y_{0i} | D_i = 0, c_{ai}\right)$ in equation 1, so as to estimate the causal impact of treatment. This counterfactual is the potential outcome for an individual in the state in which he or she is not observed.

Evaluation of TT (Treatment of the Treated), and many of the other parameters of interest, depends crucially on the assumption that the non-treated outcome for nonparticipants is the same as the non-treated outcome for participants. Using the outcomes of the non-treated as a proxy for the outcomes of the treated in the non-treated state gives:

$$E\left(Y_{1i} | D_i = 1, c_{ai}\right) - E\left(Y_{0i} | D_i = 0, c_{ai}\right) \tag{2}$$
$$= E\left(Y_{1i} - Y_{0i} | D_i = 1, c_{ai}\right) + \left[E\left(Y_{0i} | D_i = 1, c_{ai}\right) - E\left(Y_{0i} | D_i = 0, c_{ai}\right)\right]$$
$$= E\left(\Delta_i | D_i = 1, c_{ai}\right) + \left[E\left(Y_{0i} | D_i = 1, c_{ai}\right) - E\left(Y_{0i} | D_i = 0, c_{ai}\right)\right]$$

The second part of equation 2 identifies the importance of this assumption. If the assumption is valid, then the only reason for a difference in outcomes between the treated and untreated arises as a result of treatment on this base outcome $\left(E\left(Y_{0i} | D_i = 1, c_{ai}\right) - E\left(Y_{0i} | D_i = 0, c_{ai}\right)\right)$. If this term is equal to zero, then $E(\Delta_i)$ will be an unbiased estimator. However, if this outcome differs between the treated and untreated, then the right hand side in equation 2 is not equal to zero, selection bias exists and $E(\Delta_i)$ is a biased estimator. This may occur because there is a relationship between treatment status and the unobservable variables influencing the outcome.

Difference-in-Differences Matching

Matching is an estimation approach that assumes the observed variables define the sub-populations of participants and non-participants and any remaining differences can be attributed to chance (Schmidt, 1999). The impact of the treatment in these circumstances can then be estimated as the difference in the outcomes between the participant and non-participant groups. While matching has been widely used in the statistics literature, it is relatively new to economics (Dehejia and Wahba, 1998). However, the strong intuitive appeal of matching techniques is that the estimator resembles that used in ideal social experiments (Lechner, 1999; and Hujer and Caliendo, 2000). In this approach the evaluator aligns C_{it} so that the observed characteristics in the D_i=0 sample are the same as the D_i=1 sample. It has gained in popularity over the last decade as demonstrated by the number of studies that have been recently published using this technique.[7]

      The matching estimator is able to provide unbiased estimates of TT (Treatment of the Treated) given a number of important assumptions. These are that there are no general equilibrium effects arising from the training, that conditional on all covariates the outcomes are independent of assignment to treatment, known as the conditional independence assumption, and that for every participant there is a potential nonparticipant.

      A major problem that may arise with simple pair matching is dimensionality if the number of observable characteristics C_{it} is high (Dehejia and Wahba, 1998; Puhani, 1998; Lechner, 1999; and Sianesi, 2001). A solution to this issue is proposed by Rosenbaum and Rubin (1983), who suggest using the propensity score to reduce the dimensions of the matching problem and make it possible to match on a large number of covariates. The usefulness of this approach to matching is emphasised by the number of recent studies that have utilised various versions of propensity score matching (Heckman, *et al*. 1997; Dehejia and Wahba, 1998; Puhani, 1998; Lechner, 1999 and 2000; Brodaty, *et al*., 2001; Dehejia and Wahba, 2002; Lechner, 2002; and Dyke, *et al*., 2006).

      The propensity score is the conditional probability of assignment to a particular treatment given a vector of observed covariates. It is used to match participants and non-participants on their estimated probability of participation $P(C_{it})$, rather then on a vector of observed characteristics (Smith, 2000). Conditional on CIA (Conditional Independence Assumption) and the assumption that there is at least one untreated individual for every treated person, Rosenbaum and Rubin (1984) show that;

$$Y_{1it}, Y_{0it} \perp D_i \mid P\left(C_{it}\right) \tag{3}$$

      In other words, the dimensionality problem is reduced as the outcomes are independent of treatment given a single number represented by the propensity score. As with pair-wise matching, a weaker condition will suffice;

---

[7] See, for example, Card, D. and D. Sullivan (1988), Kluve, J., *et al*. (1999), Lechner, M. (1999), Eichler, M. and M. Lechner (2002), Kluve, J., *et al*. (2002), Mare, D. (2002), Hujer, R., *et al*. (2003), Meager, N., P. Bates, *et al*. (2003), Sianesi, B. (2003), Lechner, M., *et al*. (2004), Caliendo, M., *et al*. (2005) and Kluve, J., *et al*. (2005).

$$Y_{0it} \perp D_i \mid P\left(C_{it}\right) \tag{4}$$

conditioning on $P\left(C_{it}\right)$ eliminates the selection bias since

$$E\left(Y_{0it} | D_i = 1, P\left(C_{it}\right)\right) = E\left(Y_{0it} | D_i = 0, P\left(C_{it}\right)\right) = E\left(Y_{0it} | P\left(C_{it}\right)\right) \tag{5a}$$

$$E\left(Y_{1it} | D_i = 1, P\left(C_{it}\right)\right) = E\left(Y_{1it} | D_i = 0, P\left(C_{it}\right)\right) = E\left(Y_{1it} | P\left(C_{it}\right)\right) \tag{5b}$$

and the mean treatment impacts can be estimated without bias because we can use Y_{0it} for the comparison group as the counterfactual for the treated group.

Bias may arise should there be unobservables that influence participation or the outcome. The Difference-in-Differences (DID) estimator may overcome this issue by removing the effect of individual specific, time invariant unobservables. Should these be the only type of unobservables then selection bias is removed. The DID matching estimator using the propensity score requires that the difference between the outcome before (t) and after (t') the intervention time period for not participating is the same for both those who participate as for those who do not participate; that is

$$\mathrm{E}\left(Y_{0t} - Y_{0t'} | P, D = 1\right) = \mathrm{E}\left(Y_{0t} - Y_{0t'} | P, D = 0\right) = B \tag{6}$$

This condition requires that any bias (B) that exists in the time period before the intervention continues to exist in the time period after the intervention. These time specific intercepts, or fixed effects, may arise due to administrators consistently choosing participants over time based on unobserved permanent characteristics.

Following Smith and Todd (2005), the Difference-in-Differences matching estimator is given by

$$\alpha_{DDM} = \frac{1}{n_1} \sum_{i \in I_1 \cap s_p} \left\{ \left(Y_{1ti} - Y_{0t'i}\right) - \sum_{j \in I_0 \cap s_p} W(i,j)\left(Y_{1tj} - Y_{0t'j}\right) \right\} \tag{7}$$

The weights in the DID matching estimator $W(i,j)$ depend on the specific matching estimator that is chosen. For example, the DID estimator can be implemented using single nearest neighbour, multiple nearest neighbour, kernel or local linear regression, and the appropriate weighting formula would then be applied.

The use of DID estimators in matching began only in the late 1990's (Heckman, *et al.*, 1997; Heckman, *et al.*, 1998; Eichler and Lechner, 2002; and Smith and Todd, 2005). Research by Heckman, *et al.* (1998) finds that, when compared with experimental estimates, and in the presence of the influence of unobserved variables that influence participation, the DID estimator outperforms other matching estimators.

4. Data

The data for this paper were primarily obtained from the Labour Market Policy Group, New Zealand Department of Labour.[8] They come from two administrative data sets, an

---

[8] In July 2004 the Labour Market Policy Group was disbanded as the Department of Labour changed its strategic direction and reorganised.

enrolment dataset and an intervention dataset compiled from various New Zealand Employment Service (NZES) data sources, which had been collected by the NZES[9] between 1 October 1988 and 31 December 1997. This was augmented with regional data compiled by MOTU Economic and Public Policy Research Trust through their FRST-funded programme 'Understanding Adjustment and Inequality'. This section describes these data in general, are the specific dataset created to undertake the matching evaluation.

Detail, Issues and Adjustments

The Department of Labour enrolment dataset contains demographic, economic and labour market information on clients who were registered with NZES as unemployed at any time between 1 October 1988 and 31 December 1997.[10] There are 2,476,898 spells of unemployment from 1,145,168 different clients in this general dataset. The intervention dataset contains details of all interventions for NZES clients between 1 October 1988 and 31 December 1997. Each time an intervention occurs, 3,652,222 interventions in all, there is an entry in the dataset.[11] The two datasets are connected by a unique identifier.

Since there is no information on family status and number of dependent children in these datasets, which may be important influences on female involvement in the labour force, only males are included in this study. Further, due to the lack of information on participation in formal education and movement into retirement, only males who were aged between 26 and 49 on 1 January 1989 in the combined datasets are evaluated.

A major issue when combining the datasets was to identify the appropriate solution to the problem of interventions without an end date. This issue is exacerbated as there is right censoring in the data, since the interventions and unemployment spell information ends as of 31 December 1997 when either or both of these may be ongoing. There are two main options open to the researcher in these circumstances. One option is to identify a mean length of intervention and to substitute that for the missing end date. The other is to remove the observations from the dataset, treating them as erroneous data. The first option, substituting in the mean length of intervention, was used in a recent study using this dataset (Mare, 2002). However, there are a number of issues associated with this procedure. Of particular concern is how to compute the mean length of intervention, as for many of the interventions there was a wide variety of intervention lengths. For example, should the mean length of an intervention be

---

[9] The New Zealand Employment Service (NZES), part of the Department of Labour, maintained the register of all unemployed over the duration of the data set and was also responsible for administering many of the ALMPs. Unemployment benefits were administered by Income Support, part of the Department of Social Welfare. In 1998 NZES was integrated with Income Support to form Work and Income New Zealand (WINZ) and in 2001 WINZ became part of the Ministry of Social Development that had been established in 2000.

[10] The following variables are included in the dataset each time a client had an unemployment spell: start date of the spell, end date of the spell, a unique client number, date of birth, gender, ethnicity, highest educational qualification, reason for leaving the register, office at which the client is registered, preferred occupation, barriers to employment and hours available to work.

[11] The following variables are included in the dataset each time an intervention occurred: the office which manages the client, start date of the intervention, a unique client number which is the same as that in the enrolment dataset, end date of the intervention, the type of intervention and the immediate result of the intervention.

calculated for the whole sample which undertook an intervention, or calculated for each office or conditional on a set of personal characteristics for participants? The mean length obtained under each of these scenarios would be different and potentially unrepresentative what actually occurred. This could bias our estimates, and there would be no way of knowing, a priori, in what direction the bias would go. Further, some of the participants may have dropped out of the programme prior to completion. In these circumstances, including the mean length would negatively bias the estimated impact of the intervention. There is also a risk involved in the second option, removing individuals who do not have an end date, as this assumes that they did not participate in the programme. However, given the size of the sample, and the fact that the risk of bias is potentially greater from using the estimated mean length of intervention it was decided to delete individual observations from the data set that had an intervention without an end date.

When all adjustments had been completed, including removing individuals with inconsistent data and a lack of end dates to interventions, the number of males aged 26 to 49 on 1 January 1989 in the dataset fell from 257,537 to 247,507 males, representing a 3.8 per cent in individual observations.

The master dataset for this research involved establishing a continuous time daily schedule for all of the 247,507 individuals. On each day, it was possible to identify whether an individual was on or off the register, if on the register whether they were in an intervention or not, and if in an intervention the type of intervention. The decision was made to measure only the major interventions and not to include referrals to jobs or the ongoing interviews and seminars that were part of the regular operation of the New Zealand Employment Service. It was possible that NZES clients were undertaking more than one intervention on any given day. For example, they could be receive a subsidy and also undergo training. The continuous time structure of the master dataset made it possible then to aggregate the key variables of interest as proportions of the time period required, for example a year, or a quarter or a month. The dependent variable is the proportion of time on the unemployment register and the key independent variables are the proportion of time in an intervention or in specific interventions.[12]

Cohort Dataset
Difference-in-differences matching requires pre-intervention and post-intervention data for both participants and non-participants. Pre-intervention data are required to generate the propensity score for this matching. Post-intervention data are required to estimate the effects of the active labour market programmes. In order to achieve this, a 'cohort dataset' was created. This is in line with approaches adopted by several researchers since 2000 (Aakvik, *et al*., 2000; Conniffe, *et al*., 2000; Magnac, 2000; Angrist and Lavy, 2001; van Ours, 2001; Bolvig, *et al*., 2002; Bratberg, *et al*., 2002; Gerfin, *et al*., 2002; Mare, 2002; O'Connell and McGinnity, 2002; Raaum and Torp, 2002; and Regner, 2002).

---

[12] The other variables apart from the dependent variable and intervention variables in the master dataset include age, ethnicity, unemployed and no intervention, barriers to employment, year and New Zealand Employment Service office location.

In this study, two intervention years (1993 and 1994) are selected for the cohort. For the 1993 group, the pre-intervention data cover the four-period 1989 to 1992, and the post-intervention data cover the three-year period 1994 to 1996. For the 1994 group, the pre-intervention data come from the four-year period 1990 to 1993, and the post-intervention data come from the three-year period 1995 to 1997. The creation of a single cohort from these two periods requires a concept of time based around the year of intervention rather than calendar time. This is achieved by identifying the year of intervention as time t, the pre-intervention years as t-4, t-3, t-2, and t-1, and the post-intervention years t+1, t+2 and t+3. In this way, the variables containing the data for the pre-intervention, intervention and post-intervention years are all aligned (see table 3).

Table 3 - Establishing Time in the Cohort Dataset

	Pre-Intervention				Intervention	Post-Intervention		
Group	t-4	t-3	t-2	t-1	t	t+1	t+2	t+3
1	1989	1990	1991	1992	1993	1994	1995	1996
2	1990	1991	1992	1993	1994	1995	1996	1997

To be included in this cohort dataset, a participant must have received an intervention in either 1993 or 1994 and must not have received another intervention during the other seven years. The comparison group in the cohort dataset consists of those who were unemployed in the period preceding the time of potential intervention, a key determinant of participation in active labour market programmes, but who never received an intervention at any time from 1989 through 1997. In total, there are 39,275 males in the cohort dataset of which 879 participated in Training Opportunities and the balance received no intervention.

There are some noticeable differences in the characteristics of those who participated in training opportunities compared with those who did not (see table 4). While the mean age at time t is similar between the two groups, the age distribution varies. Training Opportunities participants are generally younger and have lower qualification levels. Maori and Pacific Islanders are overrepresented among Training Opportunities participants. The unemployment pattern also differs between the two groups. While both groups experienced a growth in proportion of the year registered as unemployed up to time t and a decline thereafter, the magnitudes are greater for those who participated in Training Opportunities. A further difference is that a relatively greater percentage of participants were long-term unemployed.

Table 4 - Summary Statistics

Variable	Total Sample	Training Opportunities	No Intervention
Age (as at Year of Intervention)			
Mean (yrs)	39.4	38.7	39.5
Age Groups (%):			
Less then 40	41.5	44.9	41.4
40-49 years	39.6	40.0	39.5
50+	18.9	15.1	19.1
Education (%):			
No Formal Qualification	44.9	69.1	44.4
School Qualification	33.5	22.1	33.8
Post School Qualification	21.6	8.8	21.9
Ethnicity (%):			
European/Pakeha	70.3	37.2	71.1
Maori	16.9	31.1	16.6
Pacific Islander	6.8	21.8	6.4
Other	6.0	9.9	5.9
Unemployment Pattern (%):			
Registered U t-4	17.5	27.1	17.3
Registered U t-3	18.5	35.5	18.1
Registered U t-2	20.2	48.3	19.6
Registered U t-1	19.2	68.1	18.0
Registered U t	15.7	77.6	14.2
Registered U t+1	9.2	36.7	8.5
Registered U t+2	5.4	19.3	5.0
Registered U t+3	3.8	12.4	3.6
Registered Unemployed Pre-t	18.8	42.3	18.2
Registered Unemployed Post-t	6.1	22.8	5.7
Short Term Unemployed (at t)	82.9	30.3	84.1
Long Term Unemployed (at t)	17.1	69.7	15.9
Region Population Size(%):			
Population $\leq$ 25,000	10.7	8.5	10.7
25000 < Population $\leq$ 100,000	41.9	42.4	41.9
Population > 100,000	47.4	49.1	47.4
% of Sample	100	2.2	97.8
Number of Observations	39,275	879	38,396

5. Analysis

This study evaluates the impact of male participation in Training Opportunities on the propensity to be registered as unemployed using DID matching on the propensity score with the nearest neighbour replacement estimator. Implementing this approach requires a number of aspects to be addressed. The first is estimating the propensity scores for participants and non-participants. The second is testing whether the model specification is robust and the covariates are balanced. The third involves sensitivity analysis.

Estimation of the Propensity Score

There has been much debate in the literature over the variables that should be included in the participation equation. Apart from the need to include all those variables that influence both participation and subsequent outcomes, there is no real consensus yet over this issue (Smith and Todd, 2005). The difficulty is that there is no generally accepted formal approach for identifying these variables, and failure to include all the

relevant variables violates the conditional independence assumption. Lechner (1999) suggests that these covariates can be chosen without developing a formal behavioural model, but rather simply by considering the broad 'building blocks' of this behaviour. The decisions as to which variables to include, therefore, should be based on the processes through which programme participation decisions are made. This is the major approach used in the literature and the one utilised here.

Participation in active labour market programmes generally involves the input of both the employment advisor and unemployed individual (Lechner, 1999; and Sianesi, 2003). Both dimensions to this decision need to be considered. The employment advisor potentially has an impact on the allocation of individuals to active labour market programmes in New Zealand. The employment advisor selects clients for programmes based not only on their unemployment histories, but also on subjective judgements over whether or not there are realistic chances of positive outcomes from these interventions. Factors in the decision-making process include eligibility rules, unemployment and labour market histories, age, education and other capabilities.

The decision of whether or not to participate in a active labour market programme for an unemployed individual might be based on personal assessments of the costs and benefits associated with programme participation. Demographic and human capital factors, labour market histories and external factors may influence these assessments. A number of demographic factors have been used in other matching studies, including age, ethnicity, marital status and number of children. Age and ethnicity are included in our dataset. Human capital characteristics are also potentially important as they provide information on the educational experiences and capabilities of the unemployed client. Measures of levels of education and qualifications, including work experience and training are often included as variables that influence participation (Sianesi, 2003).

Our participation model includes variables on age, age squared, dummy variables for educational qualifications and minority ethnic status, unemployment in the year of potential intervention t, unemployment in the preceding year t-1, unemployment history in the full pre-intervention period t-4 to t-1, whether or not an individual is long term unemployed prior to time $t,$ dummy variables for each region (except the base or excluded region), dummy variables for regional population size (with the omitted region being one with a population of less then 25,000) and real regional growth rates.

In the literature on matching, either probit or logit estimation is used to produce propensity score estimates for participation. There is no reason a priori to prefer one estimation technique over the other. Given that probit estimation is used in this study, the propensity score can be written:

$$\Pr(y_{it} = 1|x_j) = \Phi(\mathbf{x}_j\mathbf{b}) \tag{8}$$

where $\Phi(.)$ is the standard normal cumulative density function. Rather than reporting the parameter estimates from these regressions, we report the estimated partial derivatives evaluated at the sample means of the covariates, and their standard errors in the tables to follow. For a continuous explanatory variable, this partial derivative can be written:

$$\frac{\partial \Phi}{\partial x_1} = \phi(\mathbf{x}_j\mathbf{b})b_1 \tag{9}$$

The probit results for participation in Training Opportunities show that most of the variables are statistically significant and have the expected signs. The groups for whom the impact of participation in training is estimated, notably Maori and Pasifika ethnic groups and those with low school attainment, are more likely to participate in Training Opportunities.

Model Specification

There are two aspects of the model that need to be addressed to ascertain its robustness. The first is to determine whether the model specification is conceptually appropriate, while the second is to identify whether the matching approach balances the covariates of the matched participants and non-participants.

As mentioned earlier, the efficacy of the DID matching specification depends on the factors influencing participation and outcomes being either observable, or if they are unobservable, time-invariant. A test of the specification, as posited by Heckman and Hotz (1989) and used by Dyke *et al*. (2006), exploits the pre-training data. The idea is to test whether participation in Training Opportunities has any influence on pre-training unemployment outcomes. Should there be a non-zero impact in the pre-training period this suggests that the model has not accounted for all of the factors influencing participation and the outcome. The results of this test for each of the models estimated are shown in table 5. The only model where there is a statistically significant effect of participation in Training Opportunities in the pre-training period is for regions with a population greater than 100,000 at *t*-2. Therefore, the model appears to be robust.

Table 5 - Test of Model Specification pre *t*

Category	t-4 Coefficient	SE	t-3 Coefficient	SE	t-2 Coefficient	SE
Total	0.0187	0.0209	-0.0078	0.0236	-0.0432	0.0236
Age at Intervention:						
Age < 40 years	0.0048	0.0124	-0.0086	0.0284	0.0007	0.0339
40 years ≤ Age < 50 years	0.0236	0.0323	0.0064	0.0316	-0.0400	0.0359
Age ≥ 50 years	-0.0289	0.0394	-0.0088	0.0589	0.0058	0.0488
Education:						
No Qualification	-0.0020	0.0264	-0.0346	0.0263	-0.0457	0.0242
School Qualification	0.0470	0.0368	0.0519	0.0421	-0.0206	0.0478
Post School Qualification	0.0522	0.0557	-0.0130	0.0642	-0.0541	0.0712
Ethnicity:						
Pakeha/European	0.0399	0.0328	0.0553	0.0340	0.0128	0.0370
Maori	0.0517	0.0407	0.047	0.0373	-0.0107	0.0345
Pacific Islander	0.0041	0.0450	-0.0349	0.0475	-0.0453	0.0456
Other	-0.0178	0.0459	0.0125	0.0620	-0.0666	0.0764
Length of Time Unemployed:						
Unemployed < 26 weeks	0.0294	0.0269	0.0229	0.0244	-0.0103	0.0255
Unemployed ≥ 26 weeks	0.0289	0.0253	0.0182	0.029	-0.0272	0.0238
Population of Local Labour Market:						
Population ≤ 25,000	-0.0096	0.0755	0.0054	0.0687	-0.0091	0.0721
25,000 < Population ≤ 100,000	0.0018	0.0288	-0.0048	0.0313	-0.0184	0.0373
Population > 100,000	-0.0071	0.0261	-0.0382	0.0312	-0.0796*	0.0296

* significant 0.05. ** significant 0.01.

The second aspect of model specification is to consider whether or not the covariates of participants and non-participants are balanced. The bias, standard deviation and *t*-test for each variable in the matched participant and non-participant groups are estimated. In addition, the before-and-after matching mean bias for the sample groups as a whole and the largest and smallest bias are identified. These summary results are presented in table 6. The bias between the participant and nonparticipant groups decreases greatly after matching. This is demonstrated by the decrease in the magnitude of the mean bias and the size of the largest and smallest bias across the programmes. This analysis indicates that while matching techniques reduce imbalance, they do not remove it completely.

Table 6 - Balancing Analysis –Training Opportunities

	Unmatched			Matched		
Intervention	*Mean Bias Unmatched*	*Smallest Bias Unmatched*	*Largest Bias Unmatched*	*Mean Bias Matched*	*Smallest Bias Matched*	*Largest Bias Matched*
Training Opportunities	17.747	0.0612	240.21	3.491	0.0000	10.61

Impact Estimation

Estimation of the impact arising from participation in active labour market programmes using DID matching requires two further issues to be addressed. The first is accounting for the extra variance that occurs as a result of the matching process, with the standard approach in the literature being to use bootstrapped standard errors. This approach is used here and is implemented with 100 iterations, as there was minimal variation in the outcome once the iterations increased past 50.

The second issue involves the possibility of an 'Ashenfelter's dip'. Previous studies have found that in the period immediately prior to programme participation there is sometimes a decline in the outcome variable of interest (Ashenfelter, 1977; and Ashenfelter and Card, 1985). Upward bias in the estimated impact may arise should the before/after comparison use the period in which the dip occurs in the measure of the outcome before the intervention. Should the dip be a permanent effect then there is no bias. However, should the dip be transitory then bias will result in the estimate. Although the initial work on Ashenfelter's dip focused on wages and earnings, subsequent research indicated that there is also evidence of an increase in the propensity to be unemployed in the lead up to participation in a programme (Card and Sullivan, 1988; Heckman and Smith, 1999; and Bergemann, *et al*., 2005).

The usual approach to this problem is to choose an initial time period prior to the dip. Given that the eligibility criteria for the Training Opportunities Programme include being unemployed for greater than 13 weeks, we exclude the year immediately prior to the intervention. The pre-intervention measure of unemployment is, therefore, the average propensity to be unemployed in years *t*-2, *t*-3 and *t*-4.

The DID matching estimator identifies the impact of male participation in Training Opportunities on the proportion of a year that males spend registered as unemployed. For example, table 7 shows that the estimate at *t*+1 is -0.07113. The

Table 7 - Estimated Effects of Training Opportunities Using One-to-One DID Nearest Neighbour Estimation

Category	t		t+1		t+2		t+3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Total	-0.0105	0.0205	-0.0711**	0.0255	-0.0156	0.0235	0.0180	0.0242
Age at Intervention:								
Age < 40 years	-0.0241	0.0292	-0.0827*	0.0405	-0.0253	0.0363	0.0045	0.0364
40 years ≤ Age < 50 years	-0.0187	0.0257	-0.0340	0.0448	0.0064	0.0371	0.0183	0.0362
Age ≥ 50 years	-0.0162	0.0498	-0.0870	0.0714	-0.0389	0.0657	0.0319	0.0645
Ethnicity:								
Pakeha/European	-0.0590	0.0308	-0.1357**	0.0439	-0.0307	0.0393	-0.0203	0.0363
Maori	-0.0379	0.0339	-0.0645	0.0458	-0.0070	0.0447	-0.0096	0.0449
Pasifika	-0.0025	0.0443	-0.1048	0.0583	-0.0172	0.0564	0.0482	0.0532
Other	-0.0035	0.0423	-0.0313	0.0681	0.1156	0.0717	0.1023	0.0670
Education:								
No formal Qualification	0.0078	0.0256	-0.0707*	0.0352	-0.0115	0.0331	0.0479	0.0320
School Qualification	-0.0190	0.0410	-0.0899	0.0516	0.0088	0.0453	0.0270	0.0487
Post-School Qualification	0.0028	0.0637	-0.0909	0.0894	-0.0194	0.0809	0.0432	0.0836
Length of Time Unemployed:								
Unemployed < 26 weeks	-0.0346	0.0215	-0.1228**	0.0353	-0.0362	0.0349	-0.0256	0.0297
Unemployed ≥ 26 weeks	-0.0212	0.0261	-0.0648*	0.0325	-0.0324	0.0312	0.0025	0.0290
Population of Local Labour Market:								
Population < 25,000	0.0096	0.0648	-0.0450	0.0915	0.0041	0.0867	-0.0404	0.0499
25,000 ≤ Population < 100,000	-0.0092	0.0296	-0.0556	0.0418	-0.0207	0.0388	-0.0023	0.0389
Population ≥ 100,000	-0.0106	0.0308	-0.1063*	0.0357	0.0342	0.0314	0.0807*	0.0301

* significant 0.05. ** significant 0.01.

interpretation is that, on average, males participating in Training Opportunities reduce the proportion of the year registered as unemployed by 0.07113 (or 29.96 days) in the year following the intervention. A negative sign indicates a beneficial outcome as less time is spent registered as unemployed after the intervention, while a positive sign indicates a detrimental outcome. In this study, the impact is estimated at t, $t+1$, $t+2$ and $t+3$, making it possible to address a number of questions including contemporaneous, short term, and medium term impacts from male participation in Training Opportunities as a whole and for sub-groups.

The estimates for t indicate that there is no statistically significant locking-in effect, whereby participants decrease the likelihood of moving off the unemployment register as they become committed to the programme. Most of the signs are negative, suggesting a beneficial contemporaneous effect. However, none of the estimates is statistically significant. In the year after the intervention, year $t+1$, there is a statistically beneficial impact on participants in Training Opportunities, as pointed out above. However, this benefit is short term. At $t+2$, the estimated effect is negative, but statistically insignificant. At $t+3$, the estimated effect is positive, but also statistically insignificant. This indicates that there is only a transitory benefit from participation in the programme. This is also true for each of the sub-groups. For those sub-groups who receive a beneficial impact from participating in Training Opportunities, it only occurs at $t+1$ and disappears thereafter. The exception is in $t+3$ where there is a statistically significant detrimental impact for those living in centres with more than 100,000 people. However, as pointed out in the specification analysis, this is the one model where the specification failed to control for differences between those receiving the intervention and matched non-participants. The result for the programme differ to some extent from those that Mare (2002) finds, as his study identified a statistically negative effect from participating in the programme whereas this one finds a short term beneficial effect that then disappears. Although this study uses the same data as that used by Mare, the advantage here is that time-invariant, unobserved individual specific effects are taken into account through the DID approach.

There are some differential impacts on sub-groups at $t+1$, the year after participating in Training Opportunities. Males aged 40 or less experienced a beneficial impact from training, whereas those older than 40 did not. Men with no formal qualifications experienced a gain from participation in the programme, whereas those with at least a school qualification did not. This is consistent with the targeting of the programme at those who are most disadvantaged. The strongest impact on any sub-group is for Pakeha/European males whose time registered as unemployed decreases by 49.53 days in $t+1$.

For Maori and Pacific Islanders, there is no statistically significant effect. This is an interesting result as they represent a key government target group who have higher levels of unemployment. One possible explanation is that the definition used to identify Maori in the administrative database is that if an individual has some Maori ancestry then this dominates in the classification. Research by Chapple (1999) indicates that this definition may be too general and that a finer definition that separates out those who share a common ancestry with Pakeha from those who have a stronger Maori ancestry has a stronger explanatory power of inequality. A disaggregation of

the Maori group along these lines for estimating the impact of participation in Training Opportunities is a useful direction for further study.

The programme has an effect both on those who are long-term unemployed and those who have been unemployed for less than 26 weeks. The impact on the latter is twice that for the long-term unemployed. Participation in Training Opportunities has a statistically beneficial effect on participants in regions with a population greater than 100,000 but not in regions which are smaller.

The overall impact of male participation in Training Opportunities is transitory, since the beneficial effect does not extend beyond the year after intervention. For sub-groups some of the governments target groups, Maori and Pacific Islanders, do not have a beneficial impact while for the long-term unemployed there is a benefit but it is less than that for those who are unemployed for less than 26 weeks.

The estimates reported above were calculated using the one-to-one nearest neighbour estimator. In order to assess the sensitivity of these results to the chosen specification, the model was re-calculated using different estimators. Five and fifteen nearest neighbour and local linear regression estimators were used, with the results indicating the difference between these estimates to those obtained using the one-toone estimator is in the magnitude of the standard errors, indicating that the results are not sensitive to the estimator used.

6. Conclusion

The Training Opportunities Programme is designed to provide individuals with the skills necessary to find employment and therefore to exit from the unemployment register. It is both the primary training programme and single largest active labour market programme in New Zealand. In light of this, the impacts from participation estimated here are somewhat disappointing. The beneficial effect is transitory and not experienced by key government target groups, including Maori and Pacific Islanders.

This raises a couple of questions. Firstly, what might explain these results? It could be that primarily classroom-based programmes, such as Training Opportunities, may be less effective than work-based programmes. A related issue is the extent to which the programme provides generic or specific skills. There needs to be further evaluation of the elements of Training Opportunities, as it covers a range of generic and specific training initiatives. It may well be that certain elements have longer lasting beneficial effects but that these are counterbalanced in the estimation by those that have detrimental effects. A further possible reason for the results is that the training is too short, since the average length of time on the programme is 26 weeks. This is a limited investment for individuals who are disadvantaged in the labour market. With respect to training programmes in the United States, the

> …best summary of the evidence about the impact of past programmes is that we got what we paid for. Public sector investments in training are exceedingly modest compared to the magnitude of the skill deficiencies that policy makers are trying to address. Not surprisingly modest investments yield modest gains. (LaLonde, 1995, p. 149)

The second question arising from the results is why does the government continue to fund the programme in its current form? The answer may be that the government has not completed the analysis required in order to inform its decision making. This may well be the case since there has been relatively little formal economic evaluation of active labour market programmes in New Zealand. It may be that there are political reasons why it is useful for the government to continue to run the programme. It could also be that, even though the beneficial impact is of short duration, overall there is a positive net fiscal impact from running the programme. This is an avenue for further investigation.

This paper has also contributed to the clarification of methodological issues. The relevance of estimating the impact of programmes beyond the immediate post-intervention time period has been demonstrated. Further, the importance of paying careful attention to model specification, balancing and sensitivity analysis is a feature of the paper. Attention to these issues increases the confidence in the estimated results.

References

Aakvik, A., Heckman, J. and Vytlacil, E. (2000), Treatment Effects for Discrete Outcomes when Responses to Treatment Vary among Observationally Identical persons: An Application to Norwegian Vocational Rehabilitation Programs, NBER.

Angrist, J. and Lavy, V. (2001), 'Does Teacher Training Affect Pupil Learning? Evidence from Matched Comparisons in Jerusalem Public Schools.' *Journal of Labor Economics,* 19(2), 343-369.

Ashenfelter, O. (1977), 'Estimating the Effect of Training Programs on Earnings.' *The Review of Economics and Statistics,* 60(1), 47-57.

Ashenfelter, O. and Card, D. (1985), 'Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs', *The Review of Economics and Statistics*, 67(4), 648-660.

Barnow, B. (1987), 'The Impact of CETA Programs on Earnings: A Review of the Literature', *Journal of Human Resources,* 22(2), 157-193.

Bergemann, A., Fitzenberger, B. and Speckesser, S. (2005), Evaluating the Dynamic Employment Effects of Training Programs in East Germany Using Conditional Difference-in-Difference, Society of Labor Economics and European Association of Labour Economics World Congress, San Fransisco, California.

Bolvig, I., Jensen, P. and Rosholm, M. (2002), The Employment Effects of Active Social Policy in Denmark, Fourteenth Annual Conference of European Association of Labour Economists Conference, Paris.

Bratberg, E., Grasdal, A. and Risa, A. (2002), 'Evaluating Social Policy by Experimental and Nonexperimental Methods', *Scandinavian Journal of Economics*, 104(1), 147-171.

Brodaty, T., Crepon, B. and Fougere, D. (2001), 'Using Matching Estimators to Evaluate Alternative Youth Employment Programs: Evidence from France, 1986-1988', in M. Lechner and F. Pfeiffer, (eds.), Econometric Evaluation of Labour Market Policies, New York, Physica-Verlag, 85-123.

Caliendo, M., Hujer, R. and Thomsen, S. (2005), The Employment Effects of Job Creation Schemes in Germany - A Microeconomic Evaluation, Society of Labor Economics and European Association of Labour Economics World Congress 2005, San Fransisco, California.

Card, D. and Sullivan, D. (1988), 'Measuring the Effect of Subsidised Training Programs on Movements In and Out of Employment', *Econometrica*, 56(3), 497-530.

Chapple, S. (1999) 'Explaining Patterns of Disparity between Maori and Non-Maori Employment Chances', *Labour Market Bulletin*, (2)1, 70-100.

Cobb-Clark, D. and Crossley, T. (2002). Econometrics for Summative Evaluations: An Introduction to Recent Developments, Australia National University.

Conniffe, D., Gash, V. and O'Connell, P. (2000), 'Evaluating State Programmes: 'Natural Experiments' and Propensity Scores', *The Economic and Social Review*, 31(4), 283-308.

de Boer, M. (2003), Impact of Training Opportunities (1996-2001) on Participants' Outcomes, Ministry of Social Development, New Zealand.

Dehejia, R. and Wahba, S. (1998), Propensity Score Matching Methods for Non Experimental Causal Studies, Cambridge, MA, NBER.

Dehejia, R. and Wahba, S. (2002), 'Propensity Score Matching Methods for Non Experimental Causal Studies', *The Review of Economics and Statistics,* 84(1), 151-161.

Dyke, A., Heinrich, C., Troske K. and Jeon, K. (2006), 'The Effects of Welfare-to Work Program Activities on Labor Market Outcomes', *Journal of Labour Economics,* 24(3), 567-607.

Education Training and Support Agency, N.Z. (1993), Training Opportunity Programme: Future Directions, Wellington, Education Training and Support Agency National Office.

Eichler, M. and Lechner, M. (2002), 'An Evaluation of Public Employment Service Programmes in the East German State of Sachsen-Anhalt', *Labour Economics*, 9(2), 143-186.

Friedlander, D., Greenberg, D. and Robins, P. (1997), 'Evaluating Government Training Programs for the Economically Disadvantaged', *Journal of Economic Literature,* 35(4), 1809-1855.

Gerfin, M., Lechner, M. and Steiger, H. (2002), Does Subsidised Temporary Employment Get the Unemployed Back to Work? An Econometric Analysis of Two Different Schemes, University of St. Gallen, Department of Economics.

Heckman, J. and Hotz, V.J. (1989), 'Choosing Among Alternative Non-experimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Planning', *Journal of the American Statistical Association,* 84(408), 862-874.

Heckman, J., Ichimura, H. and Todd, P. (1997), 'Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme', *Review of Economic Studies*, 64, 605-654.

Heckman, J., Ichimura, H. and Todd P. (1998), 'Matching as an Econometric Evaluation Estimator', *Review of Economic Studies,* 65(223), 261-294.

Heckman, J., LaLonde, R. and Smith, J. (1999), 'The Economics and Econometrics of Active Labor Market Programs', in O. Ashenfelter and D. Card, (eds.), Handbook of Labor Economics, 3, New York, Elsevier Science, 1865-2097.

Heckman, J. and Smith, J. (1999), 'The Pre-Programme Earnings Dip and the Determinants of Participation in a Social Programme: Implications for Simple Programme Evaluation Strategies', *The Economic Journal,* 109(457), 313-348.

Hujer, R. and Caliendo, M. (2000), Evaluation of Active Labour Market Policy: Methodological Concepts and Empirical Estimates, IZA.

Hujer, R., Caliendo, M. and Thomsen, S. (2003), New Evidence on the Effects of Job Creation Schemes in Germany - A Matching Approach with Threefold Heterogeneity, IZA.

Kluve, J., Lehmann, H. and Schmidt, C. (1999), Active Labour Market Policies in Poland: Human Capital Enhancement, Stigmatization or Benefit Churning? London, Centre for Economic Policy Research.

Kluve, J., Lehmann, H. and Schmidt, C. (2005), Disentangling Treatment Effects of Active Labour Market Policies: The Role of Labour Force Status Sequences, Society of Labor Economics and European Association of Labour Economics World Congress 2005, San Fransisco, California.

Kluve, J., Lehmann, H. and Schmidt, C. (2002), Disentangling Treatment Effects of Polish Active Labour Market Policies: Evidence from Matched Samples, Fourteenth Annual Conference of European Association of Labour Economists, Paris.

LaLonde, R. (1995), 'The Promise of Public Sector-Sponsored Training Programs.' *Journal of Economic Perspectives*, 9(2), 149-168.

Lechner, M. (1999), 'Earnings and Employment Effects of Continuous Off-the-Job Training in East Germany after Reunification', *Journal of Business and Economic Statistics,* 17(1), 74-90.

Lechner, M. (2000), 'An Evaluation of Public Sector Sponsored Continuous Vocational Training Programs in East Germany', *Journal of Human Resources,* 35, 347-375.

Lechner, M. (2002), 'Program Heterogeneity and Propensity Score Matching: An Application to the Evaluation of Active Labour Market Policies', *The Review of Economics and Statistics*, 84(2), 205-220.

Lechner, M., Miquel, R. and Wuncsh, C. (2004), Long-Run Effects of Public Sector Sponsored Training in West Germany, Department of Economics, University of St Gallen.

Magnac, T. (2000), 'Subsidised Training and Youth Employment: Distinguishing Unobserved Heterogeneity from State Dependence in Labour Market Histories', *The Economic Journal,* 110(October), 805-837.

Mare, D. (2002), 'The Impact of Employment Policy Interventions', *Labour Market Bulletin*, 2000-02 Special Issue, 57-100.

Meager, N., Bates, P. and Cowling, M. (2003), 'An Evaluation of Business Start-Up Support for Young People,' *National Institute Economic Review*, 186, 59-72.

Ministry of Education, N.Z. (2001), Review of Training Opportunities and Youth Training: Issues Paper, Ministry of Education, New Zealand.

New Zealand Employment Service (1996), The Barriers to Employment Faced by Longterm Job Seekers: Key Findings from a Survey of New Zealand Employment Service Clients. Wellington, NZES: Market research and Evaluation Team.

Nielsen, A. (1999), Training Opportunities: Research into Longer Term Outcomes, Wellington, Skill New Zealand.

O'Connell, P. and McGinnity, F. (2002), 'Active Labour Market Policies, Market Orientation and Gender: Findings for Young People in Ireland', in K. Schomann and P. O'Connell, (eds.) Education, Training and Employment Dynamics: Transition Labour Markets in the EU, Cheltenham, Edward Elgar, 101-118.

OECD (1992), OECD Employment Outlook 1992, Paris, OECD.

OECD (1993), Economic Survey of New Zealand, Paris, OECD.

OECD (1996), Economic Survey of New Zealand 1995/1996, Paris, OECD.

OECD (2000), OECD Employment Outlook 2002, Paris, OECD.

Puhani, P. (1998), Advantage Through Training? A Microeconometric Evaluation of the Employment Effects of Active Labour Market Programmes in Poland. London, Centre for Economic Policy Research.

Raaum, O. and Torp, H. (2002), 'Labour Market Training in Norway - Effect on Earnings', *Labour Economics,* 9(2), 207-247.

Regner, H. (2002), 'A Non-experimental Evaluation of Training Programmes for the Unemployed in Sweden', *Labour Economics*, 9(2), 187-206.

Robinson, P. (2000), 'Active Labour-Market Policies: A Case of Evidence-Based Policy Making?' *Oxford Review of Economic Policy,* 16(1), 13-26.

Rosenbaum, P. and Rubin, D. (1984), 'Reducing Bias in Observational Studies Using Sub-classification on the Propensity Score', *Journal of the American Statistical Association*, 79(387), 516-524.

Rosenbaum, P. and Rubin, D. (1983), 'The Central Role of the Propensity Score in Observational Studies for Causal Effects', *Biometrica*, 70(1), 41-55.

Schmidt, C. (1999), Knowing What Works: The Case for Rigorous Program Evaluation, IZA.

Sianesi, B. (2001), An Evaluation of the Active Labour Market Programmes in Sweden, IFAU.

Sianesi, B. (2003), Differential Effects of Swedish Active Labour Market Programs for Unemployed Adults in the 1990s. London, The Institute for Fiscal Studies.

Smith, J. (2000), 'A Critical Survey of Empirical Methods for Evaluating Active Labour Market Policies.' *Schweizerische Zeitschrift fur Volkswirtschaft und Statistik*, 136(3), 1-22.

Smith, J. and Todd, P. (2005), 'Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators?' *Journal of Econometrics*, 125(1-2), 305-353.

Stanley, M., Katz, L. and Kreuger, A. (1998), Developing Skills: What We Know about the Impacts of American Employment and Training Programs on Employment, Earnings and Educational Outcomes, Cambridge, Harvard University, John F. Kennedy School of Government.

Te Puni Kokiri, N.Z. (2001), Training Opportunities and Youth Training. Wellington, Te Puni Kokiri, Ministry of Maori Development, New Zealand.

van Ours, J. (2001), 'Do Active Labour Market Policies Help Unemployed Workers to Find and Keep Regular Jobs?' in M. Lechner and F. Pfeiffer, (eds.), Econometric Evaluation of Labour Market Policies, New York, Physica-Verlag, 125-152.

Wunsch, C. (2005), Labour Market Policy in Germany: Institutions, Instruments and Reforms since Unification, Department of Economics, University of St Gallen.