

In-Work Support and Payment trial

Impact evaluation

November 2020



**MINISTRY OF SOCIAL
DEVELOPMENT**
TE MANATŪ WHAKAHIATO ORA

Authors

Waitai Rakete, Analyst, Ministry of Social Development

Marc de Boer, Principal Analyst, Ministry of Social Development

Acknowledgements

The authors wish to thank the following people for their contributions and comments in preparing this report: Bryan Ku, Clare Dominick, Kerrie O’Hara Pepper, Paul Merwood, Sarah Stacy-Baynes, Tim Maloney, David Rea, Jacob Human, Ella Cullen, Sarah Chandler and Kecia Painuthara.

Disclaimer

The views and interpretations in this report are those of the researchers and are not the official position of the Ministry of Social Development.

Published

Ministry of Social Development

PO Box 1556

Wellington

www.msd.govt.nz/insights

Published in November 2020

ISBN

978-1-99-002308-8

Contents

Contents	iii
Executive summary	6
Introduction	8
In-Work Support (IWS) and the In-Work Payment (IWP)	8
Impact method	10
Findings	14
Conclusions.....	20
Appendix 1: Recruitment into IWS process.....	22
Appendix 2: Participant demographics	23
Appendix 3: A predicted treatment variable and covariates for a regression analysis.....	27
References.....	1

Abbreviations

Abbreviation	Definition
CSRs	Customer service representatives
HEEC	High Entry-Exit Clients
IWP	In-Work Payment
IWS	In-Work Support
JS	Jobseeker Support
JS-HCID	Jobseeker Support – Health Condition, Injury or Disability
MSD	Ministry of Social Development
RCT	Randomised Control Trial
SLP	Supported Living Payment

Glossary of terms

Term	Definition
Cancelled benefit	A person's main benefit is cancelled when their circumstances change to make them ineligible eg due to starting full-time permanent employment.
Contact rate	The number of eligible people who are successfully contacted to be asked whether they want to participate in the service.
Control group	People who met the criteria to opt into the IWS and were randomly assigned to the group not to be invited to participate.
Counterfactual	The expected outcomes of people if they had not participated in the service. In this study, the counterfactual was represented by the outcomes of the control group.
Detectable impact on days off main benefit	There was a statistically significant difference (95% confidence interval) between the treatment and control groups in the average number of days spent off main benefit.
High entry-exit clients (HEEC)	High entry-exit clients are people who have shown a tendency to enter and exit a main benefit at a higher rate than usual. For the trial, HEEC was defined as being on and off a main benefit at least three times in the last two years.
Impact evaluation	A type of evaluation used to quantify the causal effects of the service on participants' outcomes, both intended and unintended. Causal effects are estimated using a counterfactual design.
Opt-in rate	Proportion of people who are invited to participate in the service and agree to do so.
Partial compliance	Some of the control group received the service.
Participation rate	Of the eligible population, the proportion who participate in the service. The participation rate is determined by the contact and opt-in rates.
Randomised control trial	An impact study that randomly allocates people into a treatment who receive an intervention or control group who do not. Randomisation ensures that, at

Term	Definition
	assignment, the two groups are statistically equivalent in their expected future outcomes. On this basis, any difference in actual outcomes can be attributed to the difference in the experience of the treatment group compared to the control group.
Statistical significance	When an observed result is unlikely to occur by chance alone. In this report, we used a p-value (probability value) of less than 0.05 as the threshold for statistical significance.
Suspended benefit	A client's main benefit can be suspended because of a temporary change in their circumstance that makes them ineligible to receive payments (eg have a brief period in work).
Treatment group	People randomised into the treatment group are allowed to opt into the IWS service.
Upper- and lower-95 per cent confidence intervals	A 95% confidence interval is the range where we are 95% confident the true value lies. The wider the range the less certain we are. The size of the confidence interval is most influenced by the number of observations, with larger samples resulting in smaller confidence intervals.
Zelen design	Eligible people are randomised into a treatment group or a control group before being contacted to take part. The treatment group is then contacted and invited to opt into the service.

Executive summary

This report presents findings from the impact evaluation of the In-Work Support service and In-Work Payment trial.

The In-Work Support service sought to help people who cycled between benefit and employment

The In-Work Support (IWS) service trial began in February 2015 and sought to improve long-term employment outcomes for a select group of Jobseekers who had frequently left and then returned to main benefit. The service aimed to help those people settle into a job, adjust to working, and remain in work by providing them with information, advice, and referrals to appropriate agencies. IWS was delivered by a team of Customer Service Representatives (CSRs) who contacted participants via phone, email, and text messages over 12 months.

An In-Work Payment (IWP) was included from April 2015

In April 2015, a one-off In-Work Payment (IWP) of \$3,000 was introduced to further encourage people to participate in the service and remain in employment. The IWP was provided as four staggered payments to participants who remained in employment for up to one year.

The trial evaluation focused on people who met the High Entry-Exit Client criteria only

All MSD clients who exited a main benefit could receive the service, but only people who met the High Entry-Exit Clients (HEEC) criteria were eligible for the IWP. The trial and this study only examined the impact of IWS for people who met the HEEC criteria and not for any other participant group.

The trial and the IWP ended in July 2019

In July 2019, the trial and IWP were discontinued but the IWS service was still available to all clients exiting a main benefit into employment.

The study used an RCT to estimate the impact of IWS/IWP

This evaluation sought to estimate the impact of IWS/IWP on the time that eligible people remained off main benefit in the following year. To answer this question, we used a Zelen randomised control trial (RCT) design. This involved randomising 6,693 eligible individuals into a treatment group who were to be invited to participate in the service. An additional 6,734 individuals were assigned to a control group who were not invited to participate in IWS/IWP for 52 weeks after assignment. The control group could still receive existing in-work support services.

An RCT design was chosen for the trial as it is the most robust method for counterfactual impact estimation. Also, the number of places available for the IWS/IWP exceed the number eligible, and therefore, the RCT could be incorporated as part of the process for selecting a manageable group of service participants.¹

The study was unable to determine whether IWS was effective

Because of the Zelen design, the following results are for all people eligible to participate in the service, of whom the participants were a subset.

We found that for people who were eligible to participate in IWS and receive the IWP:

- only 13% were successfully contacted, of whom 75% opted-in (agreed to participate), resulting in a low participation rate of 10%
- in the following year, the average time off benefit increased by 3.2 (-0.5, 6.9) days.

The small and non-significant impact from the study was primarily caused by the low contact rate reducing the statistical power of the study. On the other hand, of those contacted, a high proportion of people did agree to participate and receive the IWP.

We identified the following reasons for the low contact rate.

- Staff focused on contacting those who had cancelled rather than suspended benefits. People on suspended benefits made up 63% of the eligible group.
- Those eligible were only contacted after they had exited from a main benefit and contacts were made during work hours.
- Contacts were made through a private number. If a person missed the call, they could not easily return the call.²

Lessons learnt

The IWS trial provided us with some important lessons on how to trial and evaluate a new service.

- It is important to check the international literature on evidence and identify features that make the proposed service successful. In the IWS context, in-work support services generally recruit participants before, rather than after, benefit exit. Contact before exit results in a higher recruitment rate and gives the participants the additional incentive to search for employment to qualify for the IWP.
- After completing the service design, the next stage is to implement the service with the primary focus on ensuring it is operating as intended and make any adjustments as required.
- Only once the service is stable and operating as intended should an impact evaluation commence. At this stage all parties involved in the study need to agree to adhere to the RCT protocol.

¹ In the absence of any information on whether IWS is more effective for one group or another, a simple random selection from among the eligible group is an ethical approach.

² For privacy reasons (eg shared phones) CSRs could not leave a message.

Introduction

This report presents findings from the impact evaluation of the In-Work Support service and In-Work Payment trial. The report is divided into the following sections:

- description of the In-Work Service (IWS) and the In-Work Payment (IWP)
- outline of the impact methodology and its implementation
- findings from the analysis of the IWS's impact on off benefit outcomes
- conclusions and lessons learnt.

In-Work Support (IWS) and the In-Work Payment (IWP)

IWS was designed to reduce cycling between job seeker benefits and employment

Analysis of people on Jobseeker Support Work Ready (JS-WR) main benefits identified a group who frequently moved between being in employment and on main benefit. These people were defined as High Exit-Entry Clients (HEEC).

The In-Work Support service was designed to help people who met the HEEC criteria to increase the time they spent in employment. The trial began in February 2015 and ended in July 2019.

IWS service was available to non-HEEC clients, but these were not included in the trial

In addition to people who met the HEEC criteria, the IWS service was also available to all people exiting main benefit into employment. However, the trial and the findings in this study only apply to the people who met the HEEC criteria. For this reason, all references in this report to IWS are specific to the trial participants.

IWS was a phone-based support service

The objective of the IWS service was to help increase the time participants spent in employment after exiting from main benefit. IWS was an in and out-bound calling service run by a team of 13 Customer Service Representatives (CSRs). The CSRs provided support and referrals to services to help participants settle into a job, adjust to working, and remain in work. Participation in the service was voluntary.

The CSRs contacted participants in several ways, including up to 13 phone calls, at least eight email contacts, and unlimited texts to each participant over 12 months. Participants could also initiate contact with the service.

Types of assistance included:³

- advice about financial support a participant may qualify for
- information about other services such as budgeting or relationship services
- regular check-ins to see how things were going at work
- work-related advice such as learning the new job, training and mentoring
- advice on how to manage a disability or health condition while working including workplace modifications.

The In-Work Payment was included to increase opt-in and encourage participants to remain in employment

As part of the Government's manifesto commitments, the In-Work Payment was introduced in late April 2015, to:

- encourage eligible clients to participate in the IWS service
- provide an incentive to remain in employment.

The IWP payments were made while participants remained in employment, with:

- \$500 at the end of the first month (28 days)
- \$500 at three months (84 days)
- \$1,000 at six months (168 days)
- \$1,000 at 12 months (336 days).

The IWP was only available to participants who met the HEEC criteria and were part of the IWS trial. Because the IWP was included soon after the IWS started, it was not possible to compare the impact of the IWS with and without the IWP.

Eligibility for the IWS and IWP

People eligible for the trial had the following characteristics:

- were on Jobseeker Support (JS) with either a current, suspended or expired benefit
- were aged 18 to 64
- their current spell off a main benefit was at least 14 days.

From January 2018, participants whose main benefit was suspended were excluded from selection for the trial.

HEEC criteria definition

To meet the HEEC criteria a person had to have:

- been on and off a main benefit at least three times in the previous two years
- had at least one off-main benefit spell lasting three months or longer.

³ See <https://www.workandincome.govt.nz/work/job-support-and-advice/in-work-support.html>

Impact method

The IWS trial used a Randomised Control Trial (RCT) to identify the impact of being eligible to participate in the service on subsequent outcomes. This section provides more detail on the approach and its implementation.

Counterfactual approach to estimating impacts

RCT is based on the counterfactual framework for inferring causal effects. Counterfactual is the question, what outcomes would people eligible for IWS have achieved if the service was not available? Any difference in outcomes between these two scenarios, those who participated in the service and those who did not, would tell us the causal impact of the IWS on the eligible group's outcomes.

An RCT answers the counterfactual question by randomly assigning people to a control or treatment group. Because the assignment is based on a random draw (eg a coin flip), the two groups should have the same average profile and, most importantly, are expected to achieve similar outcomes in the future. Based on the assumption of similar expected outcomes, any differences in actual outcomes between the treatment group and control group can be attributed to differences in experience between the two. Namely, the treatment group receiving the intervention being evaluated while the control group received services as usual.

Ethical considerations in using an RCT

Unlike observational studies, RCTs have a direct impact on the allocation of services to individuals and through these services, alter their outcomes and well-being.

For this reason, we need to consider the ethics of both running an RCT as well as the IWS/IWP service itself. Starting with the IWS/IWP service, these services were authorised by elected officials (Ministers) and implemented by public officials under due delegation and oversight. Also, people could voluntarily participate in the service and choose to opt-out at any time.

The ethical question for the RCT is whether researchers have strong prior evidence to show that being assigned to the control or treatment group would result in better or worse outcomes. In the case of IWS/IWP there was little existing literature on the effectiveness of in-work support programmes in New Zealand or whether they would have a net social benefit (ie the benefits of the intervention would outweigh its cost).

A secondary justification for the use of the RCT is where a programme is oversubscribed. Here it will be necessary to select some people to receive the service and not others. In the absence of any prior information on who is likely to benefit, random selection for the evaluation provides equality of access as well as a vehicle through which to better understand who the intervention should be targeted at, if at all.

The Zelen design

For this study we used a specific design called a randomized consent design (also referred to as a Zelen design) (Torgerson & Roland, 1998). The key difference between a

Zelen design and the more common intention to treat designs is that randomisation occurs **before** people are contacted to be asked to participate (consent) in the intervention.

There are three advantages to the Zelen design:

- Impact estimates include both the effect of the IWS on those who chose to participate but also helps identify any impacts on those who do not.
- It does not require lengthy recruitment of people and having them agree to participate in both the service as well as the RCT, so avoids randomisation bias.⁴
- It is operationally easier to incorporate the randomisation at the identification of eligible people rather than needing to train frontline staff to recruit people into the study.

There is debate on the ethics of Zelen design whereby people are randomised without their consent. However, this consideration is offset against avoiding hardship by recruiting and then assigning people to a control group. In the context of the IWS, the Zelen design was justified based on the need to select a subset of eligible people for recruitment into the service.

Offsetting these advantages is the loss of statistical power. The loss of statistical power becomes more acute when either:

- the effect of the service is small (ie the change in outcomes is modest relative to other factors)
- the participation rate (ie the proportion of eligible people who participate) is low, diluting the impact of the service on participants across many non-participants.

Unfortunately, both these issues arose in the context of this study.

Implementation of the Zelen design RCT

For the IWS trial, Figure 1 summarises the allocation of people eligible to participate in the trial into the control and treatment groups. Over the trial period, a total of 13,427 people were identified as eligible. The randomisation was one to one so that the treatment group had 6,693 people and the control group had 6,734.

Assignment to each group meant:

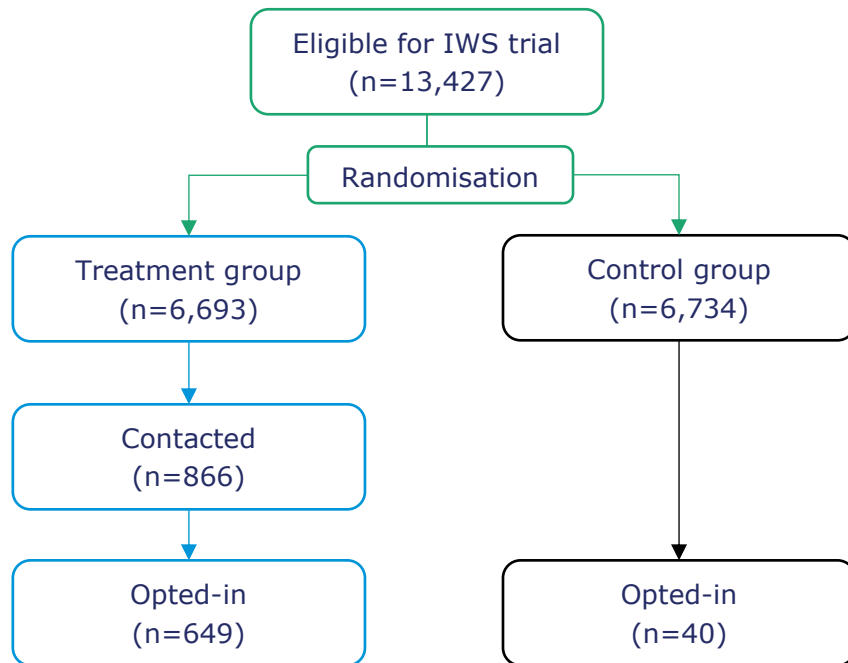
- **Treatment group:** people assigned to the treatment group would have their contact details sent to the IWS CSRs for recruitment into the service. Over the study period, people could only be assigned to the treatment group once. This was to avoid the incentive of getting the IWP and then returning to benefit to requalify for the payment again.

⁴ Randomisation bias can come in two forms. The first is that not everyone who would agree to take part in IWS/IWP would also agree to take part in the RCT. The second source is that people who are assigned to the control group change their behaviour in response to missing out on the service. For example, missing out on receiving the in-work payment may reduce the motivation to remain in work.

- **Control group:** people in the control group would have their contact details withheld from the contact lists for the IWS service for 52 weeks. Control group members continued to receive any in-work and related assistance they were entitled to other than IWS/IWP. After the holdout period, control group members could be re-randomised to either the treatment or control group if still eligible.

The impact evaluation and results are based on the difference in outcomes between those in the treatment group (who are to be invited to participate) and the control group within this 52-week period.

Figure 1: Flow diagram of the allocation of eligible people in the RCT



Less than 10% of the treatment group took part in the IWS service

The lower half of the flow chart in Figure 1 shows the subsequent experience of the two groups in participating in the IWS. CSRs only contacted 13% (866) of the treatment group. Of those contacted, 75% opted into the service. However, this means that of the treatment group (n=6,693) only 10% participated in the service.

The low contact rate and the resulting low participation rate into the IWS by the treatment group will be addressed in more detail in the findings section.

Under 1% of the control group took part in the IWS service

While the intention was that no one in the control group would participate in the IWS/IWP, in practice 40 people were identified as having participated, making up 0.5% of the control group overall.

Treatment and control group demographics were well balanced

We compared the demographic profiles of the treatment and control groups to check whether they were well balanced. If randomisation worked as intended, we expect to see no significant or substantial difference in the profiles at the point they were randomised into the control and treatment groups. We looked at both the overall treatment and control groups as well as a subset of those who had a benefit cancellation.

Overall, there was 1 percentage point or less difference between the treatment group and that of the control group profile. The exceptions are seen for people in Southern and Taranaki in the cancelled group, which differ by 1.1 percentage points. Additional demographics can be found in Appendix 2.

Findings

Here we look in detail at what happened to people assigned to the control and treatment groups in the 12 months after assignment. Before examining the impact of IWS on outcomes, we first need to understand why the IWS service achieved such a low contact rate and consequently a low opt-in rate.

There were several explanations for the low contact rate

As noted earlier, the low opt-in rate undermined the ability of the study to identify whether IWS made a difference to people's outcomes. For this reason, we examined in more detail why the contact rate was so low.

Key findings

- Many in the treatment group were not contacted to be invited to participate in the trial.
- Many of those who were not contacted had a suspended main benefit.
- Clients who had returned to a main benefit were also among those screened out.
- CSRs had trouble contacting some clients.
- Of the treatment group that were contacted, the opt-in rate was high.

Not contacted and could not contact were the primary reasons for the low contact rate

Figure 2 shows in more detail the stages between assignment to the treatment group and the opt-in stage as shown in Figure 1. The main reason for the low contact rate was the inability to contact potential participants (not yet contacted, n=2,424 and non-contact, n=1,867).

Non-contact was highest for people with suspended benefits

Feedback from CSRs was that they focused on recruiting people who had cancelled main benefit rather than those who had suspended benefits. This was a key factor since people with a suspended benefit made up nearly two-thirds (63%) of the treatment group (Table 1).

Table 1: Treatment and control group clients by cancelled and suspended benefits

	Treatment	Control
Cancelled	2,463	2,493
Suspended	4,230	4,241
Total	6,693	6,734

As shown in Figure 3 there was a marked difference in the contact rate between people who were suspended compared to those who had cancelled benefit. Over half of the treatment group with suspended benefits were "Not yet contacted", compared to only 10% of those with cancelled benefits. On the other hand, the proportion who agreed to

participate in the treatment group with a cancelled benefit was 20% compared to 3.7% for suspended.

Figure 2: Steps in contacting treatment group members to opt into IWS

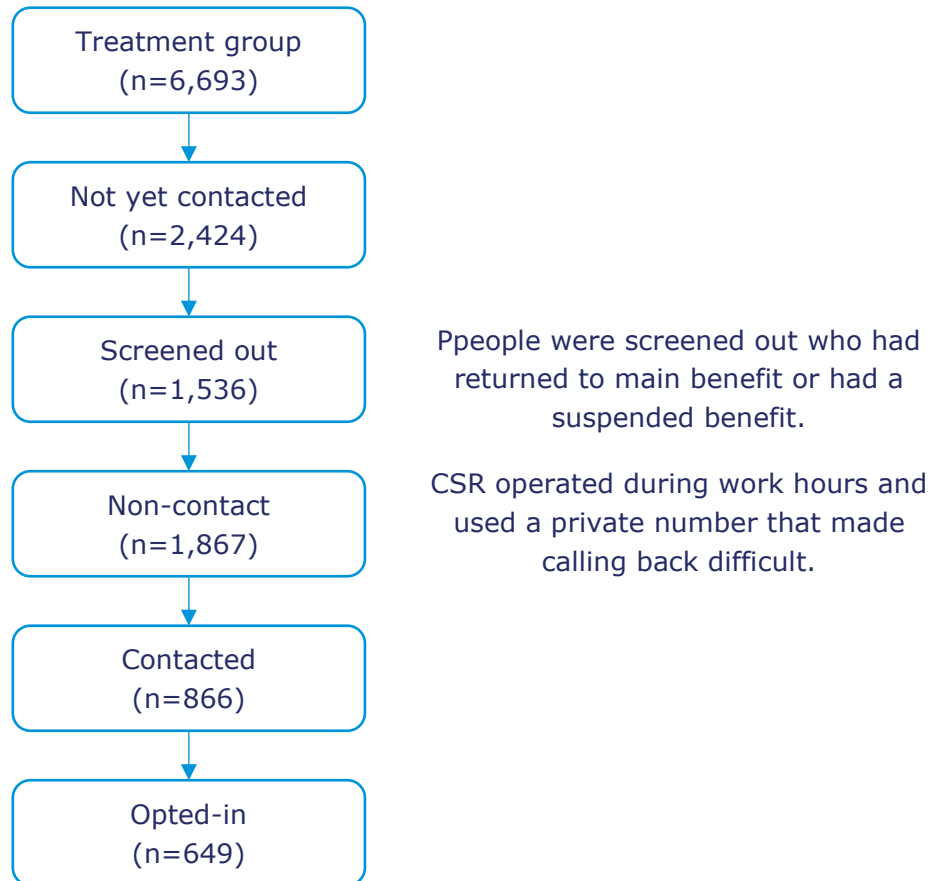
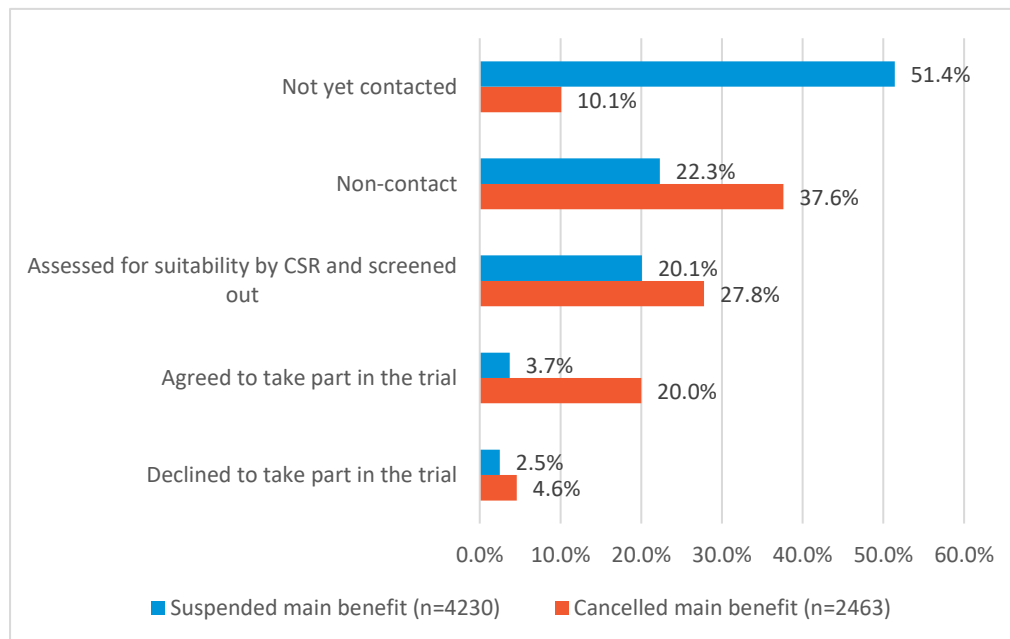


Figure 3: Participation status for people in the treatment group, by Suspended (n=4,230) and Cancelled (n=2,463) main benefits



From January 2018 people with suspended benefits were no longer eligible for the IWS trial. However, this change was too late to make a noticeable improvement to the contact rate, and the subsequent opt-in rate, before the trial ended in July 2019.⁵

People who had returned to main benefit were among those screened out

After “Not yet contacted” and “Non-contact”, Figure 3 shows the next largest group were people screened out by the CSRs as unsuitable (20.1% of suspended and 27.8% of cancelled main benefit clients). In many instances, these were people who had either already returned to main benefit or were applying for income support.

There was difficulty contacting some people

The CSRs were unable to contact a portion of eligible people (Non-Contact, 22.3% of suspended and 37.6% of cancelled main benefits). Initially, a minimum of three attempts were made, and then after conducting a text trial during July to August 2017, a text message was subsequently left with an 0800 number that people could call if they wished to opt-in.⁶ Difficulty contacting people who have left a benefit is recognised as a common issue across interventions.⁷

⁵ There were no operational changes to the service.

⁶ Malatest (2018), page 21.

⁷ Anecdotally, the CSRs operated for some time during work hours only, and initially the number used was a private number, so there may have been an aversion to answering an unknown caller. Some of these are detailed in the accompanying qualitative report by Malatest (2018).

Impact of IWS/IWP on off-benefit outcomes

Based on the observed difference in contact rate between suspended and cancelled clients, we chose to estimate the impact of the IWS/IWP on all treatments (all) and those who had their benefit cancelled at assignment (cancelled).

Being eligible for IWS/IWP resulted in a modest but non-significant increase in time off benefit

Table 2 shows the total time off main benefit in the 360 days after assignment to the control or treatment group for all and the subset of cancelled. Starting with the all group, the treatment group spent 263 days off main benefit after assignment compared to 260 days for the control group. Therefore, being eligible for the IWS/IWP resulted in an additional 3.2 days off main benefit. However, because the confidence interval includes zero, we cannot rule out the possibility that the service had no impact.

Table 2: Impact on off-main benefit outcomes at 360 days after randomisation for cancelled and all clients

	Days off main benefit	
	Cancelled	All
Treatment group	286.4	263.2
Control group	281.9	260.0
Impact (95% CI)	4.6 (-0.9, 10.1)	3.2 (-0.5, 6.9)

All: all people assigned to the control or treatment group.

Cancelled: people whose main benefit had been cancelled when assigned to the control or treatment group.

The impact of being eligible for the IWS/IWP for people who had cancelled benefit was higher, at 4.6 days. But again, we cannot rule out the possibility that the service had no impact. Reassuringly, the impact estimate among the higher opt-in group (cancelled) was higher than for the overall group.

The impact of IWS appears largest between three and nine months after assignment

Figure 4 shows the average number of days on a benefit in each 30-day period before and after assignment to the control or treatment group for those who had cancelled benefit. The zero point is the date of assignment to the control or treatment group and follows each group from 180 days before assignment to 360 days afterwards. Each data point measures the average number of days each group spent on main benefit over 30-day intervals.

Figure 4: Average number of days on main benefit per client, for each 30-day interval since randomisation, for clients with cancelled benefits

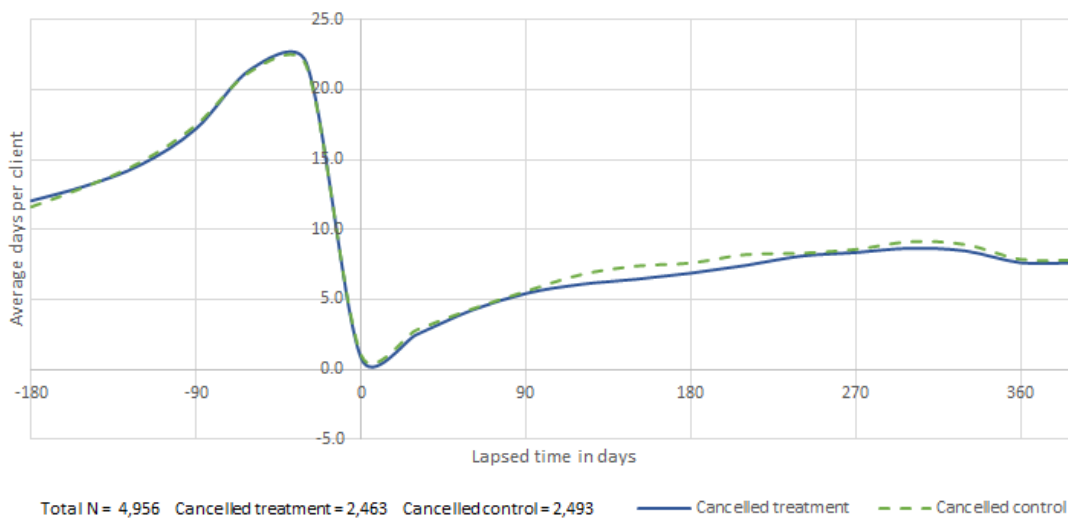
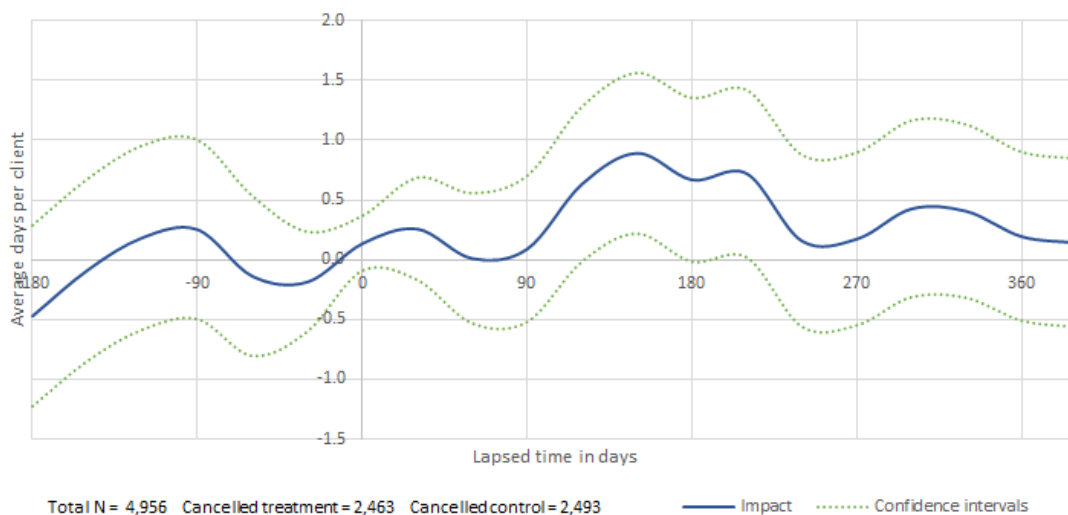


Figure 5 is based on Figure 4 and shows the difference between the treatment and control groups' days on main benefit for each 30-day interval. The blue line is the mid-point estimate while the green dashed lines are the 95% confidence intervals. The main observations from Figure 5 are:

- that the difference in off-benefit outcomes is higher for the treatment than control group indicating a positive impact, but for most intervals, the difference was not statistically significant
- the impact of the IWS/IWP appeared to be larger between three and nine months after assignment.

Figure 5: Interval impact on average days off main benefit for the cancelled group



Regression adjustments made no difference to the results

We also tested how sensitive the observed impact estimates were to the inclusion of other factors that might influence the outcomes of people in the treatment and control groups.⁸

Overall, the regression adjustment made no substantive difference to the estimated impact compared to using a difference in means values as reported in Table 2 in Appendix 3. This result indicates that randomisation was successful in ensuring the treatment and control groups were well balanced on observable characteristics and, as a result, their inclusion had no material bearing on the impact estimate.

The results of the regression analysis can be found in Appendix 3.

⁸ The assumption made here is that RCT data, which has been regressed on one or more control variables, is a form of instrumental variables.

Conclusions

The study was unable to confirm whether IWS and IWP was effective

This study was unable to confirm whether the IWS was effective on reducing the time people spent on main benefit. The inability to determine effectiveness was because the study's design relied on a high participation rate, but in practice the participation rate was only 10%.

Several operational issues hampered the impact evaluation.

- Staff focused on contacting those who had cancelled rather than suspended benefits resulting in a low overall contact rate. This reduced the statistical power of the study to identify the impact of IWS on the outcomes of those eligible for the service.
- Because eligible clients were contacted via a private number, if they missed the call, they could not make a return call.
- Eligible clients were only contacted after they had exited from a main benefit and were in employment, making contact more difficult than if they had been contacted before leaving benefit.
- A small proportion of clients assigned to the control group received the IWS service.

During the trial, improvements were made to the design to increase contact and subsequent opt-in rates (Malatest, 2018). However, these changes were too late to make a material difference to the impact evaluation.

People were willing to take part in the IWS and receive the IWP

An important finding from the study is that of those who could be contacted, 75% agreed to take part. Such a result shows that if the contact rate was improved, we could expect a high participation rate among those eligible for the service. Under these conditions, we are confident we could reliably determine whether the service was effective.

Lessons learnt

The IWS trial provided some important lessons on how to trial and evaluate a new service.

More time on reviewing the literature and service design

It is important to check the international literature on evidence and identify features that make the proposed service successful. For example, the IWS service was unusual in that it tried to recruit people after they had left benefit. Generally, in-work support and related services recruit people before they exit main benefit. In so doing, they would have more success in recruiting participants but also increase the effectiveness of the programme by including the incentive of an IWP on job search intensity.

Implementation evaluation

Once the design phase has been completed it is critical to undertake an implementation evaluation focused on ensuring the programme or service is operating as intended. Furthermore, there needs to be the governance and commitment to make the necessary changes to ensure the intervention is working well.

Evaluate effectiveness once the intervention is working well

Impact evaluation, especially RCTs, should only be considered after the implementation phase has been completed successfully. Also, RCT requires careful thought in terms of its design and implementation, taking into consideration things such as:

- whether the design and size of the RCT will provide sufficient power to identify meaningful impacts
- if the method be justified on ethical grounds
- agreement by all parties to adhere to the design protocol
- close monitoring of the RCT during its establishment phase to avoid errors that are difficult to fix at the analysis stage.

Appendix 1: Recruitment into IWS process

All clients in the table below were considered as part of the trial and were included in the quantitative analysis.

Table 3: Examples of how the recruitment process was followed by the CSRs and the resulting client status

Process followed by CSRs	Status
The eligible client was taken from the call list by the CSR, assessed as suitable, successfully contacted, invited to the trial, and client opted-in to receiving the service.	Opted-in
As above, but the client opted-out of receiving the service.	Declined
The client was on the call list but was assessed as unsuitable. ⁹	Screened out and not contacted
The client was on the call list and assessed as suitable but was not able to be contacted. ¹⁰ Additionally, some clients were contacted before the trial period started.	Non-contact
The client was on the call list but there was no record that the client was invited to participate in the trial.	Not yet contacted

⁹ One of the reasons that CSRs assessed clients as unsuitable for the trial was if they had returned to a main benefit since being randomised to the treatment group.

¹⁰ CSRs were able to contact clients three to six times.

Appendix 2: Participant demographics

Table 4: Numbers and characteristics of the population that were part of the trial, by treatment and control for cancelled and all clients

	Cancelled		All	
	Treatment	Control	Treatment	Control
Age				
20-<25 years	255 (10.4)	241 (9.7)	546 (8.2)	503 (7.5)
25-<30 years	705 (28.6)	726 (29.1)	1,807 (27.0)	1,813 (26.9)
30-<35 years	360 (14.6)	394 (15.8)	919 (13.7)	943 (14.0)
35-<40 years	237 (9.6)	242 (9.7)	653 (9.8)	628 (9.3)
40-<45 years	184 (7.5)	172 (6.9)	489 (7.3)	483 (7.2)
45-<50 years	185 (7.5)	197 (7.9)	501 (7.5)	574 (8.5)
50-<55 years	181 (7.3)	201 (8.1)	588 (8.8)	604 (9.0)
55-<60 years	169 (6.9)	157 (6.3)	521 (7.8)	544 (8.1)
60-<65 years	128 (5.2)	111 (4.5)	443 (6.6)	422 (6.3)
Over 65 years	59 (2.4)	52 (2.1)	226 (3.4)	220 (3.3)
Mean (std dev)	38.0 (12.7)	37.7 (12.3)	39.6 (13.2)	39.7 (13.1)
Gender (%)				
Male	1,677 (68.1)	1,714 (68.8)	4,356 (65.1)	4,418 (65.6)
Female	785 (31.9)	779 (31.2)	2,335 (34.9)	2,315 (34.4)
Gender diverse	1 (0.0)	0 (0.0)	2 (0.0)	1 (0.0)
Ethnicity (%)				
Māori	1,030 (41.8)	1,062 (42.6)	2,684 (40.1)	2,723 (40.4)
NZ European	832 (33.8)	829 (33.3)	2,316 (34.6)	2,340 (34.7)
Other	359 (14.6)	375 (15.0)	1,022 (15.3)	1,024 (15.2)
Pacific People	229 (9.3)	220 (8.8)	637 (9.5)	612 (9.1)
Unspecified	13 (0.5)	7 (0.3)	34 (0.5)	35 (0.5)
Location (%)				
Auckland Metro	593 (24.1)	598 (24.0)	1,711 (25.6)	1,670 (24.8)
Bay of Plenty	329 (13.4)	346 (13.9)	764 (11.4)	789 (11.7)
Canterbury	138 (5.6)	151 (6.1)	397 (5.9)	422 (6.3)
Central	141 (5.7)	139 (5.6)	390 (5.8)	369 (5.5)
East Coast	237 (9.6)	224 (9.0)	697 (10.4)	687 (10.2)
Nelson	103 (4.2)	120 (4.8)	347 (5.2)	363 (5.4)
Northland	87 (3.5)	75 (3.0)	272 (4.1)	296 (4.4)
Other	27 (1.1)	17 (0.7)	33 (0.5)	24 (0.4)
Southern	261 (10.6)	291 (11.7)	777 (11.6)	822 (12.2)
Taranaki	104 (4.2)	77 (3.1)	296 (4.4)	276 (4.1)
Waikato	277 (11.2)	305 (12.2)	571 (8.5)	576 (8.6)
Wellington	166 (6.7)	150 (6.0)	438 (6.5)	440 (6.5)
Age at first benefit				
15-<16 years	2 (0.1)	2 (0.1)	10 (0.1)	7 (0.1)
16-<18 years	475 (19.3)	468 (18.8)	1,161 (17.3)	1,130 (16.8)

	Cancelled		All	
	Treatment	Control	Treatment	Control
18-<20 years	925 (37.6)	987 (39.6)	2,374 (35.5)	2,378 (35.3)
20-<25 years	372 (15.1)	350 (14.0)	940 (14.0)	970 (14.4)
25-<30 years	195 (7.9)	231 (9.3)	563 (8.4)	620 (9.2)
30-<35 years	165 (6.7)	150 (6.0)	510 (7.6)	508 (7.5)
35-<40 years	140 (5.7)	129 (5.2)	489 (7.3)	450 (6.7)
40-<45 years	106 (4.3)	93 (3.7)	350 (5.2)	349 (5.2)
45-<50 years	30 (1.2)	47 (1.9)	122 (1.8)	152 (2.3)
50-<55 years	31 (1.3)	20 (0.8)	91 (1.4)	98 (1.5)
55-<60 years	19 (0.8)	12 (0.5)	65 (1.0)	57 (0.8)
60-<65 years	3 (0.1)	4 (0.2)	18 (0.3)	14 (0.2)
Over 65 years	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.0)
Type of first benefit				
Caring Sick Infirm	7 (0.3)	5 (0.2)	21 (0.3)	14 (0.2)
Emergency	37 (1.5)	45 (1.8)	100 (1.5)	01.8)
Emergency Maintenance	6 (0.2)	6 (0.2)	19 (0.3)	00.4)
Invalids	27 (1.1)	26 (1.0)	93 (1.4)	01.2)
Jobseeker Support	1,125 (45.7)	1,206 (48.4)	3,192 (47.7)	3,350 (49.7)
Retired	0 (0.0)	0 (0.0)	2 (0.0)	3 (0.0)
Sickness	177 (7.2)	131 (5.3)	447 (6.7)	372 (5.5)
Sole Parent	141 (5.7)	143 (5.7)	419 (6.3)	446 (6.6)
Student	233 (9.5)	219 (8.8)	613 (9.2)	552 (8.2)
Supplementary Only	404 (16.4)	441 (17.7)	1,036 (15.5)	1,082 (16.1)
Widow	3 (0.1)	0 (0.0)	11 (0.2)	4 (0.1)
Women Alone	2 (0.1)	2 (0.1)	4 (0.1)	7 (0.1)
Youth	301 (12.2)	269 (10.8)	736 (11.0)	678 (10.1)
Highest education				
None	217 (8.8)	218 (8.7)	607 (9.1)	634 (9.4)
NQF Level 1, <80 credits	140 (5.7)	158 (6.3)	381 (5.7)	403 (6.0)
NQF Level 1, 80+ credits	300 (12.2)	285 (11.4)	784 (11.7)	775 (11.5)
NQF Level 2	244 (9.9)	248 (9.9)	683 (10.2)	668 (9.9)
NQF Level 3	154 (6.3)	164 (6.6)	400 (6.0)	404 (6.0)
NQF Level 4	1,006 (40.8)	1,001 (40.2)	2,672 (39.9)	2,688 (39.9)
Degree Professional	314 (12.7)	329 (13.2)	935 (14.0)	922 (13.7)
Unknown	88 (3.6)	90 (3.6)	231 (3.5)	240 (3.6)
Total (%)	2,463 (49.7)	2,493 (50.3)	6,693 (49.8)	6,734 (50.2)

Note: Demographic figures are a proportion of the column total. Total figures are proportions of the “cancelled” population of 4,956 and the “all” population of 13,427.

Table 5: Clients by Work and Income office location for cancelled and all clients

	Cancelled		All	
	Treatment	Control	Treatment	Control
Auckland Metro	593 (24.1)	598 (24.0)	1,711 (25.6)	1,670 (24.8)
Auckland City	174 (7.1)	160 (6.4)	573 (8.6)	499 (7.4)
Franklin	22 (0.9)	28 (1.1)	61 (0.9)	74 (1.1)
Manukau City	138 (5.6)	129 (5.2)	454 (6.8)	412 (6.1)
North Shore City	67 (2.7)	67 (2.7)	167 (2.5)	172 (2.6)
Papakura	42 (1.7)	59 (2.4)	98 (1.5)	125 (1.9)
Rodney	30 (1.2)	36 (1.4)	73 (1.1)	84 (1.2)
Waitakere City	120 (4.9)	119 (4.8)	285 (4.3)	304 (4.5)
Bay of Plenty	329 (13.4)	346 (13.9)	764 (11.4)	789 (11.7)
Kawerau	15 (0.6)	16 (0.6)	45 (0.7)	39 (0.6)
Opotiki	5 (0.2)	8 (0.3)	32 (0.5)	25 (0.4)
Rotorua	72 (2.9)	64 (2.6)	142 (2.1)	138 (2.0)
South Waikato	72 (2.9)	70 (2.8)	125 (1.9)	126 (1.9)
Taupo	23 (0.9)	22 (0.9)	59 (0.9)	52 (0.8)
Tauranga	94 (3.8)	112 (4.5)	244 (3.6)	272 (4.0)
Western Bay of Plenty	27 (1.1)	26 (1.0)	50 (0.7)	51 (0.8)
Whakatane	21 (0.9)	28 (1.1)	67 (1.0)	86 (1.3)
Canterbury	138 (5.6)	151 (6.1)	397 (5.9)	422 (6.3)
Ashburton	8 (0.3)	12 (0.5)	33 (0.5)	33 (0.5)
Christchurch City	118 (4.8)	126 (5.1)	332 (5.0)	356 (5.3)
Waimakariri	12 (0.5)	13 (0.5)	32 (0.5)	33 (0.5)
Central	141 (5.7)	139 (5.6)	390 (5.8)	369 (5.5)
Horowhenua	26 (1.1)	23 (0.9)	74 (1.1)	56 (0.8)
Kāpiti Coast	22 (0.9)	24 (1.0)	60 (0.9)	61 (0.9)
Manawatu	9 (0.4)	9 (0.4)	28 (0.4)	26 (0.4)
Masterton	28 (1.1)	24 (1.0)	56 (0.8)	63 (0.9)
Palmerston North City	41 (1.7)	46 (1.8)	138 (2.1)	135 (2.0)
Taranua	15 (0.6)	13 (0.5)	34 (0.5)	28 (0.4)
East Coast	237 (9.6)	224 (9.0)	697 (10.4)	687 (10.2)
Central Hawkes Bay	27 (1.1)	25 (1.0)	58 (0.9)	50 (0.7)
Gisborne	68 (2.8)	71 (2.8)	177 (2.6)	171 (2.5)
Hastings	76 (3.1)	67 (2.7)	294 (4.4)	272 (4.0)
Napier City	57 (2.3)	49 (2.0)	130 (1.9)	150 (2.2)
Wairoa	9 (0.4)	12 (0.5)	38 (0.6)	44 (0.7)
Nelson	103 (4.2)	120 (4.8)	347 (5.2)	363 (5.4)
Buller	13 (0.5)	11 (0.4)	22 (0.3)	25 (0.4)
Grey	15 (0.6)	12 (0.5)	45 (0.7)	43 (0.6)
Marlborough	28 (1.1)	36 (1.4)	76 (1.1)	89 (1.3)
Nelson City	16 (0.6)	21 (0.8)	82 (1.2)	80 (1.2)
Tasman	31 (1.3)	40 (1.6)	122 (1.8)	126 (1.9)

	Cancelled		All	
	Treatment	Control	Treatment	Control
Northland	87 (3.5)	75 (3.0)	272 (4.1)	296 (4.4)
Far North	47 (1.9)	42 (1.7)	131 (2.0)	142 (2.1)
Kaipara	3 (0.1)	4 (0.2)	13 (0.2)	18 (0.3)
Whangarei	37 (1.5)	29 (1.2)	128 (1.9)	136 (2.0)
Southern	261 (10.6)	291 (11.7)	777 (11.6)	822 (12.2)
Central Otago	7 (0.3)	3 (0.1)	16 (0.2)	14 (0.2)
Clutha	8 (0.3)	8 (0.3)	42 (0.6)	29 (0.4)
Dunedin City	64 (2.6)	78 (3.1)	185 (2.8)	204 (3.0)
Gore	15 (0.6)	14 (0.6)	39 (0.6)	46 (0.7)
Invercargill City	108 (4.4)	112 (4.5)	311 (4.6)	281 (4.2)
Queenstown Lakes	0 (0.0)	5 (0.2)	2 (0.0)	11 (0.2)
Timaru	35 (1.4)	45 (1.8)	133 (2.0)	174 (2.6)
Waitaki	24 (1.0)	26 (1.0)	49 (0.7)	63 (0.9)
Taranaki	104 (4.2)	77 (3.1)	296 (4.4)	276 (4.1)
New Plymouth	34 (1.4)	15 (0.6)	98 (1.5)	76 (1.1)
Rangitikei	12 (0.5)	7 (0.3)	28 (0.4)	23 (0.3)
Ruapehu	6 (0.2)	2 (0.1)	13 (0.2)	9 (0.1)
South Taranaki	9 (0.4)	9 (0.4)	17 (0.3)	18 (0.3)
Stratford	6 (0.2)	10 (0.4)	21 (0.3)	25 (0.4)
Waitomo	8 (0.3)	7 (0.3)	17 (0.3)	20 (0.3)
Whanganui	29 (1.2)	27 (1.1)	102 (1.5)	105 (1.6)
Waikato	277 (11.2)	305 (12.2)	571 (8.5)	576 (8.6)
Hamilton City	171 (6.9)	180 (7.2)	369 (5.5)	365 (5.4)
Hauraki	19 (0.8)	32 (1.3)	40 (0.6)	47 (0.7)
Matamata-Piako	21 (0.9)	15 (0.6)	36 (0.5)	25 (0.4)
Thames-Coromandel	6 (0.2)	11 (0.4)	19 (0.3)	22 (0.3)
Waikato	37 (1.5)	42 (1.7)	63 (0.9)	66 (1.0)
Waipa	23 (0.9)	25 (1.0)	44 (0.7)	51 (0.8)
Wellington	166 (6.7)	150 (6.0)	438 (6.5)	440 (6.5)
Hutt City	62 (2.5)	48 (1.9)	141 (2.1)	147 (2.2)
Porirua City	28 (1.1)	40 (1.6)	71 (1.1)	84 (1.2)
Upper Hutt City	16 (0.6)	10 (0.4)	41 (0.6)	41 (0.6)
Wellington City	60 (2.4)	52 (2.1)	185 (2.8)	168 (2.5)
Other	27 (1.1)	17 (0.7)	33 (0.5)	24 (0.4)
Other	27 (1.1)	17 (0.7)	33 (0.5)	24 (0.4)

Appendix 3: A predicted treatment variable and covariates for a regression analysis

The following regression analysis was conducted to address concerns that sample selection issues and randomisation resulted in meaningful differences between the treatment and control groups. To determine whether a service is effective through an RCT, treatment and control groups need to be similar across all observable characteristics except for their participation in the service.

Methods for producing the impact outcomes were reworked and used to focus on the lapse period at 360 days for each client. This produced a dataset that had the individual client benefit durations leading up to and at 360 days by group (treatment or control). A frequency distribution of *days on a main benefit* for the first 360 days was produced, to see whether any transformations were needed before applying the regression model. Note that this part of the analysis looked at *days on a main benefit*.

The following tables were produced.

Table 6: Summary of days on main benefit at 360 days, by treatment and control groups, for cancelled and all clients

	Cancelled main benefit		All (cancelled and suspended)	
	Treatment	Control	Treatment	Control
Number of clients	2,463	2,493	6,693	6,734
Sum of days	181,147	194,770	647,576	673,227
Mean days	73.6	78.1	96.8	100
Variance in days	9,589.5	9,816.7	11,891.6	11,991.2
Std dev	97.9	99.1	109.1	109.5
Mean difference	4.5		3.2	

Figure 6 and Figure 7 describe the frequency distribution of days on main benefit, in spells of 30 days.

Figure 6: Cancelled treatment and control group days on main benefit at 360 days

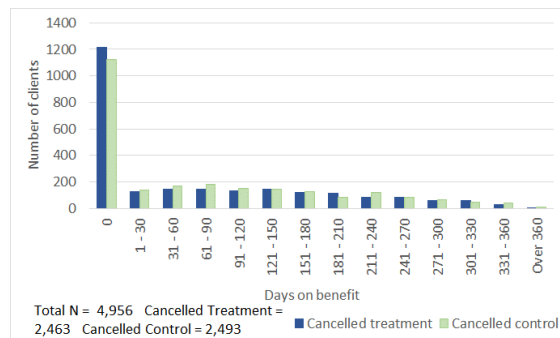


Figure 7: All treatment and control group days on main benefit at 360 days

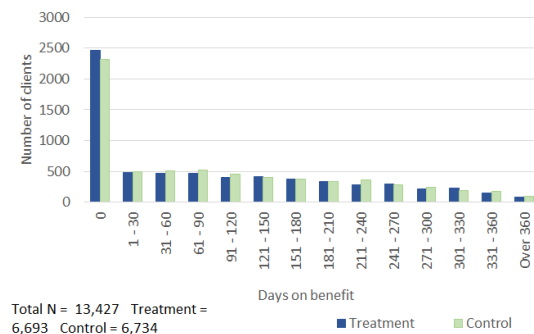


Figure 6 and Figure 7 reflect a concentration of clients around the zero days value. These are clients that have remained off benefit for the full year. Based on these resulting figures, an ordinary least squares (OLS) linear regression was modelled.

The notation for this was:

$$Y = F(\alpha + \beta_n X_n + \gamma T + \epsilon)$$

Where:

Y = the estimated number of days spent on main benefit in the 360 days after assignment

α is the intercept term

$\beta_n X_n$ refers to the other covariates or covariates in the model, which are Benefit duration¹¹, Age, Benefit type, and Location¹².

γT is the coefficient for membership of the treatment or control groups,

ϵ is the residual (which is assumed to be normally distributed with mean zero), and the OLS function is $F(\cdot)$.

The idea is that the other profile variables explain the variance in the outcome, and the γT (the expected difference in Y for being in treatment or control) has a lower standard error, with the expectation it will be significant.

For this regression it seemed intuitive to use on-main benefit outcomes.

¹¹ The current duration on a benefit at the time of randomisation.

¹² For the class variables (those variables that were not continuous) of Treatment or Control, "Control" was used as the reference value; of Benefit type, "Jobseeker Support" was used as the reference value, and; of Location, "Auckland Metro" was used as the reference value.

Table 7: Covariates are Benefit duration, Age, Benefit type, and Location, taken as at the day before randomisation to treatment or control

	Cancelled main benefit		All (cancelled and suspended)	
	Treatment	Control	Treatment	Control
Mean benefit duration (std dev) ¹³	1434.2 (1187.4)	1427.5 (1167.3)	1503.5 (1230.2)	1491.0 (1229.3)
Mean age (std dev)	34.6 (12.6)	34.3 (12.2)	35.9 (13.2)	36.1 (13.0)

Benefit Type (%)	Cancelled main benefit		All (cancelled and suspended)	
	Treatment	Control	Treatment	Control
Jobseeker Support	2 (0.1)	(0.0)	4232 (63.2)	4239 (62.9)
Off Benefit	2348 (95.3)	2375 (95.3)	2348 (35.1)	2375 (35.3)
Retired	1 (0.0)	1 (0.0)	1 (0.0)	1 (0.0)
Sickness	(0.0)	(0.0)	(0.0)	2 (0.0)
Supplementary Only	112 (4.5)	117 (4.7)	112 (1.7)	117 (1.7)
Total	2463 (49.7)	2493 (50.3)	774 (48.5)	821 (51.5)

Location	Cancelled main benefit		All (cancelled and suspended)	
	Treatment	Control	Treatment	Control
Auckland Metro	594 (8.8)	594 (8.8)	1710 (25.5)	1666 (24.7)
Bay of Plenty	333 (4.9)	354 (5.3)	775 (11.6)	795 (11.8)
Canterbury	138 (2.0)	150 (2.2)	400 (6.0)	422 (6.3)
Central	138 (2.0)	137 (2.0)	381 (5.7)	367 (5.4)
East Coast	237 (3.5)	219 (3.3)	701 (10.5)	682 (10.1)
Nelson	104 (1.5)	118 (1.8)	346 (5.2)	362 (5.4)
Northland	87 (1.3)	74 (1.1)	271 (4.0)	294 (4.4)
Other	27 (0.4)	19 (0.3)	33 (0.5)	26 (0.4)
Southern	258 (3.8)	292 (4.3)	774 (11.6)	821 (12.2)
Taranaki	103 (1.5)	80 (1.2)	296 (4.4)	281 (4.2)
Waikato	276 (4.1)	305 (4.5)	565 (8.4)	573 (8.5)
Wellington	168 (2.5)	151 (2.2)	441 (6.6)	445 (6.6)

¹³ Mean benefit duration is the average of the sum of Jobseeker Support benefit duration and Sickness benefit duration.

Total	2463 (36.6)	2493 (37.0)	6693 (100.0)	6734 (100.0)
--------------	--------------------	--------------------	---------------------	---------------------

Note: Demographic figures are a proportion of the column total. Total figures are proportions of the treatment and control population of 13,427.

Regression analysis verified the randomisation to treatment and control groups

Regression analysis was carried out for both the all (cancelled and suspended) treatment and control groups, and the cancelled only treatment and control groups. For each analysis, this consisted of trying to fit the treatment members and controls in to both the model without covariates, and then the model with covariates.

The first half of Table 8 (the treatment and control groups) shows that the overall F statistic is insignificant for the model without covariates. The mean difference (impact) for this regression adjusted estimate between the treatment and control groups is -3.2 days with a 95% probability that the true impact was between -6.9 days and 0.5 days. This means that those in the treatment group spent on average 3.2 less days on a main benefit than those in the control group, which is the same result as the simple comparison of mean impact analysis we did for the treatment and control groups. However, statistically, this is not significant.

The second half of Table 8 shows that the overall F statistic is significant for the model with control variables. Interestingly, it shows that adding covariates barely made a difference to the estimate, which reflects that the treatment and control groups have closely balanced demographics. However, the mean difference is not statistically significant.

Table 8: Treatment and control groups main benefit duration comparisons after regression

Treatment (6693) and Control (6734)	Mean days (std err)	LCL - UCL	t Value (Probability > t)
Model: Without control variables, F Value is 2.91, Probability > F is 0.0878			
Treatment	96.8 (1.3)	94.1 – 99.4	-1.71 (0.0878)
Control	100 (1.3)	97.4 – 102.6	75.1 (< .0001)
Difference/impact	-3.2 (1.9)	-6.9 – 0.5	-1.71 (0.0878)
Model: With control variables, overall F Value is 45.94, overall Probability > F is <0.0001			
Treatment	96.7 (1.3)	94.2 – 99.3	-1.81 (0.0708)
Control	100 (1.3)	97.5 – 102.6	23.2 (< .0001)
Difference/Impact	-3.3 (1.8)	-6.9 – 0.3	-1.81 (0.0708)

Attempting to fit the model shows from Figure 8 and Figure 9 that the residuals (the error terms) of the dependent variable without covariates are not normally distributed.¹⁴ With control variables added, in Figure 9 the residuals of the dependent variable are slightly less skewed. However, the sum and the mean of the residuals in both figures is almost zero. Therefore, despite the residuals not being normally distributed, this partially satisfies the conditions to validate this model.

Figure 8: Treatment and control groups residuals without covariates

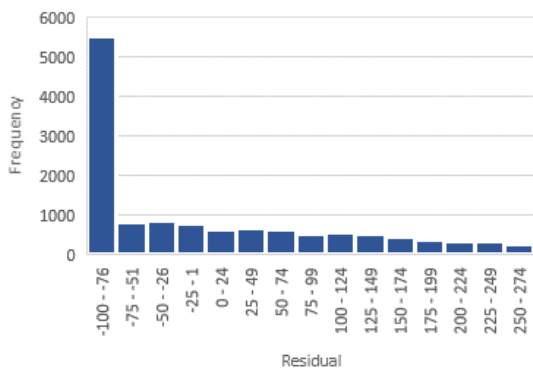
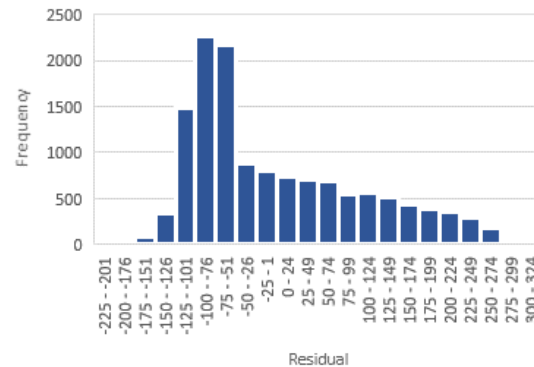


Figure 9: Treatment and control groups residuals with covariates



The first half of Table 9 (the cancelled treatment and control groups) shows that the overall F statistic is insignificant for the model without control variables. The mean difference (impact) between the cancelled treatment and cancelled control is -4.6 days with a 95% probability that the true impact was between -10.1 days and 0.9 days. This means that those in the cancelled treatment group spent on average 4.6 less days on main benefit than those in the cancelled control group, which is the same result as the simple comparison of mean impact analysis we did for the cancelled treatment and cancelled control groups. Statistically this too is not significant.

The second half of Table 9 shows that the overall F statistic is significant for the model with control variables. Interestingly, it shows that adding covariates barely made a difference to the estimate, which reflects that the cancelled treatment and cancelled control groups have closely balanced demographics. However, the mean difference is not detectable.

Table 9: Cancelled treatment and cancelled control groups main benefit duration comparisons after regression

Cancelled treatment (2463) and Cancelled control (2493)	Mean days (std err)	LCL - UCL	t Value (Probability > t)
Model: Without control variables, F value is 2.68, Probability > F is 0.1018			

¹⁴ In regression testing, if residuals are normally distributed the assumptions and inferences from the model are valid.

Cancelled treatment	73.5 (2.0)	69.7 – 77.4	-1.64 (0.1018)
Cancelled control	78.1 (2.0)	74.3 – 82.0	39.60 (< .0001)
Difference/impact	-4.6 (2.8)	-10.1 – 0.9	-1.64 (0.0878)
Model: With control variables, overall F value is 13.44, overall Probability > F is <0.0001			
Cancelled treatment	73.7 (1.9)	70.0 – 77.5	-1.56 (0.1191)
Cancelled control	78.0 (1.9)	74.2 – 81.8	-0.04 (0.9648)
Difference/impact	-4.3 (2.7)	-9.7 – 1.1	-1.56 (0.1191)

Attempting to fit the model shows from Figure 10 and Figure 11 that the residuals (the error terms) of the dependent variable without covariates are not normally distributed. Additional control variables made minor difference. However, the sum and the mean of the residuals in both figures is almost zero, partially satisfying the conditions to validate the analysis.

Figure 10: Cancelled treatment and cancelled control groups residuals without covariates

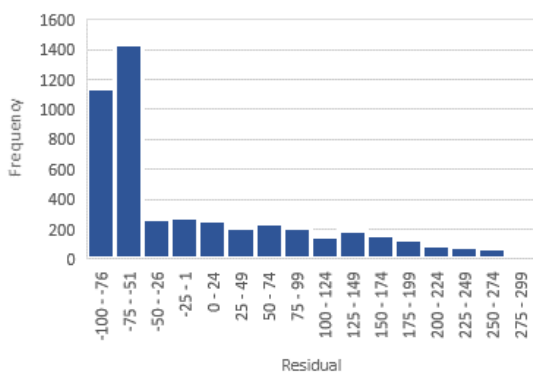
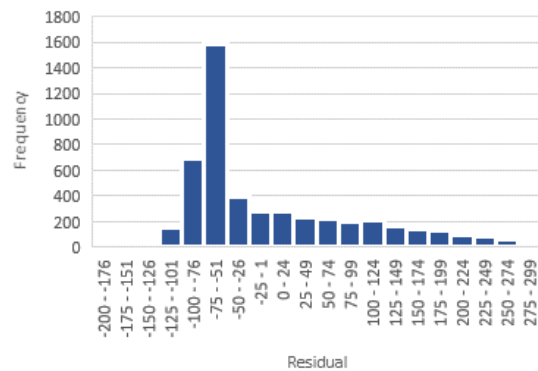


Figure 11: Cancelled treatment and cancelled control groups residuals with covariates



References

- Meyer, B. (1988). "Implications of the Illinois Reemployment Bonus Experiments For Theories of Unemployment and Policy Design," NBER Working Papers 2783, National Bureau of Economic Research, Inc.
- Cebulla, A., Flore, G., & Greenberg, D. (2008) *The New Deal for Lone Parents, Lone Parent Work Focused Interviews and Working Families' Tax Credit: a review of impacts*, DWP Research Report 484, https://www.researchgate.net/publication/235359885_The_New_Deal_for_Lone_Parents_Lone_Parent_Work_Focused_Interviews_and_Working_Families'_Tax_Credit_A_review_of_impacts.
- Dolton, P., & Smith, J. (2011), *The Impact of the UK New Deal for Lone Parents on Benefit Receipt*, Discussion Paper Series, IZA DPO No. 5491, February 2011, from <http://ftp.iza.org/dp5491.pdf>.
- Malatest (2018), *In Work Support and Payment Trial: Evaluation*. Ministry of Social Development, Wellington.
- Michalopoulos et al (2002), *Making Work Pay: Final Report on the Self-Sufficiency Project for Long-Term Welfare Recipients*. Social Research and Demonstration Corporation (SRDC), Ottawa.
- Taylor Fry. (2013). *The Treasury New Zealand Actuarial valuation of the Benefit System for Working-age Adults as at 30 June 2013*. Retrieved March 19, 2018 from <https://www.msd.govt.nz/documents/about-msd-and-our-work/publications-resources/evaluation/valuation-reports/valuation-benefit-system-working-age-adults-30-june-2012.pdf>.
- Taylor Fry (2017), *Off-benefit transitions: Where do people go?* Table 5-1 Proportion of people in each activity category who have been in that category for the last 3 months, p 30. Social Policy Evaluation and Research Unit, Wellington. Retrieved August 08, 2018 from <https://thehub.swa.govt.nz/assets/Uploads/REPORT-Off-benefit-Transitions.pdf>
- Torgerson D J, Roland M. *What is Zelen's design?* BMJ 1998; 316 :606
doi:10.1136/bmj.316.7131.606