# WHAT WORKS FOR WHOM?: THE EFFECTIVENESS OF DIFFERENT EMPLOYMENT PROGRAMMES

**David C Maré**

*Labour Market Policy Group*
*New Zealand Department of Labour*

## Abstract

*This paper estimates the impact of different categories of employment policy interventions on job-seekers' subsequent experience. We generate a range of estimates to help us distinguish programme effects from selection effects. We also examine the robustness of our findings for a range of sub-populations. There is limited variation in programme effects for people with different observed characteristics. Some evidence suggests that programme effectiveness may be countercyclical.*

Keywords: *Jobseeker, programme effects, selection effects*

Each year, the New Zealand government spends around $600m on labour market policies to improve the job prospects of unemployed or disadvantaged job-seekers. This paper investigates one important dimension of how these policies change job-seekers' prospects. The paper analyses administrative records collected by the New Zealand Employment Service (NZES) to learn about the effect that job-seekers' receipt of labour market assistance has on their subsequent contact with NZES. We consider five different types of assistance, and estimate separately the effects on job-seekers with different observable characteristics.

## Methods

Ideally (at least from a researcher's perspective), people would be assigned randomly to interventions, so that any difference in outcomes could be attributed to the effects of the intervention. Such an approach to evaluating the effectiveness of interventions is described as 'experimental'. Because of the random selection, there would be no reason to expect average outcomes to be different for those who received assistance and those who did not. For a range of practical and ethical reasons, there is no history of reliable experimental evaluations in New Zealand, although the approach is widely used in other countries, notably the USA.[1]

The approach taken in this paper is to use what are termed 'quasi-experimental methods' to control for the fact that expected outcomes for those who receive assistance are likely to be different from expected outcomes for those who do not receive assistance, regardless of the impact of the assistance. The logic of targetting assistance to those most in need, which has been a strong theme in New Zealand labour market policies, makes such differences almost inevitable.

Our aim in this paper is to evaluate the impact of a programme in terms of the "effect of the treatment on the treated" - i.e.: compare the outcome for an individual who received assistance with the outcome that *that individual* would have experienced had he or she not received assistance. Obviously, we do not observe the same individual both receiving assistance and *not* receiving assistance, and this is the central identification problem for the evaluation.

Table 1 illustrates the identification problem. The shaded areas show the outcomes that we do *not* observe directly. The first is the outcome that would have occurred in the absence of assistance, for those who receive assistance. The second is the outcome that would occur following assistance, for those who do not receive assistance. The true measure of the effect of treatment on the treated is the difference between the two cells in the first row:

$$\begin{aligned}\alpha &= E[Y_{i1} - Y_{i0} \mid D_i = 1] \\ &= \underline{E[Y_{i1} \mid D_i = 1] - E[Y_{i0} \mid D_i = 1]}\end{aligned} \qquad (1)$$

The individuals who receive the treatment may be quite different from those who do not, and the difference in observed outcomes $\underline{E[Y_{i1} \mid D_i = 1] - E[Y_{i0} \mid D_i = 1]}$ may be due to these differences rather than to the effect of treatment. The term 'selection bias' is used to describe the difference between the true $\alpha$ (see equation 1) and the difference in observed outcomes.

Because we do not observe $E[Y_{i1} - Y_{i0} \mid D_i = 1]$, we need to impose some plausible but essentially untestable assumptions in order to get an estimate of it.[2] In the results presented below, we present three different estimates of programme effects, which reflect alternative assumptions:

**Table 1. Matrix of possible outcomes**

| | Outcome if treatment is *not* received ($Y_{i0}$) | Outcome if treatment *is* received ($Y_{i1}$) |
|---|---|---|
| **Individuals who receive treatment ($D_i$=1)** | $E[Y_{i0} \mid D_i = 1]$ | $E[Y_{i1} \mid D_i = 1]$ |
| **Individuals who *do not* receive treatment ($D_i$=0)** | $E[Y_{i0} \mid D_i = 0]$ | $E[Y_{i1} \mid D_i = 0]$ |

*Notation:*
*i denotes an individual*
*E[A|B] denotes the expected value of variable A given that event B occurs.*

a) naïve estimator: the difference in observed outcomes between those who do receive assistance (the 'treatment group') and those who do not (the 'comparison group').
b) regression estimator: the difference in outcomes between the treatment and comparison groups, controlling for the relationship between observable characteristics and expected outcomes, as estimated by linear regression.
c) propensity estimator: the difference in outcomes between the treatment and comparison groups, controlling for the relationship between observable characteristics and the probability (propensity) of selection, estimated semi-parametrically.
These methods will be explained by example in the text below.

## Data

The data that we use are administrative records collected by NZES between 1 October 1988 and 31 December 1997. The records contain information on (almost) all spells of unemployment and labour market assistance received by anyone who enrolled with NZES during this period. This provides us with a detailed, and often sobering record of what are in many cases extensive interactions with NZES. Gobbi and Rea's paper in this volume provides a summary of the data, and discusses patterns of duration and multiple spells of unemployment.

We group different forms of labour market assistance into five types - referrals to job vacancies; NZES interviews with job-seekers; wage subsidies; training, and work experience courses. The stated objectives of the specific programmes that were offered during the period covered by this study vary greatly. They range from building self-esteem, to placing job-seekers into full-time unsubsidised employment. The overall logic of the interventions was to move job-seekers closer to employment. Job-seeker needs are heterogeneous. There was therefore a range of forms of assistance on offer, and different expectations of what each could achieve.

The current paper considers only one measure of the effectiveness of the interventions - whether the receipt of assistance reduced the time that the job-seeker subsequently spent either unemployed or in receipt of some form of active labour market assistance (described loosely as 'contact' with NZES). The specific outcome measure used for most of the results is the number of weeks of contact with NZES in the 36 months after assistance commences. The major weakness with this outcome measure is that we do not know whether a lack of contact means that the job-seeker is in employment or has left the labour market.

The main results in the paper relate to interventions that commenced in the first quarter of 1993. This period was chosen because it is mid-way through the period that the dataset covers, and therefore allows us to observe job-seekers for a number of years both before and after receiving assistance.

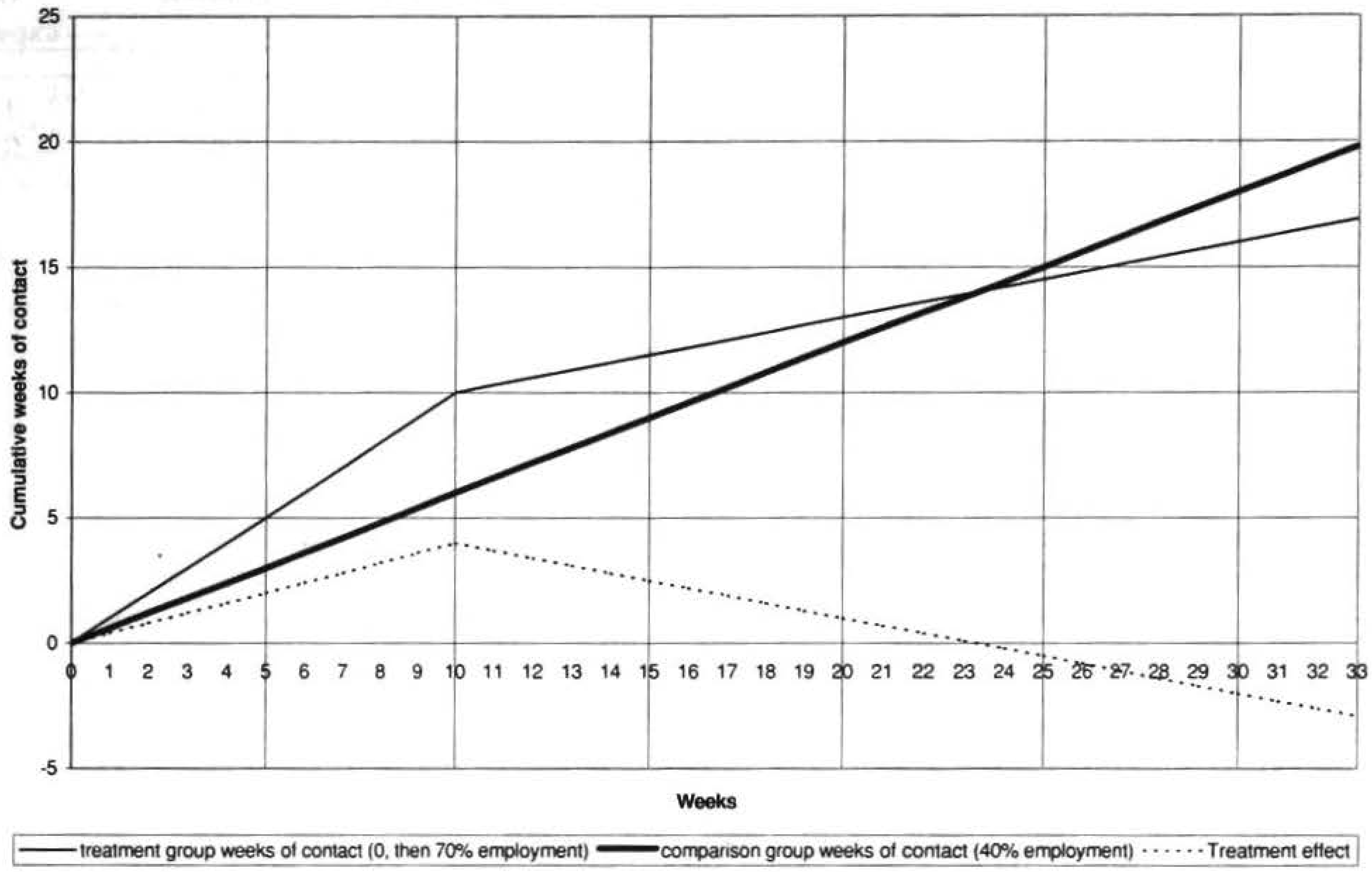
The dataset contains a range of observable characteristics of each job-seeker, measured at the beginning of a spell of unemployment or assistance. These variables are captured as sets of indicator (dummy) variables, and one dummy variable is omitted for each set of characteristics, to avoid perfect collinearity. The included observable characteristics are:

- Age: [15-19; 20-24; 25-29; 30-39; 40-49; 50+]
- Education: [No formal school qualifications; fewer than 3 School Certificate subjects; 3 or more School Certificate subjects; sixth form certificate/ UE; other school qualifications; post-secondary or trade qualifications; degree or professional qualifications]
- Ethnicity: [NZ European/ Päkehä; Sole Mäori; Mixed Mäori; Pacific Islands peoples; Other]
- Location: [Northland; Auckland North; Auckland Central; Auckland South; Waikato; East Coast; Bay of Plenty; Central; Taranaki; Wellington; Nelson; Canterbury; Southern; Not recorded]
- Occupation: NZES occupation codings for preferred occupation
- Barriers: Psychiatric disability; physical disability; intellectual disability; Education/ learning/ literacy/ ESOL barrier; Alcohol and drugs barrier; Multiple disability; no barrier

In addition, we calculate summary measures of the job-seeker's prior contact with NZES. For this, we include the number of weeks of unemployment in the previous 36 months, the duration of the current spell of unemployment, and indicator variables for each form of assistance, indicating whether the job-seeker received that form of assistance in the previous 36 months. These "history" variables raise the explanatory power of the regression equations by about 5 percentage points above what is explained by other personal characteristics. Generally, around 15 percent of the variation in outcomes is accounted for by the regression models.

The 'treatment' group is selected by identifying all job-seekers who commenced a spell in receipt of a particular form of labour market assistance during the first quarter of 1993. There is a different treatment group selected for each type of assistance. The comparison group is selected

**Figure 1. An Effective Intervention**




by identifying all job-seekers who were registered as unemployed mid-way through the first quarter of 1993 and who did not receive any form of assistance during that quarter. The same comparison group is used for the analysis of all forms of treatment. In order to ease the computational burden, we randomly select a sample of comparison observations.

## Does the Type of Assistance Matter?

Figure One illustrates the way that we think about programme effectiveness. The horizontal axis measures calendar time (in weeks). The vertical axis measures the cumulative number of weeks of contact time for each group. The darkest line represents the experience of the comparison group. For simplicity, they are shown as spending 40 percent of their time in employment. After 10 weeks of calendar time, they have therefore had 6 weeks of contact with NZES. In contrast, the lighter solid line shows the experience of a treatment group that receives 10 weeks of assistance. After 10 weeks of calendar time, they have spent all 10 weeks in contact with NZES.

In order for the intervention to be effective, it must improve the employment chances of the treated group. In the graph, we assume that, following the assistance, the treatment group spend 70 percent of their time in employment. They therefore accumulate contact time more slowly than does the comparison group. After about 23 weeks, the lines cross, indicating that The dotted line on the graph shows the difference in contact time between treatment and comparison groups. A negative number indicates that the treatment group has experienced fewer weeks of contact, which is a favourable outcome. As noted above, the main outcome measure used in this paper is the net difference in contact time after 3 years (156 weeks).

Table Two summarises the number of observations in each of the treatment and comparison groups, and shows the average outcome for each group. For instance, the first row shows that there were 59,767 referrals to vacancies during the quarter. On average, those receiving referrals spent 51.8 weeks of the following 36 months either unemployed or in receipt of some form of assistance. In contrast, the comparison group spent 76 weeks of the following 36 months in contact with NZES. The difference of 24.2 weeks is a combination of the programme effect, and the selection effect (those who were referred to vacancies had better prospects than those who were not referred, even before the referral). It would be naïve to attribute all of the 24.2 week difference in outcomes to the referral.

Similarly, it would be naïve to infer from the other results in the right-most column of Table Two that training increases subsequent contact time by 19 weeks, work experience by 16.1 weeks, and interviews by 6 weeks. All of these forms of assistance were targeted to disadvantaged job-seekers, whom we would expect to have more weeks of contact time. The raw difference between outcomes for the subsidy treatment and control groups may appear a little surprising, given that subsidies too are targeted to disadvantaged job-seekers.

Note that the contact time measured in our outcome vari-

**Table 2. Summary of interventions - 1993Q1 Intervention Cohort**

| | Vacancies | Interviews | Subsidies | Training | Work experience |
|---|---|---|---|---|---|
| **Sample size** | | | | | |
| Treatment group (#) | 59767 | 32285 | 7988 | 12576 | 1786 |
| Comparison group (#) | 20644 | 20644 | 20644 | 20644 | 20644 |
| **Outcome = weeks of contact in subsequent 3 years** | | | | | |
| Treatment group | 51.8 | 82 | 65.2 | 95 | 92.1 |
| Comparison group | 76 | 76 | 76 | 76 | 76 |
| Raw diff | -24.2 | 6 | -10.8 | 19 | 16.1 |

able includes the time spent in receipt of the treatment. For instance, the 10.8 fewer weeks of contact that the subsidy treatment group received occurred despite the fact that the average duration on subsidy was 21.1 weeks. For work experience assistance, the raw difference of 16.1 more weeks of contact was around the same as the average length of assistance (16.5 weeks). The raw difference for training interventions was 19.0, compared with an average course length of 10.4 weeks.

There are reasons other than selection bias to be wary of interpreting the raw outcome differences in Table Two. For instance, the treatment and control groups are selected in different ways. As noted above, the treatment group is selected on the basis of job-seekers starting a spell of assistance. This is termed 'flow-sampling'. The comparison group is selected on the basis of being unemployed at a particular point in time. This is commonly referred to as 'stock-sampling'. It is a standard result in duration analysis that average duration in a stock-sampled sample is higher than that in a flow-sampled sample. It is hard to tell a priori which way the sampling methods bias the results in our particular application. Most of the treatments (with the exception of vacancy referrals) are targeted at longer duration job-seekers. The stock sampling of the comparison group gives more weight to job-seekers with longer duration.

Any sensible comparison of outcomes for treatment and comparison groups must take account of duration composition. Similarly, training assistance is understandably targeted to job-seekers with relatively low levels of qualifications, so we need to allow for the different qualification profiles of treatment and comparison groups, at least to the extent that they are related to outcomes or receipt of treatment. In practice, we wish to control for a broad range of observable attributes that may differ between treatment and comparison groups.

The first method used to control for these differences is linear regression. The outcome variable (weeks of contact during a three-year period) is modelled as a linear function of observable characteristics. The regression is estimated using a sample of observations on the comparison group, and on a selected treatment group. An indicator variable is added to the regression to capture whether the observation is for a comparison or treatment spell. The estimated coefficient on this indicator ($\gamma$ in the following equation) represents the mean difference in outcomes between the treatment and comparison groups, allowing for differences in characteristics ($X_i$) and the average relationship that exists between the observable characteristics and the outcome ($Y_i$). In the notation introduced in the methods section of this paper, our estimate of the (unobserved) $E[Y_i^0 \mid D_i=1]$ is the predicted outcome for someone who has the same characteristics as an assisted job-seeker but who did not receive assistance. The regression implicitly puts more weight on characteristics that are strongly related to outcomes.

$$Y_i = \alpha + \gamma D_i + X_i\beta + \varepsilon_i \qquad (2)$$

where: i indexes observations,

$Y_i$ is the outcome variable (weeks of contact in 3 years)

$X_i$ is a vector of observable characteristics on observation i (as listed in the data section above.

$D_i$ is an indicator (dummy) variable that equals one for a treatment group observation and zero for a comparison group observation.

Table 3 presents estimates of the parameter $\gamma$ for various regressions of this form. The first row reports estimates from regressions that do not contain any observable characteristics. The estimate is the same as the mean difference reported in the right-most column of Table Two. The second row reports estimates of the parameter $\gamma$ from regressions containing a full set of covariates (X s). Each of the numbers in Table Three is from a separate regression. Numbers in the same row use the same treatment group (the comparison group is always the same). Numbers in the same column use the same regression specification.

By comparing the estimates from the first row (naïve estimates) with those from the second row (regression estimates), we get an indication of the size and direction of selection bias. The most pronounced change is for the estimate of the effect of referrals to vacancies. The estimated impact drops from -24.2 weeks, to -7.9 weeks. This implies that a good deal of the initial favourable estimated impact is due to the fact that the sort of job-seekers who were referred to vacancies had characteristics associated with good outcomes. This is confirmed by examining average characteristics of the vacancy treatment and com-

**Table 3. Estimated impacts of interventions (Total and for selected subgroups)**
**Dependent Variable: Outcomes in the 36 months after the reference date**

| | Vacancies | Interviews | Subsidies | Training | Work Experience |
|---|---|---|---|---|---|
| **Total** | | | | | |
| **Naïve estimate (raw difference)** | -24.2 | 6.0 | -10.9 | 19.0 | 16.1 |
| | (0.4) | (0.4) | (0.6) | (0.6) | 1.3) |
| **Regression Estimate** | -7.9 | 8.6 | -6.2 | 20.1 | 15.1 |
| | (0.4) | (0.5) | (0.7) | (0.7) | (1.2) |
| **Propensity Estimate** | -11.7 | 6.3 | -13.9 | 18.8 | 14.6 |
| | (0.8) | (1.1) | (3.3) | (2.0) | (1.3) |
| **Sub-groups** | **(propensity** | **estimates)** | | | |
| ***Gender*** | | | | | |
| **Male** | -11.5 | 3.6 | -13.0 | 11.6 | 12.9 |
| | (1.2) | (1.7) | (3.4) | (1.7) | (1.6) |
| **Female** | -11.4 | 7.6 | -7.3 | 24.2 | 19.0 |
| | (1.5) | (1.8) | (2.1) | (2.5) | (2.1) |
| ***Ethnicity*** | | | | | |
| **Mäori** | -11.6 | 8.1 | -6.5 | 24.7 | 12.0 |
| | (1.5) | (1.6) | (2.0) | (2.1) | (2.1) |
| **Pacific Peoples** | -9.2 | 12.8 | -7.2 | 15.5 | 17.9 |
| | (2.7) | (2.8) | (4.3) | (4.2) | (6.0) |
| ***Qualifications*** | | | | | |
| **Low qualifications (less than UE)** | -13.2 | 6.0 | -11.9 | 17.6 | 10.8 |
| | (1.3) | (1.5) | (2.1) | (3.1) | (1.7) |
| **High qualifications (UE or above** | -11.9 | 5.3 | -11.2 | 16.0 | 15.9 |
| | (2.4) | (5.1) | (3.5) | (2.4) | (3.0) |
| *Unemployment Duration* | | | | | |
| 0-13 weeks | -12.0 | 0.8 | -26.9 | 4.30 | 14.8 |
| | (4.3) | (10.8) | (27.1) | (10.7) | (3.6) |
| 13-26 weeks | -4.3 | 5.1 | 1.0 | 22.6 | 12.6 |
| | (1.9) | (6.9) | (2.9) | (3.7) | (3.3) |
| 26-52 weeks | -6.5 | 11.4 | 0.8 | 16.3 | 22.0 |
| | (4.3) | (2.1) | (3.1) | (10.0) | (4.3) |
| 52-104 weeks | -2.7 | 15.6 | -8.6 | 3.4 | 15.9 |
| | (2.4) | (5.6) | (2.6) | (3.1) | (3.6) |
| 104+ weeks | -7.0 | -1.3 | -17.2 | 0.4 | 11.0 |
| | (2.0) | (2.3) | (2.8) | (3.0) | (4.1) |
| | weeks of "any contact" | | | weeks of registered unemployment | |
| *Location* | | | | | |
| Northland | -16.2 | 5.6 | -10.6 | 13.3 | 8.9 |
| | (6.4) | (4.7) | (4.3) | (7.0) | (6.2) |
| North Auckland | -16.2 | 3.4 | -6.7 | 20.9 | 21.2 |
| | (3.2) | (4.0) | (4.7) | (5.2) | (7.6) |
| Auckland Central | -10.9 | 0.1 | -8.7 | 12.2 | 18.3 |
| | (4.5) | (3.9) | (3.4) | (3.9) | (7.8) |
| South Auckland | -11.0 | 7.8 | 1.4 | 17.7 | 18.0 |
| | (3.1) | (3.5) | (3.3) | (0.7) | (5.2) |
| Waikato | -8.9 | 8.8 | 1.6 | 20.2 | 17.0 |
| | (3.3) | (3.5) | (4.0) | (3.7) | (5.6) |
| Central | -12.0 | 9.9 | 1.2 | 25.2 | 21.1 |
| | (4.4) | (3.8) | (4.3) | (5.1) | (7.2) |
| Bay of Plenty | -1.8 | 11.3 | -9.0 | 25.4 | 7.2 |
| | (2.4) | (3.6) | (3.4) | (4.3) | (5.3) |
| East Coast | -8.8 | -1.2 | -13.4 | 12.5 | 32.6 |
| | (3.7) | (4.2) | (4.6) | (5.6) | (7.0) |
| Taranaki | -23.0 | 3.6 | -15.9 | 8.3 | 17.9 |
| | (3.3) | (5.4) | (4.8) | (4.0) | (8.6) |
| Wellington | -9.9 | 3.8 | -20.0 | 20.1 | 9.3 |
| | (3.2) | (3.5) | (13.0) | (4.5) | (9.4) |
| Nelson | -7.5 | 3.1 | -19.4 | 23.8 | 29.9 |
| | (3.0) | (5.1) | (5.9) | (5.1) | (12.7 |
| Canterbury | -14.3 | 11.8 | -6.5 | 16.8 | 13.8 |
| | (3.3) | (4.7) | (3.0) | (5.6) | (4.8) |
| Southern | -6.4 | 11.4 | -16.7 | 12.9 | 11.3 |
| | (2.9) | (3.6) | (7.7) | (4.0) | (4.2) |

Notes: Each cell in this table is from a separate OLS regression. Cells in the same column relate to the same type of intervention. Cells in the same row share a common sample definition – comprising the treatment and comparison groups for the sub-population listed in the first column.

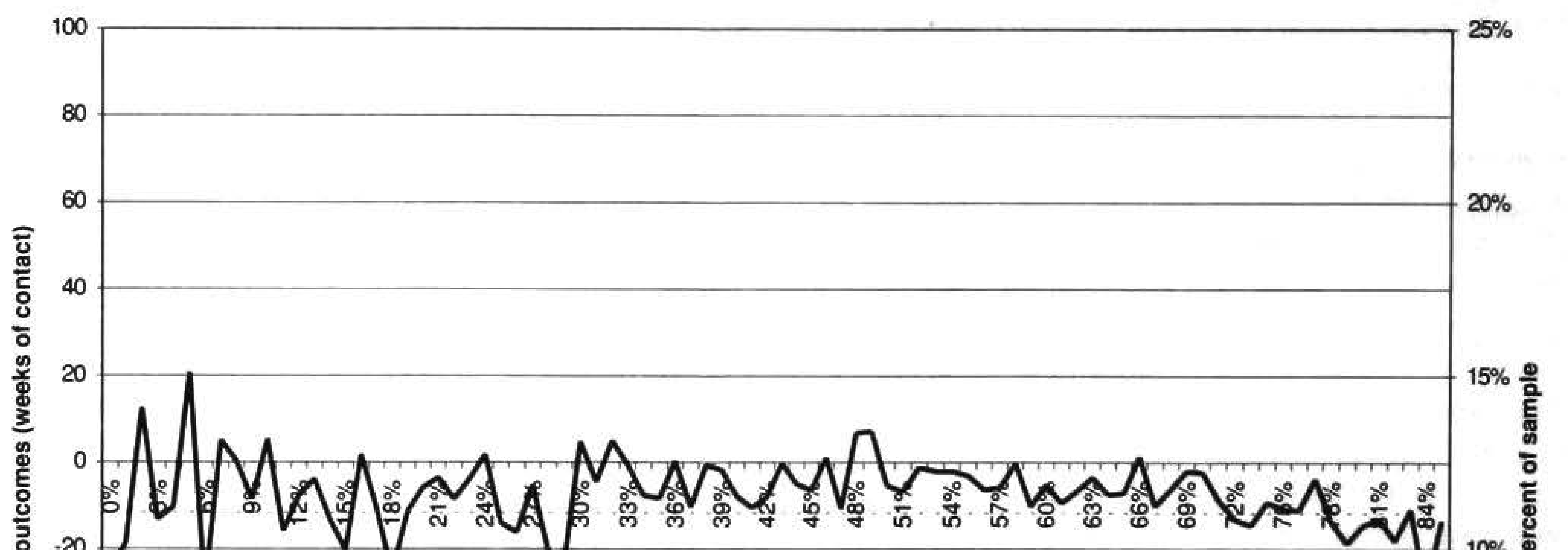
parison groups. The treatment group, for instance, has much lower average unemployment duration. This is likely to be due both to actual differences in the sort of job-seekers who are referred to vacancies, and also to the way that the sampling was done, as noted above. We rely on the regression method to remove both of these effects.

The removal of selection effects makes less of a difference for estimates of the impact of the other forms of assistance. Using regression to control for differences in characteristics generally reduces the estimated effect of interventions in reducing subsequent contact time. Subsidies are estimated to reduce subsequent contact by 6.2 weeks over the 3 years following assistance, compared with the raw estimate of 10.9 weeks. Interviews and training are followed by a larger number of contact weeks (8.6 and 20.1 respectively, compared with the naïve estimates of 6.0 and 19.0). Episodes of work experience are estimated to result in 15.1 more weeks of contact, which is insignificantly different from the naïve estimate of 16.1 weeks.

The third row of Table 3 contains estimates based on propensity matching. Instead of modelling outcomes, as was done for the regression modelling just described, propensity matching relies on modelling the probability (or 'propensity') that a job-seeker receives assistance. The first step in deriving a propensity-matched estimate is to calculate the propensity score. Using the same definition of treatment and comparison groups as used for the regression matching, we estimate the probability that each observation is in the treatment group. To do this, we use a logistic regression, with the full set of covariates as used for regression matching. We can then compare outcomes for those in the treatment group with outcomes for those in the comparison group who were most likely to receive assistance (but did not).[3]

Figure 2 helps to illustrate the method. Having obtained a predicted probability of treatment for each observation, we can rank job-seekers according to this prediction. The horizontal axis is the predicted probability, or propensity score. The two lower lines (using the right axis) show the densities for the treatment and comparison samples - the proportion of each sample with a particular propensity score.[4] The dotted line shows that the treatment group is concentrated at the right of the graph, with high probabilities of being treated. The comparison group is more evenly spread across predicted probabilities. The treatment group accounts for 74.3 percent of the full sample, so the regression-matching coefficients are strongly weighted towards the treatment group. The weighting for the propensity matched estimate is based on the density of the treatment group. The most weight is therefore given to outcome differences between those in the treatment and comparison groups with treatment propensities of around ten percent and sixty percent.

**Figure 2. Propensity Score estimation for Vacancies**



The formula for the propensity matched estimate is thus a weighted average of the form:

$$\alpha_p = \sum_{s^i} \Delta^i \lambda_T^i \qquad (1)$$

where $s^i$ = the i'th distinct propensity score;
$\Delta_i$ = the difference in outcomes between treatment and comparison groups with propensity score $s^i$;
$\lambda_T^i$ = the proportion of the Treatment group that has propensity score $s^i$

The line in the centre of the graph, labelled "difference in outcome" is the difference between outcomes for treatment and comparison groups with the same propensity score. The lightly dotted line at around -12 weeks is the weighted average of these differences, using the density of the treatment group as weights.

The difference in outcomes does not differ greatly across different propensity scores. Referral to vacancies appears to have the same favourable effect on subsequent outcomes regardless of how likely a job-seeker is to be referred to vacancies. There appears to be little evidence that those who are referred to vacancies are those who are likely to benefit most.

The third row of Table 3 contains propensity estimates of the effects of the different types of interventions. The patterns are broadly similar to the regression estimates. Our preferred method of controlling for differences in characteristics produces point estimates that are uniformly more favourable than the regression estimates. The propensity estimates show greater reductions, or smaller increases in subsequent contact time.

It is clear from both the regression and propensity estimates that the different forms of intervention have quite different effects on subsequent contact time. Referrals to vacancies and the receipt of wage subsidies are the only forms of intervention associated with lower levels of subsequent contact. As noted earlier, not all interventions are designed to or expected to reduce the amount of subsequent contact or assistance that job-seekers receive. Some forms of assistance serve largely to prepare recipients for other forms of assistance. A positive coefficient does not necessarily mean that the assistance is ineffective - it does mean that it does not have the effect of reducing subsequent contact time over a limited (three year) period.

## Do the Jobseeker's Characteristics Matter?

In one sense, job-seeker characteristics do matter. The likelihood of receiving any particular form of assistance, or the expected number of weeks of contact, are definitely related to job-seeker characteristics. It is these relationships that we use in the regression and propensity methods to allow us to estimate treatment effects.[5]

However, when we look at the effect of different forms of assistance, observable characteristics do not appear to matter as much, at least not in any systematic way. While there are some differences in estimated effects for different sub-groups, the general pattern of effects across different forms of assistance holds for most of the groups considered. What differences there are between sub-groups are often not statistically significant.

The remainder of Table Three presents propensity estimates based on separate samples of job-seekers with particular characteristics - by ethnicity, qualifications, unemployment duration, and location.

The fourth and fifth rows of the table show results for males and females separately. A male treatment group is compared with a male comparison groups, and a female treatment group is compared with a female comparison group. This method allows the estimated relationship between observable characteristics and the probability of selection to be different for males and females, as well as allowing the effects of treatment to differ. The effect of assistance in reducing subsequent contact appears to be stronger for males than for females. The male results show consistently larger reductions or smaller increases than the results for females, although only for training is the difference statistically significant.

There are no strong patterns of differing effects for different ethnic or qualifications groups. Results are generally most favourable for long duration (104 weeks or more) job-seekers, although there is not a clear universal duration gradient.[6] The results across NZES regions vary widely. The ranges of estimates for the different interventions are: vacancies (-1.8 to -23.0), interviews (-1.2 to 11.8), subsidies (-20.0 to 1.6), training (8.3 to 25.4), and work experience (7.2 to 32.6).

## Does The Labour Market Matter?

It is plausible to expect that the effect of interventions in reducing subsequent time will vary depending on the state of the labour market. It is less clear whether to expect interventions to be most effective at reducing subsequent contact time in tight or in loose labour markets. Certainly we would expect outcomes to be better in a tight labour market but this applies to both treatment and comparison groups, and our measure of effectiveness is based on the difference between the two. We have limited information on the links between labour market conditions and programme effects. We rely on variation over time, across locations, and between different types of intervention to provide indirect and circumstantial evidence on these links.

Table 4 contains estimates of the effect of assistance on subsequent contact time, for cohorts selected in the first quarter of each year.[7] These data are graphed in Figure 3, along with the employment growth rate in the year to March for each year. Although estimates are not yet available for all years, the pattern at this stage strongly suggests that

**Table 4. Estimated impacts of interventions (For different periods)**
**Dependent Variable: Outcomes in the 24 months after the reference date**

| | Vacancies | Interviews | Subsidies | Training | Work Experience |
|---|---|---|---|---|---|
| **Year of sample (Regression estimates)** | | | | | |
| 1990 | -9.6 | -2.6 | -13.0 | 7.2 | 3.8 |
| | (0.5) | (0.5) | (1.1) | (1.1) | (2.1) |
| 1991 | -11.0 | -0.2 | -16.5 | 5.3 | 6.9 |
| | (0.4) | (0.4) | (0.8) | (0.8) | (0.0) |
| 1992 | -7.5 | -0.1 | -10.0 | 8.4 | 6.8 |
| | (0.3) | (0.3) | (0.5) | (0.5) | (1.0) |
| 1993 | -5.7 | 7.0 | -3.8 | 12.6 | 12.4 |
| | (0.3) | (0.4) | (0.5) | (0.5) | (0.8) |
| 1994 | 23.9 | 29.3 | 22.8 | 33.7 | 28.9 |
| | (0.3) | (0.3) | (0.6) | (0.4) | (1.0) |

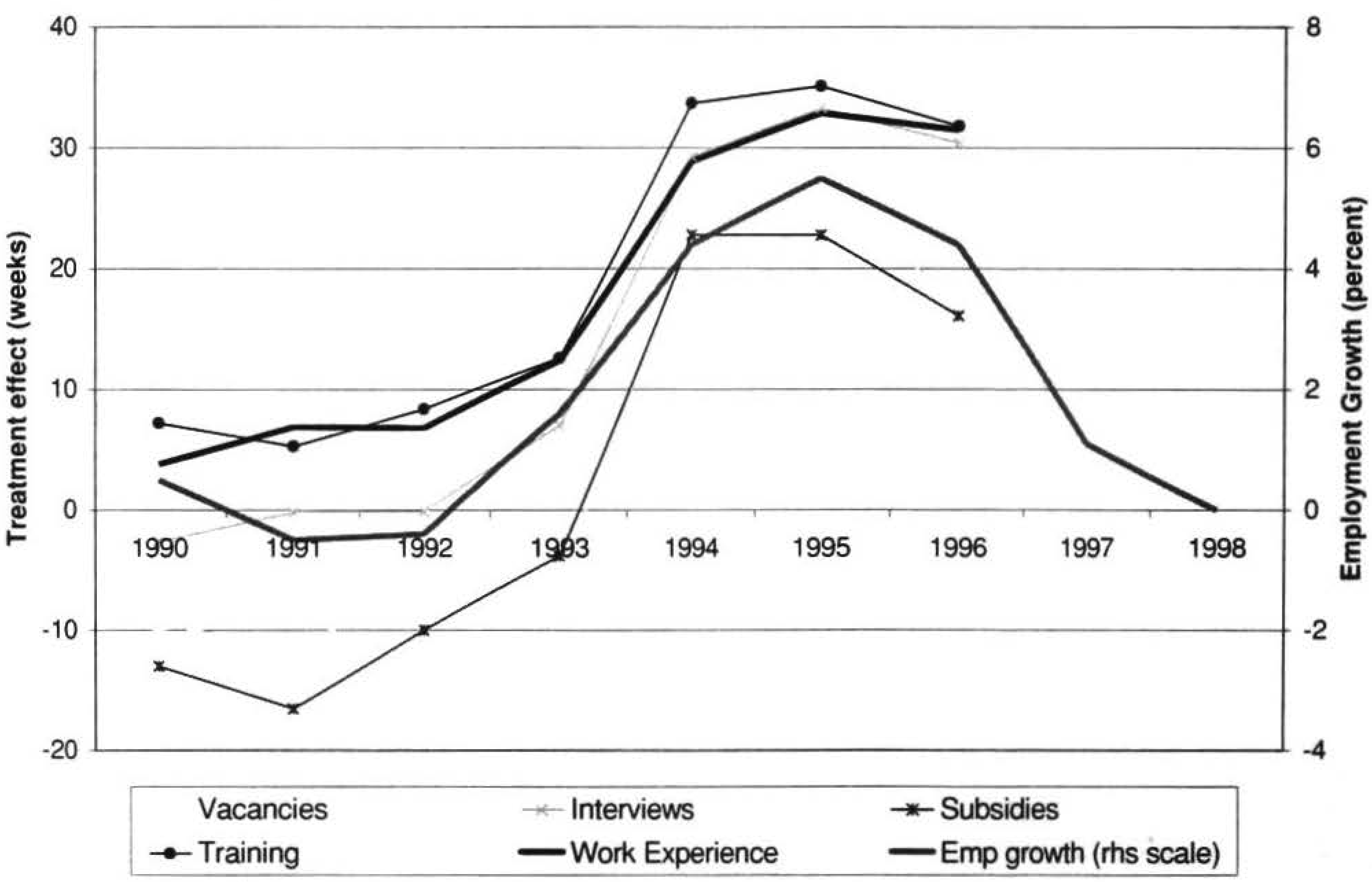
Notes: Each cell in this table is from a separate OLS regression. Cells in the same column relate to the same type of intervention. Cells in the same row share a common sample definition – comprising the treatment and comparison groups for the sub-population listed in the first column.

rapid employment growth is associated with *longer* subsequent contact for the treatment group, relative to the comparison group.

Between 1993 and 1994, when the growth in employment and the deterioration of the effectiveness of interventions was greatest, there was both a drop in the average number and a rise in the average weeks of subsequent contact for each of the treatment groups. It is possible that the deterioration in outcomes for the treatment groups is a result of selection effects. Faced with strong employment growth, it may be that those who remained unemployed in 1994 and 1995, and received assistance, were the hardest to place job-seekers. The regression method controls for such selection to the extent that it is evident in observed characteristics, but we have no way of convincingly controlling for unobserved heterogeneity.[8]

We also inspected the results by location, which were presented in Table 3 and discussed above, for evidence on a

**Figure 3. Employment Growth and the Effects of Interventions Over Time**



link between intervention effects and the local labour market. Unfortunately, we do not have employment data for the regions used for this analysis. Instead, we compared the estimated effects with the relative unemployment level in each regions - i.e. number of people unemployed in the region, compared with the average level of unemployment for that region. There was no discernible pattern of effects related to the level or the change in unemployment.

The other piece of circumstantial evidence on the importance of the labour market is the fact that referrals to vacancies and wage subsidies are the two forms of assistance that lead to the greatest reductions in subsequent contact time. These are also the forms of assistance that are dependent on employers being willing to take on an unemployed worker.

## Other Outcomes

This final section provides some evidence on the extent of 'staircasing' - interventions may move job-seekers closer to work and independence by preparing them to receive other forms of assistance for which they are not initially ready. This may be an appropriate strategy for some job-seekers. In terms of our chosen outcome measure, staircasing would lead to more rather fewer weeks of subsequent contact.

As a crude indicator of the degree of staircasing, we report in Table 5 estimates of the effect of past receipt of different forms of assistance on the likelihood of receiving various forms of assistance. The reported estimates are the proportional changes in a job-seeker's likelihood of receiving assistance given that the job-seeker received a specific form of assistance in the previous three years.[9]

There is clear evidence of job-seekers receiving multiple spells of the same sort of assistance. For instance, having received a subsidy in the past increases your chances of receiving one again by 53 percent. The exception is interviews, which generally occurred after a set number of weeks of unemployment, and are therefore less likely to be affected by prior assistance. Having had an interview increases the chance of receiving further assistance, which is not surprising given that interviews generally entailed taking stock of a job-seeker's needs and directing them to appropriate assistance. It appears that receiving a subsidy, training, or work experience increases the chance of receiving any of those three forms of assistance. The exception is that subsidies are less likely to be followed by further training.

The results in this section help to shed light on why some forms of assistance appear to lead to longer periods of subsequent contact. They are not intended to be a sound analysis of staircasing. Such an analysis is beyond the scope of the current paper.

## Conclusions

Different forms of active labour market policies have quite different effects on the amount of contact that job-seekers have with the public employment agency. Referrals to vacancies and wage subsidies appear to be most likely to reduce subsequent contact time. Other forms of assistance may serve to help job-seekers into further assistance rather than directly into jobs.

The effects of interventions on subsequent contact time does not appear to vary much according to personal attributes, although there is variation across job-seekers with different unemployment duration and across different locations. The limited data analysed here suggest that interventions may become less effective when employment growth is strong, probably because of the improved prospects for those who are not receiving assistance.

## Future Research

The research reported in this paper is the result of a significant allocation of resources - not only in modelling and estimation, but more fundamentally in assembling a research database from administrative data. Nevertheless, the work done to date has barely scratched the surface of what can be learnt from the data. In this section, I briefly outline some possible directions for future work. The list is far from exhaustive!

- Updating: For reasons of consistency of data, the current study uses data only up until 1997. More recent DWI data could be used to analyse more recent pat-

**Table 5. Staircasing - Effect of Previous Assistance on Probability of Receiving Assistance**

| | P[Referral to vacancy] | P[Interview] | P[Subsidy] | P[Training] | P[Work Experience] |
|---|---|---|---|---|---|
| *Assistance received in the previous three years* | | | | | |
| Referral to vacancy | 38% | 9% | 37% | 21% | 44% |
| Interview | 10% | -2% | 46% | 71% | 68% |
| Subsidy | 3% | 0% | 53% | -11% | 25% |
| Training | 0% | 8% | 24% | 53% | 38% |
| Work Experience | 9% | 2% | 48% | 22% | 363% |

terns of effects.

- Labour market links: Figure 3. suggests a strong link between programme effectiveness and employment growth, but we do not at this stage understand the reasons for this relationship. Further work is clearly needed to understand this issue, and why regional patterns show no obvious link to labour market conditions.

- Intensity of treatment: In the current paper, we do not use the information that is available on the length of treatment. For instance, subsidies, training, and work experience courses vary in length. Future work could examine whether longer periods of assistance have stronger effects.

- Other outcome measures: One of the main weaknesses of the current dataset is the lack of confirmed labour market outcomes - we observe only whether someone is still in contact with NZES. Future could supplement the administrative data by collecting outcome information from people who are no longer in contact with the agency. The impact on wages, employment, hours worked, etc. would be particularly valuable. Other administrative data, such as benefit records, could be used to supplement the data used in this project, and provide additional information about those who lost contact with NZES.

- Individual programmes: This study has grouped interventions into generic types. It is possible to identify specific schemes and programmes. Analysing specific forms of assistance would allow more careful incorporation of rules and practices about who received assistance. It could also allow a closer examination of the features of schemes that appear most effective.

- Staircasing: The current paper has not carefully analysed patterns of staircasing, or whether particular sequences of assistance are more effective than others. Given how common it is for assisted job-seekers to receive more than one form of assistance, such an analysis would be a valuable input into operational policy.

## Notes

1. See, for instance various studies by the Manpower Demonstration Research Corporation, documented at www.mdrc.org

2. For further discussion, See Rubin (1974)

3. For a fuller discussion of this method, see Dehejia & Wahba (1998), or the discussion in Angrist & Krueger (1999)

4. The predicted probability is rounded to the nearest percentage point, so there is a maximum of 100 possible distinct values. The approach thus uses non-overlapping rectangular kernels of width 0.01, centred on multiples of 0.01. There are many alternative (and arguably superior) ways in the literature of smoothing estimates across the propensity domain.

5. Around 15 percent of the variation of outcomes can be accounted for by differences in charcateristics.

6. There are relatively few assisted jobseekers in the low-duration sub-group, presumably due to duration being the main targeting criterion. The small size of the treatment group is reflected in the large standard errors on the coefficients.

7. Note that the estimates in Table 4. are not directly comparable with the estimates in Table 3., for two main reasons. First, the outcome definition is weeks of contact in the *two* years following the start of assistance, and controlling for only one year of prior intervention experience. The results in Table 3. use a period of three years for each of these measures. This is done to avoid having to drop as many years at either end of the sample period. Second, the estimates are from regression rather than propensity estimation. This is due to the fact that the propensity estimation is very time-consuming and propensity estimates has not yet been completed.

8. Methods do exist to adjust for such heterogeneity, but generally rely on observing an attribute that is related to participation but unrelated to outcomes. We were unable to find such a variable.

9. A separate linear probability model is used to estimate the probability of being in each treatment group. The model contains a full set of covariates, including dummy variables capturing whether the job-seeker has received each of the forms of assistance in the previous three years. The reported estimate is one plus the ratio of the coefficient on each of these dummies to the mean probability of being in the treatment groups. This gives an estimate of the proportional change in probability.

## References

**Angrist, J. & Krueger, A.** (1999) Empirical Strategies in Labour Economics. Ashenfelter, O. & Card, D. (eds) *Handbook of Labor Economics: Volume 3A*, Elsevier, Amsterdam. Ch 23: 1277-1366.

**Dehejia, R & Wahba, S.** (1998) Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programs' *NBER Working Paper 6586*.

**Rubin D.** (1974), Estimating Causal effects of Treatments in Randomized and non-Randomized Studies. *Journal of Educational Psychology* 66 pp. 688-701.

## Author

David C Maré
Advisor
Labour Market Policy Group
New Zealand Department of Labour
P O Box 3705
Wellington
Dave.Mare@lmpg.dol.govt.nz