



**MINISTRY OF SOCIAL
DEVELOPMENT**

TE MANATŪ WHAKAHIATO ORA

Effectiveness of the Limited Service Volunteer programme in 2014/15

Evaluation report

August 2018

Authors

Marc de Boer, Principal Analyst, Ministry of Social Development

Bryan Ku, Analyst, Ministry of Social Development

Statistics New Zealand IDI disclaimer

The results in this report are not official statistics, they have been created for research purposes from the Integrated Data Infrastructure (IDI) managed by Statistics New Zealand.

The opinions, findings, recommendations and conclusions expressed in this paper are those of the author(s) not Statistics New Zealand.

Access to the anonymised data used in this study was provided by Statistics New Zealand in accordance with security and confidentiality provisions of the Statistics Act 1975. Only people authorised by the Statistics Act 1975 are allowed to see data about a particular person, household, business or organisation and the results in this paper have been confidentialised to protect these groups from identification.

Careful consideration has been given to the privacy, security and confidentiality issues associated with using administrative and survey data in the IDI. Further details can be found in the Privacy impact assessment for the Integrated Data Infrastructure available from www.stats.govt.nz.

Inland Revenue IDI disclaimer

The results in this report are based in part on tax data supplied by Inland Revenue to Statistics NZ under the Tax Administration Act 1994. This tax data must be used only for statistical purposes, and no individual information may be published or disclosed in any other form, or provided to Inland Revenue for administrative or regulatory purposes.

Any person who has had access to the unit record data has certified that they have been shown, have read, and have understood section 81 of the Tax Administration Act 1994, which relates to secrecy. Any discussion of data limitations or weaknesses is in the context of using the IDI for statistical purposes, and is not related to the data's ability to support Inland Revenue's core operational requirements.

New Zealand Defence Force IDI disclaimer

The New Zealand Defence Force has consented to the release of the IDI results by Statistics New Zealand as part of this report.

Published

Ministry of Social Development

PO Box 1556

Wellington

www.msd.govt.nz

Published in August 2018

ISBN

Online 978-1-98-854136-5

Contents

Summary.....	4
The Limited Service Volunteer programme aims to motivate young job seekers.....	4
The LSV programme has mixed effects.....	4
Subgroups were impacted differently.....	4
Estimating the effectiveness of the programme.....	5
Findings.....	7
Primary and secondary outcomes.....	7
Outcomes at the subgroup level.....	9
Actionable insights.....	12
LSV participant cost.....	13
Future work.....	14
Appendix 1: Summary of technical notes.....	15
Impact method.....	15
Caveats.....	15
Impact results and participant outcome values.....	16
Appendix 2: Technical notes.....	19
LSV participation.....	19
Outcome measures.....	19
Effectiveness rating.....	22

Summary

In this report, we examine the effectiveness of the Limited Service Volunteer (LSV) programme. We measured the impact of the programme on approximately 13,000 participants' outcomes – focusing on income, employment and dependence on welfare assistance.

The Limited Service Volunteer programme aims to motivate young job seekers

The LSV programme is a residential motivational training scheme run by the New Zealand Defence Force. The LSV programme targets people on income support aged between 17 and 25 years of age. Participants of the programme stay at an army barracks for six weeks and undertake basic military training (excluding weapons training). The aim of the programme is to increase the number of young job seekers entering employment or training by improving their self-discipline, self-confidence, motivation and initiative. We expected the programme to increase participants' job searching activities, resulting in participants moving into employment or training within 16 weeks of programme completion.

The LSV programme has mixed effects

Overall, the effectiveness of the LSV programme is mixed because the analysis showed both significant positive and negative effects on different outcomes. See Table 1.

We estimated the effectiveness of the programme by comparing the outcomes of the participants to a matched group of people that did not take part.

We examined a range of primary outcomes including employment, welfare, and income. Secondary outcomes included education and justice.

Relative to the comparison group, we found the LSV programme:

- increased both the time that participants spent in employment and their net income from employment
- slightly reduced the time that participants spent in prison
- increased the time participants spent in training or education
- led to a greater dependence on welfare assistance.

Subgroups were impacted differently

We compared outcomes between participant subgroups. We considered the effectiveness of the LSV programme for participants depending on their gender, age, ethnicity and their likelihood of being on a benefit long-term. The Ministry of Social Development (MSD) groups people into three categories based on their likelihood of receiving a benefit in the next two years (low, medium and high). The LSV programme does not target people at high risk of being on a benefit long term.

We found the LSV programme had:

- stronger positive impacts for participants that were more likely to receive a benefit in the long term for income and employment outcomes, compared to those with a lower likelihood
- stronger positive effects on net income from employment for female participants than male participants
- a positive impact on participation in education of Māori participants
- a negative impact on the percentage of younger participants obtaining National Qualification Framework (NQF) levels
- a positive impact on the time spent in any corrections services for older participants
- a positive impact on justice domain outcomes for male participants but not for female participants
- no consistent or significant difference in the effectiveness of the programme by the year participants started.

Table 1 - Summary of observed impacts

Average impact on participants relative to the comparison group	Direction of impact
Participants spend three more weeks being employed over 2.5 years	Positive
Participants earn \$1,300 more over 2.5 years	Positive
Participants spend 0.7 fewer weeks in prison over 2.5 years	Positive
Participants spend 1.4 more weeks in training or education over 2.5 years	Positive
Fewer participants (one percentage point less than comparison group) who achieved a NQF level 4 qualification over 2 years	Negative
Participants spend 4.6 more weeks being dependent on welfare assistance over 3 years	Negative

The impacts listed above are statistically significant – we are more confident that an impact was truly observed and not due to chance. The average impact is measured with reference to a comparison group constructed from propensity score matching. We used an alpha level of 0.05 to determine statistical significance.

Estimating the effectiveness of the programme

The comparison group resembled the participants on a range of observed characteristics (for example, similar duration on a benefit, benefit type, age, gender, and education level). We constructed the comparison group using propensity score matching, and measured the outcomes of participants as well as the comparison group using linked administrative data in the Statistics New Zealand Integrated Data Infrastructure (IDI).

We are confident about the estimated impacts of the LSV programme on most reported outcomes. However, we were not able to perfectly match the participants to the

comparison group so some of the results reported are likely to overstate the effectiveness of the programme. In particular, we are cautious about the LSV programme's effects on justice outcomes because the comparison group had a higher rate of corrections contact before the start of the programme.

We were not able to observe the impact of the LSV programme on intermediate outcomes such as self-discipline and self-confidence, as these psychological characteristics were not included in this study.

Findings

In this section, we provide more detail on the impact of the LSV programme on participants' outcomes. We examine the primary and secondary outcomes before evaluating the effect of the programme at the subgroup level.

We followed the outcomes of a sample of approximately 13,000 MSD clients who participated in the programme between 2000 to 2014. The follow up period is between 2 to 12.5 years (the outcome window) depending on how recently participants started the programme.

Primary and secondary outcomes

We examined the effectiveness of the LSV programme against three primary outcomes.

- Employment: time in employment either on or off income support assistance.
- Income: income from all sources – employment and income support assistance.
- Independent of welfare assistance: time spent off main benefits or receiving employment assistance – Jobseeker Support, Sole Parent Support, or wage subsidies.

We also examined two secondary outcomes.

- Education: the educational qualifications measured in this report are defined according to the National Qualifications Framework (NQF).
- Justice: any spell under corrections supervision including periods of custodial and non-custodial supervision – prison, community service, home detention, remand, parole and periodic detention.

Therefore, the LSV programme is rated ¹as having mixed effectiveness.

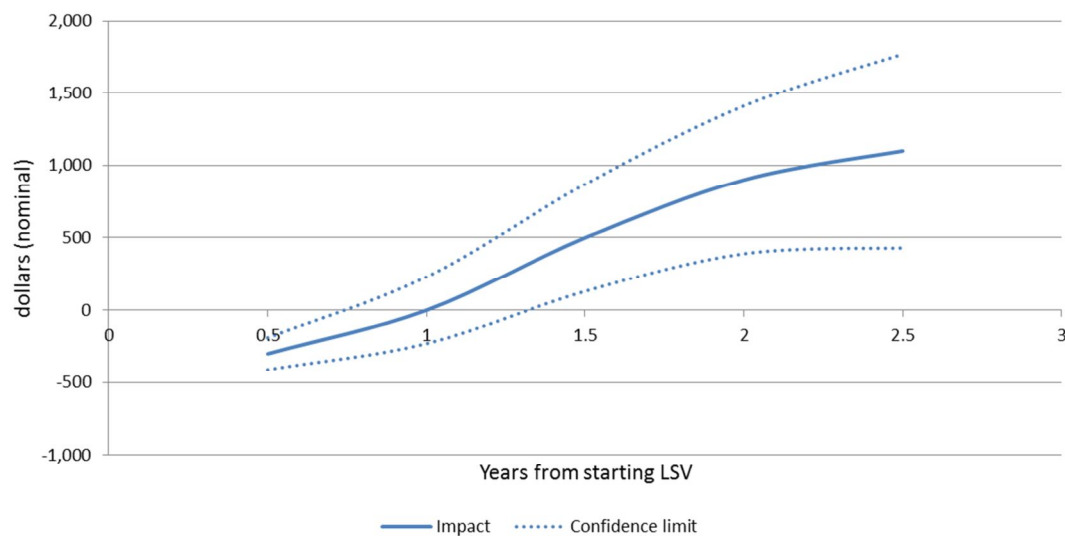
LSV increased time in employment and total income

The LSV programme increased participants' time in employment relative to the comparison group by an average of three weeks over 2.5 years. The programme also increased their total income by an average of \$1,300 over 2.5 years.

Figure 1 illustrates the positive impact on income by showing the difference in cumulative income at six monthly intervals after starting the programme. At 1.5 years, we are fairly confident that participants' total income is higher than the comparison group. This impact steadily increases up to 2.5 years from starting LSV. Because the impact on cumulative income is still increasing after 2.5 years we conclude we have not observed the full impact of the programme on participants' income.

¹ See page 22 for how the rating was done.

Figure 1: Difference in cumulative net income between LSV participants and the comparison group



Participants spent less time independent of welfare assistance

The LSV programme had a significant negative impact on the welfare assistance outcome. On average, participants were dependent on welfare assistance for 4.6 more weeks than the comparison group over three years. Participants are less likely to come off benefit while they are on the programme compared to the comparison group.

We describe this observation as a 'lock-in' effect. It is notable for the LSV programme since the participants are based at the military camp for up to 6 weeks. Therefore, when reading the result on welfare assistance, it is important to keep in mind that participants could have spent up to six weeks on main benefit. During that same period, the comparison group could use the time to look for employment, and are more likely to be independent of welfare assistance.

LSV may have reduced time spent in corrections services

The programme made little difference to the offending rate of participants (0.04 more offenses over 2.5 years). Participants spent 0.71 fewer weeks on average in prison and 3.6 fewer weeks in any corrections spell after 2.5 years, relative to the comparison group.

However, we are cautious about attributing these differences to the LSV programme because the comparison group had a previously higher level of corrections contact. The comparison group spent approximately 0.5 days to one day more on average in corrections services, relative to LSV participants before the start of the programme. For this reason, and despite our best endeavours, we cannot rule out the possibility that shorter corrections spells for the participants after the LSV programme is because of prior differences in the propensity towards offending among the two groups at the start of the programme.

LSV made little difference to educational outcomes

On average, participants spent slightly more time in education or training relative to the comparison group (1.4 weeks cumulatively over 2.5 years). However, the programme had a small negative impact on educational achievement. We found one percent fewer participants obtained a NQF level 4 qualification relative to the comparison group (see full report). With these results in mind, the overall effectiveness of the LSV programme on educational outcomes is inconclusive.

Figure 2: Primary outcomes: income, employment and off welfare summarised

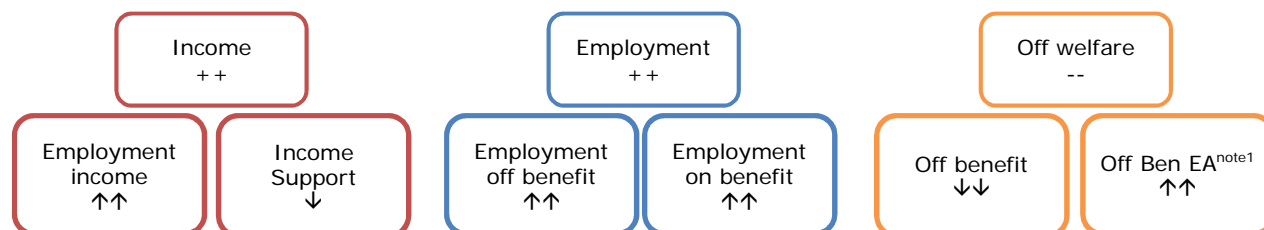
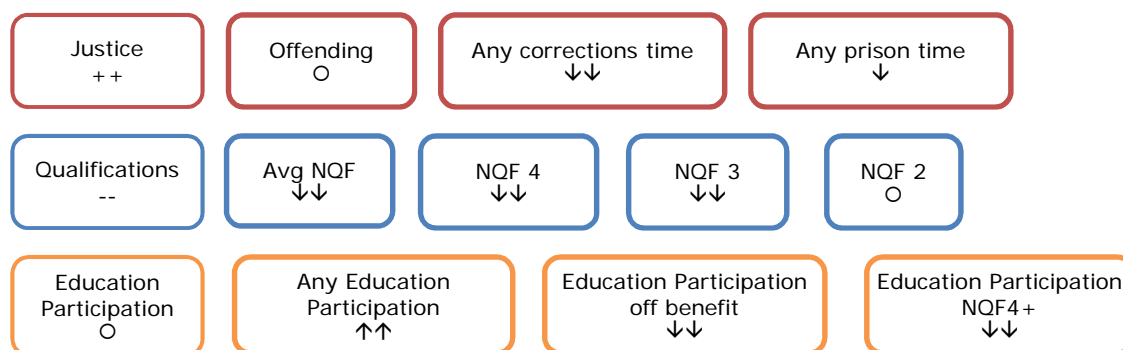


Figure 3: Secondary outcomes: justice and education summarised



++: effective, +: promising, 0: makes no difference, -: likely ineffective, --: ineffective

↑↑: increase, ↑: likely to increase, ○: no difference, ↓: likely to decrease, ↓↓: decrease

Note 1: Off benefit but receiving employment assistance eg a wage subsidy

See Rating the overall effectiveness of an intervention for details on how these ratings were constructed.

Outcomes at the subgroup level

In the following sections, we highlight some of the statistically significant differences between participants and the comparison group at the subgroup level. Subgroups considered include:

- gender – male or female
- age – 19 years and younger, or 20-24 years
- ethnicity – Māori or New Zealand European
- likelihood of long-term benefit receipt (LLTBR): this is a risk score of how likely a person will remain on benefit – low or medium.

The LSV programme has significantly stronger positive impacts on participants at higher risk of long term benefit receipt for income and employment outcomes. The programme

also has significant positive impacts on female participants' income. We could not detect differences in the impact on income and employment outcomes by age or ethnicity.

We found participants under the age of 20 years had poorer outcomes for educational achievement. Māori participants show more positive impacts from the programme compared to those of European ethnicity. In terms of justice outcomes, we observe a significant positive impact for males and older participants (aged 20-24 years).

Participants with medium LLTBR spent more time in employment

We observed significant differences in impact for income and employment outcomes between groups of different LLTBR levels. On average, participants with medium LLTBR levels spent 5.7 more weeks in employment over 1.5 years relative to people in the comparison group with similar levels of LLTBR (while being independent of welfare assistance). There was no impact for the same outcome for participants with low LLTBR levels.

Participants with medium LLTBR levels earned more net income on average (\$2,500) from employment relative to the comparison group over 1.5 years. We found no impact in net income from employment for participants with low LLTBR levels.

Participants with low LLTBR spent less time independent of welfare assistance

On average, participants with low LLTBR levels spent 3.6 fewer weeks being independent of welfare assistance over 1.5 years relative to the comparison group with similar levels of LLTBR. These participants also obtained more of their total income from income support relative to the comparison group over the same time period (\$500 more on average). People at lower risk of longer term benefit receipt are more impacted by the lock-in effect of the programme compared to those at higher risk.

Female participants earned more than those in comparison group

Female participants earned \$3,900 more on average compared to females in the comparison group over 1.5 years. While male participants did not show a significant impact on increasing income from employment, they did receive an additional \$600 in welfare payments, on average. Overall, gender differences for the effects of the LSV programme seem to be mixed when secondary outcomes are also considered.

Younger participants spent less time in education

Participants under the age of 20 years spent three fewer weeks in education or training at NQF level 4 and above relative to their peers in the comparison group after three years.

We found three percent fewer participants under the age of 20 years achieved an NQF level 3 qualification or higher relative to people in the same age band from the comparison group. This means younger participants were less likely to hold a high NQF qualification on average. There were no significant impacts on educational achievement outcomes for older participants.

Older participants spent less time in any corrections service

Participants between 20-24 years of age spent 7.1 fewer weeks on average in any corrections service compared to people of a similar age in the comparison group. We found no impact on time spent in any corrections services for younger participants.

Māori participants spent more time in education

The LSV programme had a significant positive impact on the education of Māori participants. On average, Māori participants spend a total of 2.9 more weeks in education or training after two years compared to Māori in the comparison group. LSV does not appear to have an impact on the same outcome for participants of European ethnicity.

Male participants spent less time in any corrections service

We observed gender differences for justice domain outcomes. On average, male participants spent 4.7 fewer weeks in any corrections service or 0.9 fewer weeks in prison compared to males in the comparison group over 2.5 years from starting the programme. There was no detectable impact on the same outcomes for female LSV participants.

Actionable insights

We think that increasing the proportion of participants who have a higher risk of remaining on benefit in the long-term may improve the impact of the programme on time spent independent of welfare assistance. This may reduce the lock-in effect and result in a larger post-participation effect.

This statement is based solely on the findings from the Outcomes at the subgroup level section report, and we are aware of other things that may impact on referral decisions. Therefore, we present this as something to consider when thinking about who would benefit the most from LSV. We also do not make any stronger recommendations because more information is needed about the social returns of the programme and the effect of the programme on the intermediate outcomes of its participants (for example, motivation and confidence).

LSV participant cost

The LSV programme is a high-cost intervention. We define whether an intervention is high-cost or low-cost according to the programme's cost per participant in comparison to that of other programmes run during the same time period. Each time a client participated in LSV in the 2014/2015 financial year, it cost MSD \$6,077. This value is in the highest 25 percent of costs when compared to other employment assistance programmes in that financial year.

Table 2 shows the estimated total cost of running the programme for three financial years. At that time, the number of participants remained relatively stable while there was a slight increase in cost from the 2012/2013 financial year to the 2013/2014 financial year.

Table 2 - Average cost per start from 2012/2013 to 2014/2015 financial years

Financial year	Total estimated nominal cost	Number of participant starts	Average nominal cost per start
2012/2013	\$7,688,382	1,460	\$5,266
2013/2014	\$8,481,832	1,380	\$6,146
2014/2015	\$8,453,671	1,391	\$6,077

Future work

MSD is currently working on analysing and reporting the cost effectiveness of employment assistance programmes. The first phase of the work focuses on producing return-on-investment results based on welfare costs (for example, costs to MSD). This is followed by return-on-investment results that also include other costs to society such as costs to the justice system.

Appendix 1: Summary of technical notes

Impact method

We estimate LSV's effectiveness using a counterfactual design. The term counterfactual refers to the question: what would have happened in the absence of the intervention? By definition, it is not possible to observe the counterfactual outcomes of participants. The solution is to identify a proxy for the counterfactual, usually a group of non-participants whose outcomes we use for comparison purposes. The challenge is to ensure that the comparison outcomes are an accurate representation of participants' counterfactual outcomes. Specifically, other than programme participation, are there other reasons for any differences between the outcomes of participants and those of the comparison group (ie selection bias)?

We have used a propensity matching technique to estimate the effectiveness of LSV on outcomes. Propensity matching estimates the counterfactual by constructing a matched group of non-participants who have the same (or similar) observed characteristics as the participants (eg similar duration on benefit, benefit type, and age).

Caveats

For the justice domain measures, we detected a pre-existing difference in the outcomes between the participant group and the comparison group up to 12 months before participation start date. This is very likely caused by the propensity matching process only using MSD administrative data. In future evaluations of employment assistance interventions, we plan to run the propensity matching process using variables from other government agencies as well.

Impact results and participant outcome values

Table 3 below shows all available impact results and participant outcomes organised by outcome domain. The Period (years) column shows the number of years (after starting LSV) that the impacts and participant outcomes are measured over. Both the Impact and Participants' outcomes values columns show the point estimate and the 95% confidence interval.

Table 3: Impact results by domain

Name of outcome	Period (years)	Impact	Participants' outcome values
Earnings (++)			
Net total income*	2.5	\$1,100 ± \$670	\$39,200 ± \$390
Income from employment*	2.5	\$1,300 ± \$780	\$25,900 ± \$460
Income from welfare payments	2.5	\$0 ± \$330	\$12,600 ± \$180
Employment (++)			
Any employment*	2.5	3.00 ± 1.29 wks	59.3 ± 0.7 wks
Employment off welfare*	2.5	1.86 ± 1.29 wks	48.0 ± 0.7 wks
Employment while on welfare*	2.5	1.14 ± 0.43 wks	11.3 ± 0.3 wks
Independent of Welfare Assistance (--)			
Independent from all welfare*	3	-4.57 ± 1.57 wks	86.7 ± 0.8 wks
Off main benefit*	3	-2.71 ± 1.57 wks	92.9 ± 0.8 wks
Participating in employment assistance*	2.5	1.59 ± 0.33 wks	5.3 ± 0.2 wks
Justice (++)			
Offending rate	2.5	0.04 ± 0.10	1.14 ± 0.07
Time spent in any corrections service*	2.5	-3.57 ± 0.86 wks	12.3 ± 0.4 wks
Time spent in prison*	2.5	-0.71 ± 0.29 wks	1.4 ± 0.1 wks
Education achievement (--)			
Average NQF level*	2	-0.08 ± 0.05	1.12 ± 0.03
% achieved NQF4 *	2	-1.00 ± 0.80 ppt	7.8 ± 0.50 ppt
% achieved NQF3*	2	-2.00 ± 1.00 ppt	20.2 ± 0.70 ppt
% achieved NQF2	2	-1.00 ± 1.00 ppt	34.9 ± 0.90 ppt
Education participation (++)			
Any participation in education*	2.5	1.43 ± 0.86 wks	27.4 ± 0.6 wks
Participation while independent from welfare*	2.5	-1.43 ± 0.71 wks	15.4 ± 0.4 wks
Participation in NQF level 4 or higher*	2.5	-1.43 ± 0.57 wks	7.7 ± 0.4 wks

Statistically significant impact results are marked by an asterisk (*)

++: effective, +: promising, 0: ineffectual, -: likely ineffective, --: ineffective

Ppt: percentage point (eg 50% minus 40% is 10 ppt)

Table 4 and Table 5 below also present impact and outcome results. Unlike Table 3 results are organised according to subgroup comparisons. For example, LSV can have a positive impact on one subgroup and a negative impact on another (see the last two rows of Table 4 where females obtain less welfare payments on average and males obtain more). Since there are many possible combinations of outcomes and subgroups,

we only included results where at least one of the subgroups had a statistically significant impact.

Table 4: Subgroup impact results (primary outcomes)

Name of outcome	Subgroup	Period (years)	Impact	Participants' outcome values
Any time in employment	LLTBR-low	1.5	0.4 ± 1.1 wks	33.9 ± 0.7 wks
	LLTBR-medium*	1.5	4.3 ± 2.6 wks	35.7 ± 1.7 wks
Employment while off welfare	LLTBR-low	1.5	0.1 ± 1.0 wks	27.1 ± 0.6 wks
	LLTBR-medium*	1.5	5.7 ± 2.4 wks	27.1 ± 1.6 wks
Independent from all welfare	LLTBR-low*	1.5	-3.6 ± 1.0 wks	41.4 ± 0.6 wks
	LLTBR-medium	1.5	1.4 ± 2.6 wks	40.0 ± 1.6 wks
Income from employment	LLTBR-low	1.5	\$100 ± \$680	\$15,200 ± \$400
	LLTBR-medium*	1.5	\$2,500 ± \$1,500	\$14,200 ± \$960
Income from welfare payments	LLTBR-low*	1.5	\$500 ± \$250	\$8,100 ± \$140
	LLTBR-medium*	1.5	-\$800 ± -\$660	\$8,300 ± -\$390
Income from employment	GENDER-female*	2.5	\$3,900 ± \$2,300	\$23,200 ± \$1,600
	GENDER-male	2.5	\$500 ± \$1,000	\$28,100 ± \$600
Income from welfare payments	GENDER-female*	2.5	-\$2,300 ± -\$1,400	\$16,300 ± -\$820
	GENDER-male*	2.5	\$600 ± \$340	\$11,400 ± \$200

Statistically significant impact results are marked by an asterisk ()*

It is also possible for LSV to have a positive or negative impact on one subgroup, and no impact on the other. See the first two rows of Table 4 for an example. We report the results for both subgroups even though we detected an impact for only one of them.

Table 5: Subgroup impact results (secondary outcomes)

Name of outcome	Subgroup	Period (years)	Impact	Participants' outcome values
% achieved NQF2	AGE-20-24	2.5	4.00 ± 4.00	53.00 ± 3.00 ppt
	AGE-low-19*	2.5	-3.00 ± 2.00	38.00 ± 1.00 ppt
% achieved NQF3	AGE-20-24	1.5	2.00 ± 3.00	31.00 ± 2.00 ppt
	AGE-low-19*	1.5	-3.00 ± 1.00	15.60 ± 0.90 ppt
average NQF level at school	AGE-20-24	2	0.07 ± 0.09	1.05 ± 0.06
	AGE-low-19*	2	-0.06 ± 0.04	0.79 ± 0.03
average NQF level	AGE-20-24	2.5	0.10 ± 0.10	1.71 ± 0.08
	AGE-low-19*	2.5	-0.10 ± 0.06	1.17 ± 0.04
Time spent in any corrections service	AGE-20-24*	3	-7.1 ± 2.6 wks	12.9 ± 1.4 wks
	AGE-low-19	3	-1.4 ± 1.4 wks	16.3 ± 0.9 wks
Participation in NQF level 4 or higher	AGE-20-24	3	0.0 ± 1.9 wks	9.7 ± 1.1 wks
	AGE-low-19*	3	-3.0 ± 1.0 wks	9.6 ± 0.6 wks
Participation while off main benefit	AGE-20-24	3	0.0 ± 2.0 wks	16.1 ± 1.3 wks
	AGE-low-19*	3	-3.7 ± 1.3 wks	19.6 ± 0.7 wks
Any participation in education	ETHNIC-European	2	-1.4 ± 2.0 wks	18.4 ± 1.1 wks
	ETHNIC-Māori*	2	2.9 ± 1.4 wks	25.6 ± 0.9 wks
Time spent in any corrections service	GENDER-female	2.5	0.3 ± 1.3 wks	4.0 ± 0.9 wks
	GENDER-male*	2.5	-4.7 ± 1.1 wks	14.9 ± 0.7 wks
Time spent in prison	GENDER-female	2.5	0.0 ± 0.1 wks	0.1 ± 0.1 wks
	GENDER-male*	2.5	-0.9 ± 0.4 wks	2.0 ± 0.3 wks

Statistically significant impact results are marked by an asterisk (*)

Appendix 2: Technical notes

This section provides more detail on the following:

- participation in LSV
- the outcomes measures used in the analysis
- methods used to estimate the impact of LSV
- method for estimating unobserved future impacts
- the process used to rate the effectiveness of LSV.

LSV participation

Participation in LSV is based on administrative records of successful referrals to the programme. We do not remove participants who may have ended the programme early. If a person participates more than once in a cohort year, we select the first commencement date.

Outcome measures

In the current report, we measured the impact of LSV across a range of outcome domains. Here we describe each outcome measure and how it was constructed.

Income

Net income from all sources

Net income from all sources is the main income outcome. It includes all sources of income, but excludes the drawdown of student loans. Income is net of tax. The measure was based on Inland Revenue and MSD data provided to the Statistics New Zealand Integrated Data Infrastructure (IDI). Current income information includes:

1. Employer Month Schedule (EMS): New Zealand operates a Pay As You Earn (PAYE) tax system. Accordingly, all employers provide Inland Revenue with monthly schedules of the earnings of all their employees. In addition to employee earnings, the Employer Month Schedule also includes taxable income support, accident compensation and pension payments.
2. Self-employment and company earnings: People who run their own business or company are also required to file annual tax returns. In the analysis, these annual returns are converted into monthly spells with annual total split equally across these months. There can be considerable lags in the lodging of self-employment earnings, that can mean measures of income for the most recent periods underestimate actual income. However, because we update the analysis on a regular basis the results incorporate these lags in reported earnings in subsequent updates.
3. Non-taxable income support payments: Not all income support payments are subject to tax. In particular, second tier assistance such as the Accommodation Supplement and third tier or hardship assistance such as Emergency Food Grants are not taxed. For hardship payments, we exclude Recoverable Assistance, as these are advances

on main benefits. Recoverable payments will either be reflected in lower main benefit payments, or, if the person moves off main benefit, in the form of an income support debt. At present, we do not have reliable data on income support debt.

Income from employment

From all income measures, the net income from employment measure only includes income from employment, self-employment, business income and paid maternity leave.

Income from welfare

The income from welfare measure includes all working age main benefits and superannuation payments as well as second and third tier assistance payments.

Employment

Any time in employment

Employment is based on the period that people declare income from employment or from self-employment. Note that employment spells are based on either monthly or annual periods so we may be over or understating the actual time a person is in employment depending on where in the month or tax year they started employment. At present, we have not attempted to adjust for this (for example, looking at the following or subsequent month to identify likely start and end periods).

There are also lags in lodging tax returns, with these most pronounced for annual returns. We choose not to censor our analysis period to accommodate these lags and instead rely on regular updates to the analysis to incorporate delayed tax data into the results.

Employment not on main benefit / on main benefit

Employment spells for a particular person can occur when the person is on a main benefit, or not on a main benefit. We analyse these conditions as separate outcomes. This enables us to analyse the proportion of time LSV participants spend in employment while being on a main benefit, as well as the proportion of time they spend while *not* being on a main benefit.

Independent from welfare

Independent of Work and Income assistance

We measured the time people are dependent on welfare assistance by the period they were entitled to a main benefit and whether they were participating in employment assistance interventions. The inclusion of the latter is to cover instances where people are receiving employment assistance whilst off main benefit (for example a wage subsidy).

A limitation of this measure is that it fails to account for negative destinations. For example, people who move from main benefit into prison would appear to be off welfare assistance. In subsequent versions of this report, we plan to include time in Corrections service well as other negative destinations into this measure.

Off main benefit

Time spent off main benefit is a simple measure of the time people are not entitled to a main benefit (for example, Jobseeker Support, Sole Parent Support or Support Living Payment).

Off main benefit and participating in employment assistance interventions

Time that people participate in employment assistance interventions such as wage subsidies and not on main benefit.

Justice

We have two sources of information on justice outcomes: police offending and periods under Corrections supervision. These data are also obtained from SNZ IDI.

Any offence

This measure is based on police data of people who are arrested for an offence (but may not result in a prosecution). Note that offending data is only available from 2009. We then calculate offending rates for participant and comparison groups in each month from the start of LSV.

Corrections spell

Any spell under corrections supervision and covers periods of custodial and non-custodial supervision (such as prison, community service, home detention, remand, parole and eriodic detention).

Prison spell

An important subset of Corrections supervision spells is time spent in prison. Accordingly, we also include time spent in prison as a separate outcome.

Education qualifications

Educational achievement information is based on secondary and tertiary qualifications achieved. We include school, tertiary, industry training and targeted training qualifications data. There is a considerable reporting lag for qualifications data in the SNZ IDI, normally qualifications data are out of date by over 12 months. In addition, qualification data only provides the year the qualification was attained. In our analysis, we make the assumption that the qualification was attained at the end of the year (ie 31st of December).

Qualifications achieved at NQF level 2, 3, 4

For each person, we construct spells when they have achieved a specified minimum NQF level. NQF levels start at one (first school national assessment) through to nine (doctorate). For each individual, we identify the date they first achieved the specific NQF level.

Highest NQF level

The highest NQF level is the highest NQF level achieved by a person at a specific date. From this measure, we can calculate the average NQF level achieved by the participant and comparison group of an employment assistance intervention.

Education participation

Participation in further education and training provides an early indication of whether people are engaged in developing their human capital. The unit of measurement for this outcome is time in number of days.

Any education participation

For any education participation, we combine all education spells in school, tertiary, industry training and targeted training.

Education participation off benefit

Education participation spells where a person is also off main benefit (based on benefit entitlement spells).

Education participation NQF4

Time spent participating in education courses at NQF level 4 or above (broadly equivalent to University degree level).

Effectiveness rating

Rating the effectiveness of LSV is a two-step process. The first step is to estimate the observed impact of LSV on participants' outcomes to date. The second step is to apply standard rules to determine the effectiveness rating of LSV.

Estimating the observed impact of LSV

The first step in rating the effectiveness of LSV is to determine the impact of the programme on outcomes to date. In this analysis, we estimate effectiveness using counterfactual designs. The term counterfactual refers to the question: what would have happened in the absence of the intervention?² By definition, it is not possible to observe the counterfactual outcomes of participants. The solution is to identify a proxy for the counterfactual, usually a group of non-participants whose outcomes we use for comparison purposes. The challenge is to ensure that the comparison outcomes are an accurate representation of participants' counterfactual outcomes. Specifically, other than programme participation, are there other reasons for any differences between the outcomes of participants and those of the comparison group (ie selection bias)?

There is no foolproof means to remove selection bias; rather, various methods are able to control for it to a greater or lesser degree. To assist readers on judging the robustness of a particular counterfactual design, we categorise methods according to the Scientific Maryland Scale (SMS). The SMS scale ranks counterfactual designs from 1 (least robust) to 5 (most robust). Robust in this context refers to the level of confidence we have that the impact estimate of a design provides an accurate measure of the quantitative causal effect of the intervention on the outcome.

In the current report, we used a propensity matched comparison group and this is considered to be an SMS level 3 design. To substantially increase our confidence in the

² It is important to emphasise that counterfactual designs are not the only or primary evaluation method. To fully understand the effect of an intervention requires both quantitative estimation of its impact on outcomes, but equally important is information on the context and the operation of the intervention itself to understand why the intervention has the impacts that it does.

estimated impact of LSV would require implementing a Randomised Control Trial (SMS level 5) evaluation.

Propensity matching

Propensity matching is a common alternative to Randomised Control Trial (RCT). It estimates the counterfactual by constructing a matched group of non-participants who have the same (or similar) characteristics as the participants.

Before outlining propensity matching, it is useful to think of an intuitively appealing alternative of exact matching. Exact matching, as the term suggests, is to match a participant to a comparison who has the same characteristics (eg same age, gender, benefit history and so on). However, exact matching is limited by the probability that two people share the same set of observable characteristics (and it is also unnecessarily restrictive).³ The more characteristics included in the exact match, the less likely it is to find a comparison person with the exact same characteristics for each participant. As a result, these methods require the arbitrary selection of only a few matching variables.

Propensity matching overcomes this problem by using a logistic regression model to relate observable characteristics to participation in LSV. The logistic regression produces an estimate of the probability that a given individual is a participant of LSV. It is possible to use this probability (called “the propensity score”) to match participants and non-participants based on the similarity of their propensity scores. If the propensity score is properly specified, the participants and matched comparison groups will have a similar observable characteristic profile (eg similar duration, benefit type, age, number of children).

Conditional Independence Assumption

The Conditional Independence Assumption (CIA) states that controlling for differences in observable characteristics between the participant and comparison groups also controls for unobserved differences between the two groups. Estimating the impact by controlling for observable characteristics requires that the CIA hold. If it holds, the only statistically significant difference between the participant and comparison groups will be their participation in LSV. Any resulting estimates would be unbiased. In other words, the only explanation of differences in outcomes between the two groups would be whether they participated in the programme. If the CIA fails, the estimate will be biased. Here differences in outcomes could be due to unobserved differences between participants and their comparisons, as well as the impact of the programme.

The main limitation of the propensity matching method is that it relies on available and measurable information about people likely to participate in the employment assistance intervention. It is rare that comprehensive information exists about the types of people who participate in the programme or those who could form part of the comparison group. The analysis relies on the information available on MSD’s administrative databases. This increases the risk of biased estimates. The second limitation of the CIA is that it is not possible to determine whether it has been violated or to what extent if it has.

³ Within a randomised control treatment group, the two groups share the same statistical profile, not that each treatment group member has an identical twin in the control group.

Table 6 summarises the variables included in the propensity matching. The emphasis is on historical variables and, in particular, the four years prior to the start date.⁴

Table 6: Observable characteristics included in the propensity matching of the comparison group

Area	Variable	Presentation of variable in the analysis
Demographics	Gender	Female, Male
	Age	Age in years Age group (16–<18 yrs, 18–<20 yrs, 20–<25 yrs, 25–<30 yrs, 30–<35 yrs, 35–<40 yrs, 40–<45 yrs, 45–<50 yrs, 50–<55 yrs, 55–<60 yrs, 60–<65 yrs)
	Ethnicity	Māori, NZ European, Pacific people, Other, Unspecified
Residency	Migrant	Yes, No
	Current Migrant	Yes, No
	English preferred	Yes, No
	Refugee	Yes, No
	Time in NZ	1–2 yrs, 4–8 yrs, 8–12 yrs, 12+ yrs, New Zealand
Labour market skills	Education	None; NCEA Lvl 1, <80 credits, NCEA Lvl 1, 80+ credits; NCEA Lvl 2; NCEA Lvl 3; Other school qualifications; NCEA Lvl 4; Post-secondary; Degree/prof qualifications
	Numeracy literacy barrier	Yes, No
	Language verbal barrier	Yes, No
	Income in six months prior to benefit commencement	No income, Under \$250, \$250 to \$499, \$500 to \$749, \$750 to \$999, Over \$1,000
	Family status	Individual has an identified partner
	Age of youngest child	0–5 yrs, 6–13 yrs, 14+ yrs, No child
	Number of children	Categorical (ie No child, 1 child, 2 children, etc)
Health and disability	Employment barriers identified: Disability, Alcohol and drug, Intellectual, Mental illness, Mobility and agility, Sensory, Unspecified (7 variables)	Yes, No
	Number of current incapacities	0 incapacity, 1 incapacity, 2 incapacities, 3 incapacities, 4 incapacities
	Primary incapacity	Unspecified, No incapacity, Cancer, Intellectual, Schizophrenia, Congenital, Alcohol, Anxiety, Anxiety Depression, Circulatory NFD, Circulatory Other,

⁴ Start date refers to the date participants commenced the programme (the actual date is usually three days prior to recorded participation start) or the date the non-participants were selected for inclusion in the comparison group.

Area	Variable	Presentation of variable in the analysis
		Depression, Diabetes, Drugs, Endocrine Other, Heart Disease, Infectious Parasitic, Mental Other, Nervous Epilepsy, Nervous Other, Non Organic Psychoses NFD, Stress, Nervous Hearing, Nervous Sight, Stroke, Blood Diseases, Mental NFD, Bipolar, Genitourinary, Injury NFD, Injury Other, Musculoskeletal NFD, Respiratory NFD, Vertebral Column, Skin, Digestive, Musculoskeletal Other, Pregnancy Normal, Pregnancy Complications, Arthropathies Osteopathy, Fractures Dislocations, General, Respiratory COPD, Rheumatism Not Back, Strains Sprains. Respiratory Other
	Current incapacity 1 to 4 (4 variables)	Same as primary incapacity
	Identified incapacity in the previous five years: Unspecified, No incapacity, Cancer, Intellectual, Schizophrenia, Congenital, Alcohol, Anxiety, Anxiety Depression, Circulatory NFD, Circulatory Other, Depression, Diabetes, Drugs, Endocrine Other, Heart Disease, Infectious Parasitic, Mental Other, Nervous Epilepsy, Nervous Other, Non Organic Psychoses NFD, Stress, Nervous Hearing, Nervous Sight, Stroke, Blood Diseases, Mental NFD, Bipolar, Genitourinary, Injury NFD, Injury Other, Musculoskeletal NFD, Respiratory NFD, Vertebral Column, Skin, Digestive, Musculoskeletal Other, Pregnancy Normal, Pregnancy Complications, Arthropathies Osteopathy, Fractures Dislocations, General, Respiratory COPD, Rheumatism Not Back, Strains Sprains. Respiratory Other	Yes, No
	Invalid's Benefit reassessment period	Never, 2 years, 5 years, Not indicated, Not applicable
	Medical assessment of time until part time work	Now, <1 month, 1-<3 month, 3-<6 months, 6 or more months, Unlikely in the foreseeable future, No indication, Not applicable.
	Medical Assessment of time to selected duties	
	Medical Assessment of time to work planning	
Labour market context	Territorial local authority area	64 categories
	Work and Income region	12 categories
	Quarter of start date	2004Qtr1, 2004Qtr2, 2004Qtr3, etc

Area	Variable	Presentation of variable in the analysis
Other	Ex-prisoner Time since last prison event	Yes, No No duration, < 3 months, 3-6 months, >6 mths-1 yr, >1-2 years, >2-3 years, >3-4 years, >4-5 years, >5-6 years, >6-8 years, >8-10 years, Over 10 yrs
Independence from Work and Income Assistance	Dependent on Work and Income Assistance in each of the 48 months prior to start date (48 variables)	Yes, No
Benefit information	Current benefit	Unemployment/Independent Youth, Domestic Purposes/Widow's/Emergency, Sickness, Invalid's, Supplementary only, No benefit
	Primary status	Primary, Partner, Single
	Current benefit status	Current, Cancelled, Suspended, Registered, No benefit
	Duration on current benefit	Categorical (<=3 months, >3-6 months, >6-12 months, >1-2 years, >2-4 years, >4-6 years, >6-8 years, >8-10 years, Over 10 yrs, No duration) Continuous (days)
	Continuous duration on benefit	
	Duration off-benefit	Categorical (<=3 months, >3-6 months, >6-12 months, >1-2 years, >2-4 years, >4-6 years, >6-8 years, >8-10 years, Over 10 yrs, On benefit) Continuous (days)
	Last benefit	On benefit, Unemployment/, Independent Youth, Domestic Purposes/Widow's/Emergency, Sickness, Invalid's, Supplementary only, No benefit
	Years on main benefit over previous 10 years	Categorical (0 years, <1 year, 1 year, 2 years, ..., 10 years)
	OnBenAt18	Yes, No, Too old
	Benefit status in each of the 48 months prior to start date (48 variables)	Unemployment, DPB related, Sickness, Invalid's, NZSuper Vets TRB, Widow's, Youth, No benefit
Duration on each main benefit group:	Unemployment/Independent Youth, Domestic Purposes/Emergency, Widow's, Sickness, Invalid's	Categorical (<=3 months, >3-6 months, >6-12 months, >1-2 years, >2-4 years, >4-6 years, >6-8 years, >8-10 years, Over 10 yrs, No duration) Continuous (days)
Register duration	Current register duration (if participated before 2007)	Categorical (<=3 months, >3-6 months, >6-12 months, >1-2 years, >2-4 years, >4-6 years, >6-8 years, >8-10 years, Over 10 years, Unspecified) Continuous (days)

Area	Variable	Presentation of variable in the analysis
Employment programme participation	Current participation in: Into-work support, Job search, Matching and placement, Training, Wage subsidy, Work confidence, Work experience, Other (8 variables)	Yes, No
	Participation in the previous 5 years in: Into-work support, Job search, Matching and placement, Training, Wage subsidy, Work confidence, Work experience, Other (8 variables)	No participation, Under 1 month, 1 to 3 months, 3 to 6 months, 6 months to 1 year, 1 to 2 years
	Programme participation in each of the 48 months prior to start date (48 variables)	Into-work support, Job search, Wage subsidy, Work confidence, Work experience, Training, Matching and placement, Other, No participation
Participation in tertiary study	Received student loans or allowances in each of the 48 months prior to start date (48 variables)	Yes, No
	Proportion of time receiving student loans and allowances in last 5 years or since 2000	Categorical (0 years, <1 year, 1 year, 2 years, ..., 5 years)
Part-time work	Average weekly declared earnings in each of the 48 months prior to start date (96 variables)	Categorical (No income, >\$0–\$80, >\$80–\$180, >\$180–\$300, >\$300) Continuous (nearest dollar)

Rating the effectiveness of interventions

The last step in the process is to systematically rate the effectiveness of interventions based on their impacts on selected outcomes. The goal here is to ensure that all employment assistance interventions (including LSV) are rated in the same way and that the rating process is transparent.

Rating by outcome domain

For each employment assistance intervention, we have one outcome measure grouped under each broad outcome domain. In the current effectiveness report, we focus on three outcome domains: income, employment and independence from welfare.

At present, we select one outcome measure to provide the summative assessment for the impact of each employment assistance intervention on that domain. In the current analysis:

- income effectiveness is based on the LSV's impact on net income from all sources
- employment effectiveness is based on the impact on any time in employment
- independence from welfare assistance is based on time spent independent from Work and Income Assistance (ie not on main benefit or participating in employment assistance interventions).

Rating the overall effectiveness of an intervention

Once we have an effectiveness rating for each outcome domain we then combine these ratings to arrive at an overall rating of a programme. Because we are combining three outcome domains, the number of combinations of results becomes much greater.

Table 7 shows how we rate employment assistance interventions based on the rating for one or more of the outcome domains as well as the observed outcome period. The Domain 1 to Domain 3 columns can refer to any combination of the three outcome domains used in our analysis, the focus here is on the combination of positive and negative impacts between the three.

Table 7: Employment assistance intervention effectiveness rating code table

Outcome domain			Outcome period	Rating
Domain 1	Domain 2	Domain 3		
++			Any period	Effective
0	++	++	Any period	Effective
+	++	++	Any period	Effective
++	++	++	Any period	Effective
0	+	++	2+ years	Promising
0	0	++	2+ years	Promising
--	--	++	2+ years	Mixed
--	++	++	2+ years	Mixed
--	-	++	2+ years	Mixed
--	0	++	2+ years	Mixed
-	0	++	2+ years	Mixed
-	++	++	2+ years	Mixed
-	+	++	2+ years	Mixed
--	0	+	2+ years	Mixed
--	+	++	2+ years	Mixed
0	0	0	2+ years	Makes no difference
0	0	+	2+ years	Makes no difference
0			2+ years	Makes no difference
--	0	0	2+ years	Likely negative
--	--	+	2+ years	Likely negative
--			2+ years	Negative
--	--	--	2+ years	Negative
--	--	-	2+ years	Negative
--	-	-	2+ years	Negative
--	--	0	2+ years	Negative
--	-	0	2+ years	Negative

Outcome domain rating: ++: effective, +: promising, 0: no difference, -: likely to be negative, --: negative.

Effective: Employment assistance interventions are rated effective only if they are effective against the majority of outcome domains and they show no sign of having a negative impact on any other outcome domain. We do not wait two years before rating a programme as effective.

Promising: Promising programmes are those that are effective or likely effective for at least one outcome, and show no negative effects. We wait until we have two years of outcome data before rating an intervention as promising.

Mixed: Mixed covers interventions that show both positive and negative effects across outcome domains. We wait until we have two years of outcome data before rating a programme as mixed.

Makes no difference: Includes all employment assistance interventions that have no effect on any outcome domain. We wait until we have two years of outcome data before rating a programme as making no difference.

Likely negative: Interventions are in this group because either a minority of outcome domains are rated as negative with the remainder having no impact. Or, the majority are negative, with a minority having the possibility of being positive. We wait until we have two years of outcome data before rating a programme likely negative.

Negative: Interventions where the majority of outcome domains are rated as negative. We wait until we have two years of outcome data before rating a programme negative.

Too soon to rate: With the exception of interventions rated as effective, interventions with less than two years of observed impacts are rated as too soon to rate. The reason for waiting at least two years is that the majority of employment assistance interventions have negative effects in the short-term (for example lock in effects) and it is necessary to wait sometime after commencement before positive effects are potentially observed.