Using IDI Data to Estimate Fiscal Impacts of Better Social Sector Performance

Analytical Paper 16/04

November 2016

New Zealand Government

ANALYTICAL PAPER 16/04

Using IDI Data to Estimate Fiscal Impacts of Better Social Sector Performance

MONTH/YEAR

November 2016

AUTHOR

Robert Templeton

ISBN (ONLINE)

978-0-947519-43-8

URL

Treasury website at November 2016:
http://www.treasury.govt.nz/publications/research-policy/ap/2016/16-04

NZ TREASURY

New Zealand Treasury
PO Box 3724
Wellington 6008
NEW ZEALAND

Email information@treasury.govt.nz

Telephone 64-4-472 2733

Website www.treasury.govt.nz

PURPOSE OF THE ANALYTICAL PAPERS SERIES

The Treasury's aim in publishing the Analytical Papers series is to make this analysis available to a wider audience and to inform and encourage public debate, with the ultimate aim of informing our policy advice.

Analytical Papers are commissioned as part of the Treasury's core function in developing and providing advice to Ministers. They include work undertaken by staff at the Treasury or other government departments, as well as work undertaken for the Treasury by external researchers or consultants.

Analytical Papers do not themselves represent policy advice.

Executive summary

The note uses data from the IDI (Integrated Data Infrastructure at Statistics New Zealand) to provide measures of potential fiscal impacts of four aspirational social investment scenarios that are outlined in Burton (2016). It is supplementary to a suite of background papers that underpin the Treasury's updated Long Term Fiscal Statement. The note also provides a descriptive picture of potential non-fiscal outcomes relating to the same scenarios.

The fiscal impacts from this analysis are expressed as cost-ratio parameters used in the calculations of the long term fiscal trajectories related to each scenario. This note aims to provide detail of the methods used in calculating these ratios. The descriptions of the scenarios and the resulting estimated fiscal trajectories can be found in Burton (2016).

This work is a first attempt aimed at giving some sense of how IDI data could be used in this sort of fiscal analysis. It is expected that as social investment initiatives come closer to implementation, with details more clearly specified, that these methods and calculations would be re-visited.

Contents

List of figures

Abstract

The analysis was undertaken as part of the work to update Treasury's long term fiscal modelling.

As part of this a scenario analysis of the potential fiscal and non-fiscal benefits of social investment was undertaken and is described in 'The benefits of improved social sector performance', a detailed background paper to *He Tirohanga Mokopuna*, Treasury's long-term fiscal statement (LTFS).

The paper provides information about how we used data from SNZ's Integrated Data infrastructure (IDI) to inform the scenario analysis.

The paper also provides a descriptive picture of potential non-fiscal outcomes relating to the same scenarios. The results of this work have been incorporated into *He Tirohanga Mokopuna* and the background papers but this report is being made available for those interested in more of the details behind some of the scenario analysis.

JEL CLASSIFICATION

C55 Mathematical and Quantitative Methods – Econometric Modelling – Large Data Sets: Modelling and Analysis

I38 Health, Education and Welfare – Welfare and Poverty – Government Policy: Provision and Effects of Welfare Programs

1 Introduction

Greater use of population data in the IDI (Integrated Data Infrastructure held at Statistics New Zealand) supports analysis that enables the public sector to identify sub-groups of the population who are at risk of poorer education, welfare, health and corrections related outcomes. Treasury's analytical paper 16/01 describes in detail the creation of the datasets that underpin the work described in this note:

Characteristics of Children at Greater Risk of Poor Outcomes as Adults
(http://www.treasury.govt.nz/publications/research-policy/ap/2016/16-01)

In this note we build on that analysis to examine what might be the impact of better social sector performance on long term fiscal trajectories. A number of scenarios are outlined in Burton et al (2016) "The benefits of improved social sector performance". The analysis in this note includes four sub-populations which are the focus of the scenarios outlined in that report.

The earlier analytical work had three distinct phases:

1. Individuals in two cohorts (born in 1990/91 or in 1993) were observed in the IDI data from birth through to their early twenties. This enabled us to understand the association between various explanatory characteristics (gender, ethnicity, region, contact with CYF, family welfare history, caregiver corrections contact etc) and education, health, welfare and corrections outcomes as young adults. These relationships are summarised in a set of regression models. We constructed separate models for males and females at each year of age.

2. The second phase involved estimating longer term outcomes for these people based on statistical techniques that involved matching individuals from younger cohorts to individuals in older cohorts. This enabled us to estimate likely future cost trajectories out to age 34 for each individual.

3. Finally a "current" (2013) population of children (aged 0 to 14) had their risk of poorer outcomes estimated using the models from the first phase. Future cost trajectories were also estimated for these children using similar statistical matching techniques between cohorts. This analysis gave us a more contemporary picture of risks and future costs.

This note uses the datasets created in this earlier work combined with a strengths-based measure of being "on track at 21". This measure represents what we hope to see if the cumulative impact of individual, family, community factors and government services mean young people are "on track" for success in adulthood. We define this to be:

- having attained or enrolled in a course at level four or above (training for skilled employment) *or*
- being employed and earning more than two-thirds of median wage for most of their 21st year (approximately the "living wage") *or*
- being self-employed

(Note: we exclude those who served a custodial sentence in their 21st year)

We begin by reviewing the descriptive analysis of the 1993 cohort, showing their interactions with different government agencies up to their early adult years. This is largely drawn from the previous analysis but we have added some hospital event-related data and the “on track at 21” measure. We focus on sub-groups relevant to four scenarios discussed in Burton et al (2016).

We then show how we have constructed risk measures and projections of future fiscal costs for each individual in both the 1993 cohort dataset and the current population dataset.

We then describe how we calculate parameters (cost ratios) which are used in the long-term fiscal modelling of the impacts for each of the four scenarios. These ratios reflect how much we might expect spending on welfare and corrections to reduce under the four scenarios. We do this by modelling changes in the risk distribution and use the observed relationships between risk and future costs in the micro-level population datasets we have created to estimate the possible reductions in future costs.

It should be noted that the results reported in this note have a heavy focus on one cohort (1993). The cohort’s interactions with government social agencies reflect the cyclical economic conditions and social policy settings in place during their upbringing, and the quality of the administrative data systems across a twenty year period. Other population cohorts are likely to experience different economic conditions and social policy settings during their lifetimes and some measures of interactions with government agencies will be better recorded in more recent data. For example, more recent cohorts will not necessarily experience the same level of associations between the factors recorded and the labour market or tertiary education participation rates.

2 Non-fiscal indicators

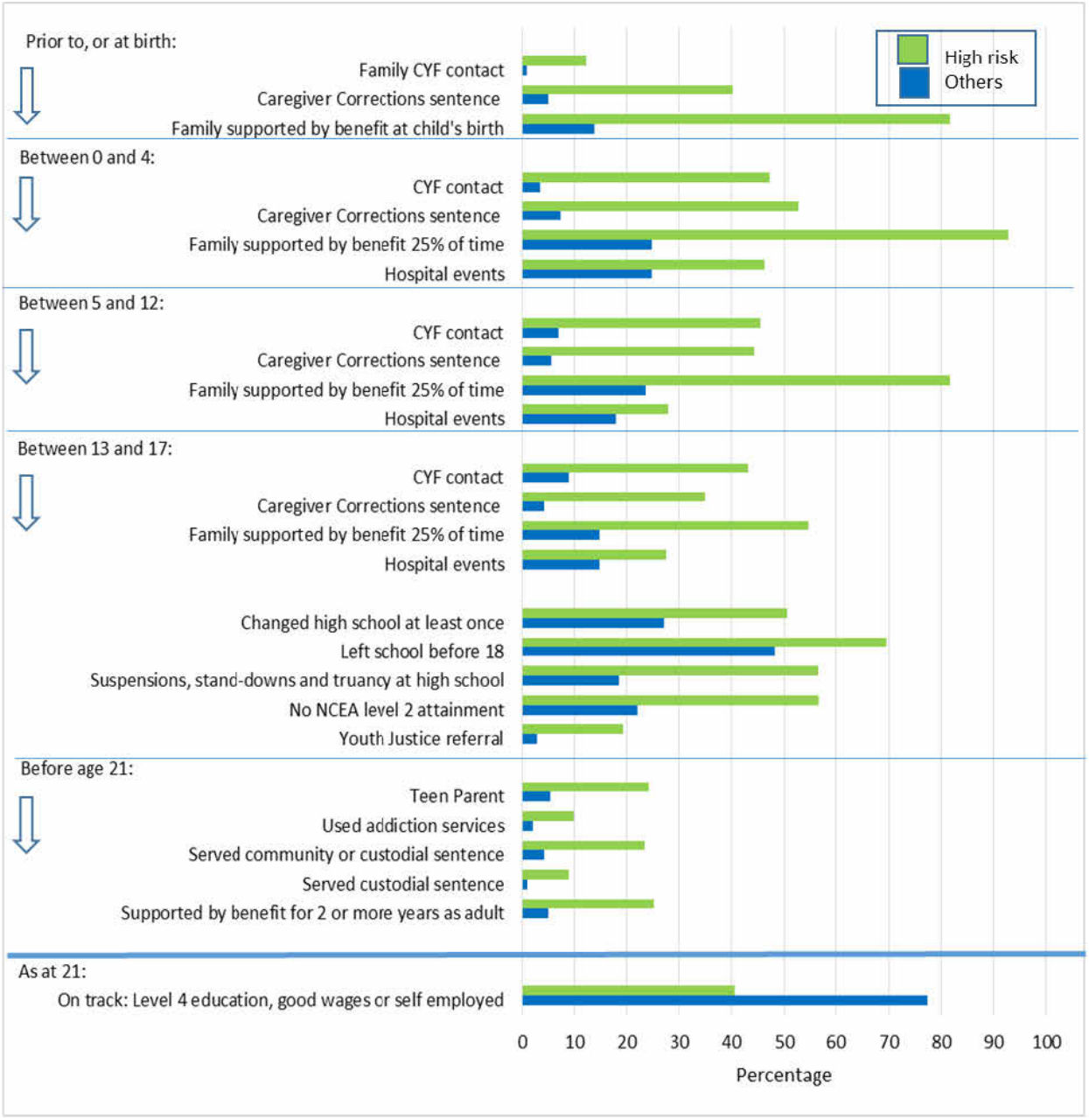
We begin with a description of the characteristics and prevalence of life events for the 1993 cohort. We focus on subgroups relevant to the social investment scenarios in Burton et al (2016).

- **Minimise childhood vulnerability**: Those children we identify at birth as being in the top 10 per cent in terms of risk of poorer welfare and justice outcomes.
- **Equitable Māori outcomes:** We compare Maori and non-Maori children.
- **Broader investment in human capital:** We look at those individuals in the cohort who do not achieve a qualification at level two NCEA or above.
- **Regional convergence:** We compare people living in the 3 largest urban areas (Auckland, Wellington and Christchurch) with the rest of New Zealand.

(Note that modelling the longer term fiscal impacts of each scenario are separate exercises. These subgroups are not mutually exclusive and the resulting impacts cannot be added together to get a combined impact of one or more scenarios.)

Figure 1 shows prevalence of different indicators for those children identified at birth as having a higher risk of poor outcomes as young adults, and compares these to all other children. The identification of the risk groups arose from the regression modelling undertaken in the previous analytical work. This allowed us to construct an equation for each individual that could be used to allocate them a risk score for each outcome of interest (in this case adult welfare receipt and corrections sentences) based on their age and gender as well as a wide range of other characteristics. To help illustrate the '**Minimise childhood vulnerability**' scenario we chose to portray the children with the highest 10% of these risk scores in the 'at birth' model as our 'high risk' children.

Figure 1: Minimise Childhood vulnerability: Comparing children identified at birth as high risk with all others

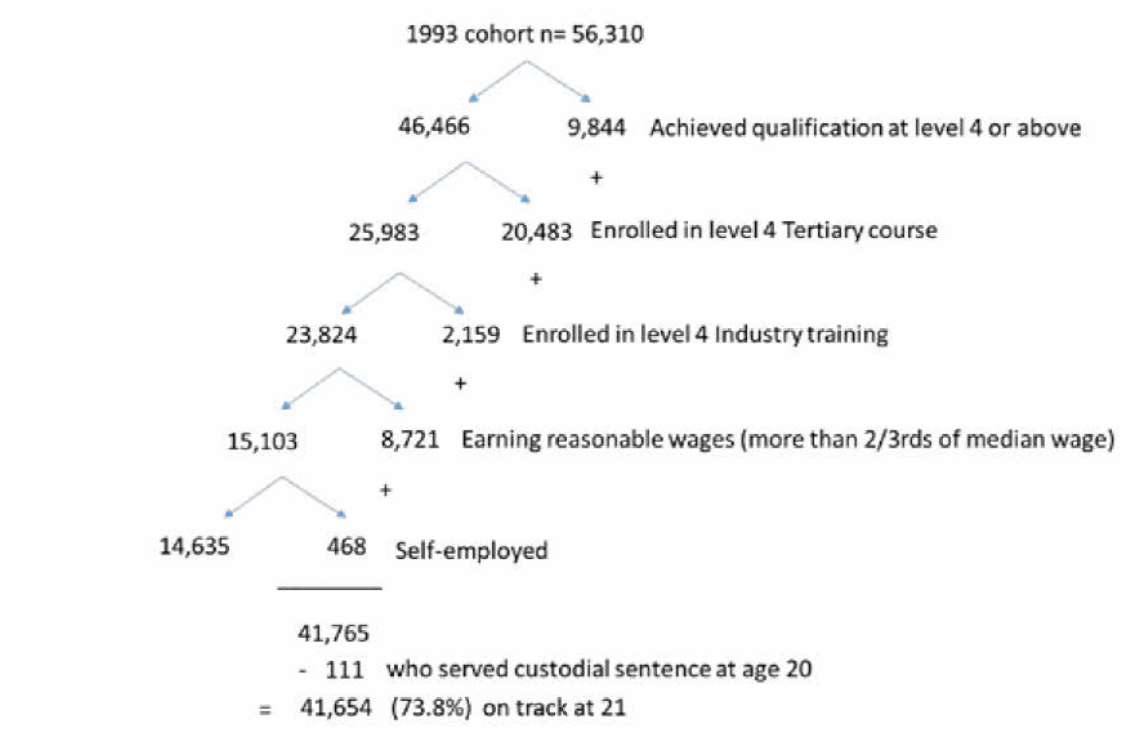


The chart illustrates that we can describe and monitor the prevalence of an array of life events throughout this cohort's life. These include health, education, family welfare, child protection and justice-related events.

Showing the contrast between those identified as at risk at birth and the others provides a sense of the improvements in non-fiscal outcomes that are the aspirational goals under the "minimise childhood vulnerability" scenario. The green bars are the levels for the target population and the blue bars represent the rest of the population (the aspirational benchmark).

The last bars on the chart related to the "on track" measure. Figure 2 illustrates how this has been constructed.

Figure 2: Components of the "on track" measure



Of the original cohort, we identify those who have already achieved a qualification at level four. From the remaining group we find any who are currently enrolled in a level four qualification. Of those without that level of education, we find who earns reasonable wages or who is self-employed. Finally we exclude anyone who has served a custodial sentence in their 21st year. This gives us just under 74 per cent of the cohort who meet the definition of "on track".

Figures 3, 4 and 5 show similar comparisons between the target subgroup and the aspirational benchmark for each of the other scenarios. Figure 3 contrasts Māori and non-Māori. Figure 4 contrasts those who have not achieved NCEA at level 2 to all others. Figure 5 contrasts those in the three main urban centres (Auckland, Wellington and Christchurch) with the rest of New Zealand.

Figure 3: Equitable Māori Outcomes: Comparing Māori and non-Māori

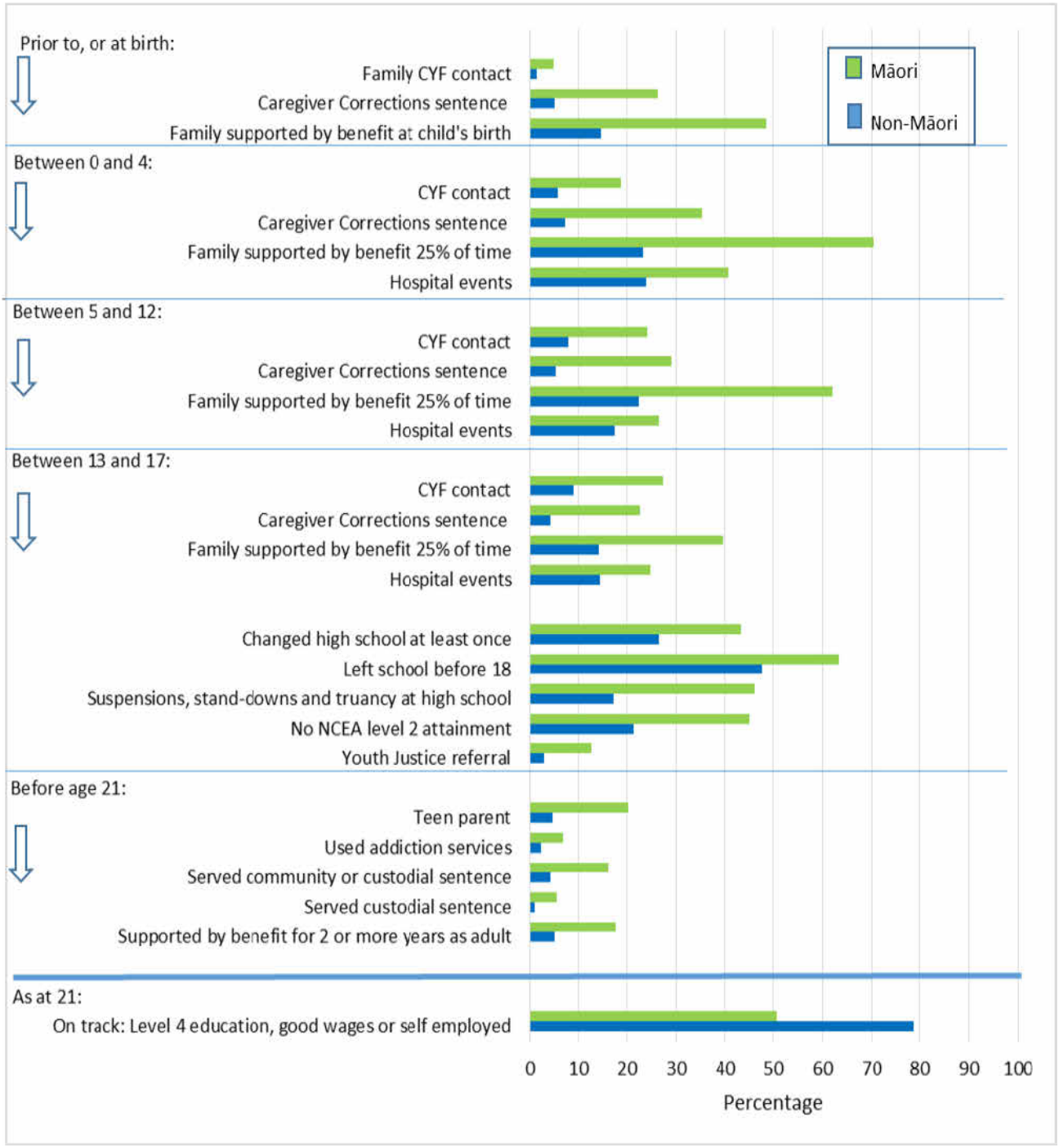


Figure 4: Broader human capital investment: Comparing those who did not achieve NCEA level 2 with those who did

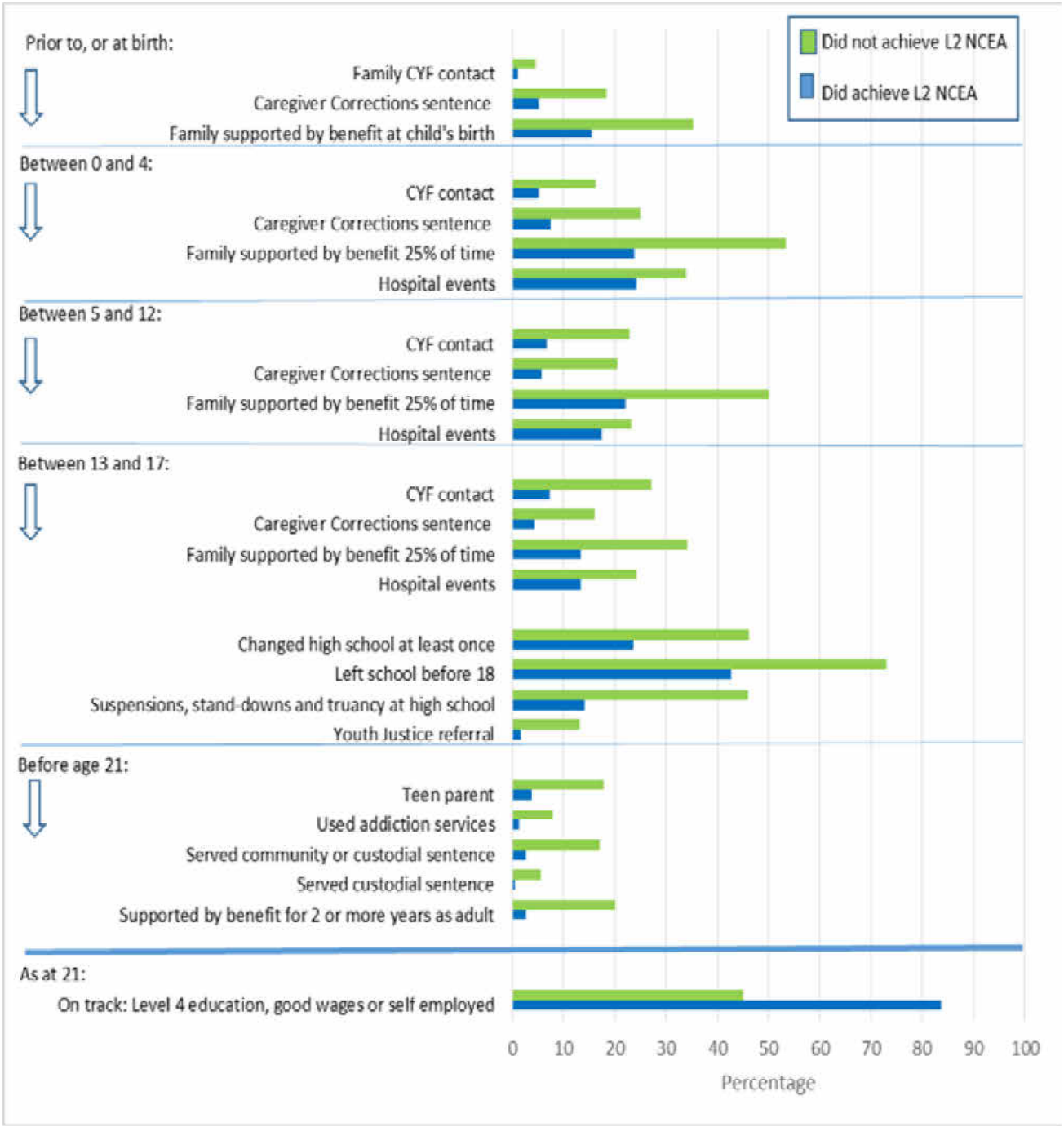
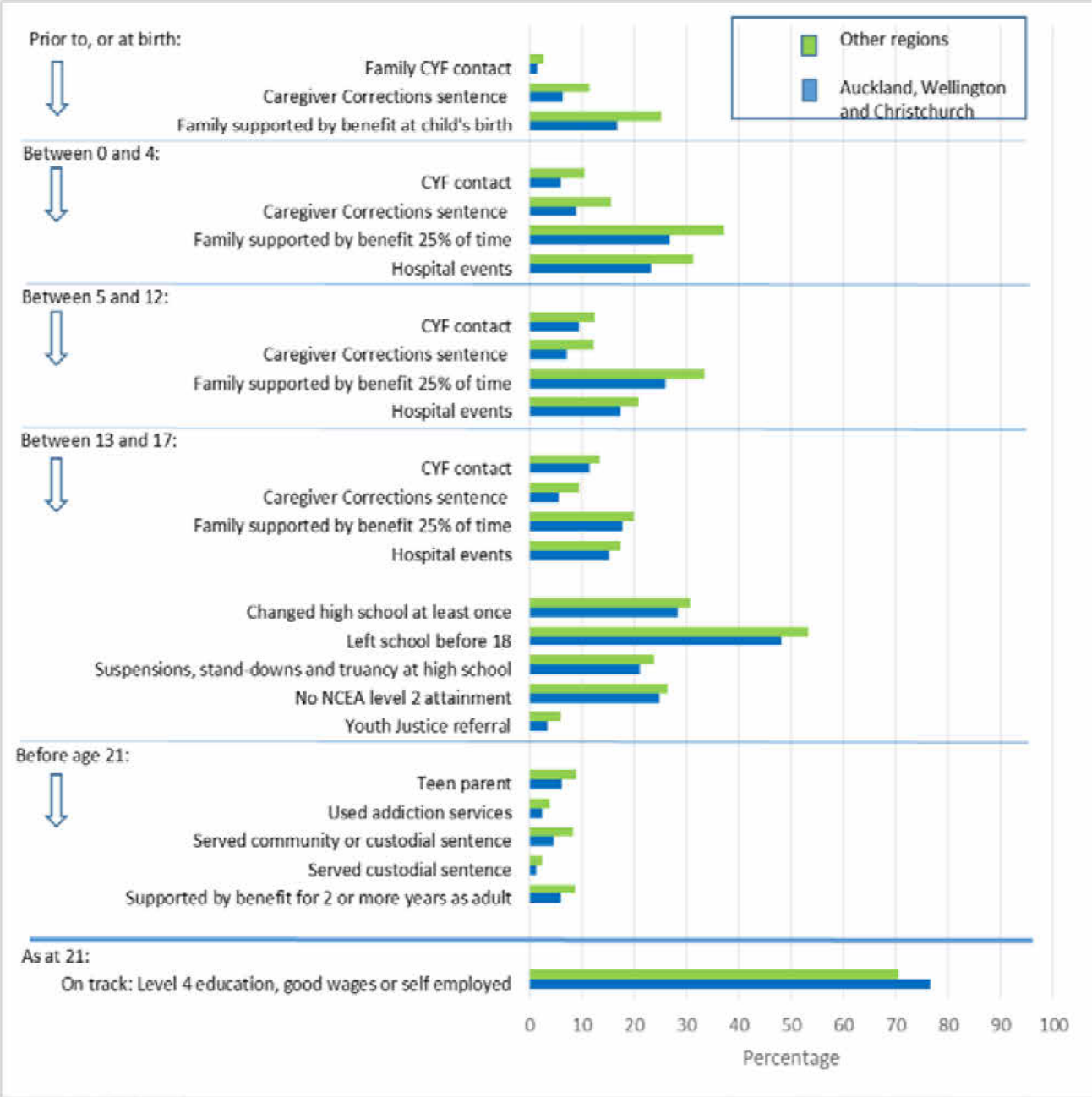


Figure 5: Regional convergence: Comparing the 3 main urban areas with the rest of New Zealand



The figures (1, 3, 4, and 5) are simply devices to present a profile of the current outcomes for the target groups compared to the aspirational benchmark for each scenario. Descriptive comparisons like these are at some risk of being mis-interpreted. Differences in composition of the two groups we are comparing will explain much of the difference in the various indicators we have presented. We are not implying that there are independent educational, regional, ethnic or early age risk effects of this magnitude.

3 Calculating risk at a micro-data level

3.1 Calculating risk

In Treasury's earlier analytical work predictive modelling was used to investigate the extent to which various characteristics (observed through ages 0 to 14) were associated with poor outcomes as young adults.

Four outcome measures were selected and defined as follows:

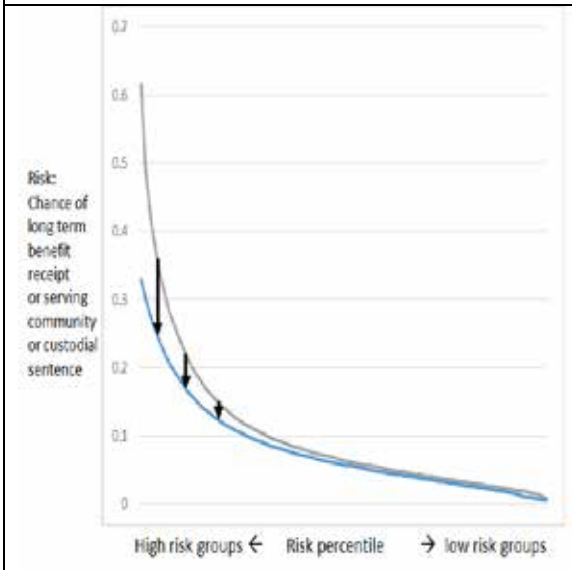
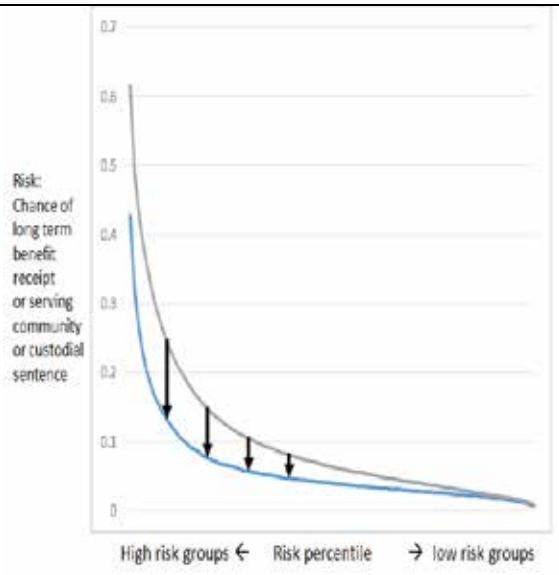
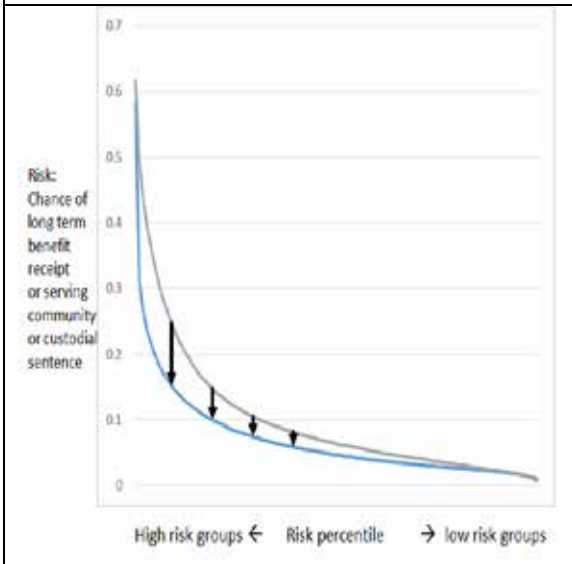
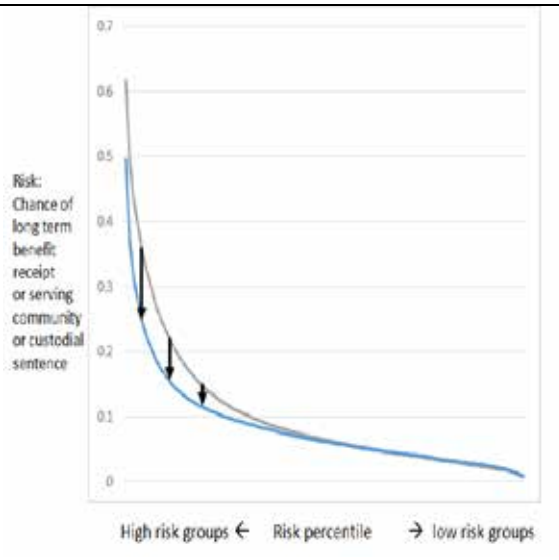
1. not achieving at least a Level 2 education qualification by age 19
2. use of mental health or addiction services whilst aged between 18 and 20
3. receiving a custodial or community sentence before age 21
4. being on benefit for 2 years or more before age 21.

Logistic regression models were run at each year of age for females and males separately for four outcome measures. Forward selection was used to select the model. This process allowed us to identify the key indicators for each age/gender combination and outcome measure, and calculate a predicted risk score for each outcome for each individual in the population.

In this note the maximum risk score across the (welfare and corrections outcomes) was used to identify the 10 per cent of children with the highest predicted score. This score indicates the combination of risk factors most associated with the projected fiscal costs that we modelled.

3.2 Risk profiles for the different scenarios

Figure 6: Risk profile impact for each scenario

<table>
<tr><td>Minimise childhood vulnerability scenario</td><td>Equitable Māori outcomes scenario</td></tr>
<tr><td> 
</td><td> 
</td></tr>
<tr><td>Broader human capital investment scenario</td><td>Regional convergence scenario</td></tr>
<tr><td> 
</td><td> 
</td></tr>
</table>

The modified risk profiles (blue curve) show the expected impact of each of the social investment scenarios, compared with the original risk distribution (grey curve):

1. The **Minimise childhood vulnerability** scenario targets services to those children and young people at highest risk of long-term unemployment and poor criminal justice outcomes. We create a benchmark where the risk profile is "translated" across the horizontal axis. This means that the risks for those identified as at the highest risk (100th to 95th percentiles) are reduced and now reflect the next level of risk (ie, the 94th to 90th percentiles). All other risk levels are similarly translated across.

2. The **Equitable Māori outcomes** scenario focuses on attempting to deliver services for Māori with fiscal and non-fiscal benefits so that Māori experience the same outcomes as the rest of the population. For modelling purposes we assign to Māori the levels of risk that are currently observed by the non-Māori population.

3. The **Broader investment in human capital** scenario focuses on policies to improve the health and educational components of human capital. Its results are based on the assumption that improved health and educational services could lead to higher labour market participation without reducing overall productivity. The long-term fiscal modelling uses a variety of approaches to capture the different aspects of this scenario. The IDI analysis provided estimates of the impact of higher levels of NCEA attainment at levels 2 and 3. This is modelled by looking at the risk levels associated with each level of NCEA attainment in the 1993 cohort and re-weighting the overall risk profile to reflect the new target NCEA attainment levels.

4. The **Regional convergence** scenario focuses on closing the gap in outcomes between people living in the regions with those living in Auckland, Christchurch and Wellington. For modelling purposes the benchmark profile reflects the current risk profile of these three largest urban areas.

4 Projecting fiscal costs at a micro-data level

The scenarios we examined using the IDI dataset for the LTFM analysis each identified a sub-population that would be the main focus for improved social services delivery:

The individual-level datasets (both the 1993 cohort and the current dataset of children) enabled us to identify these populations directly, and estimate their risk profiles and their projected fiscal costs.

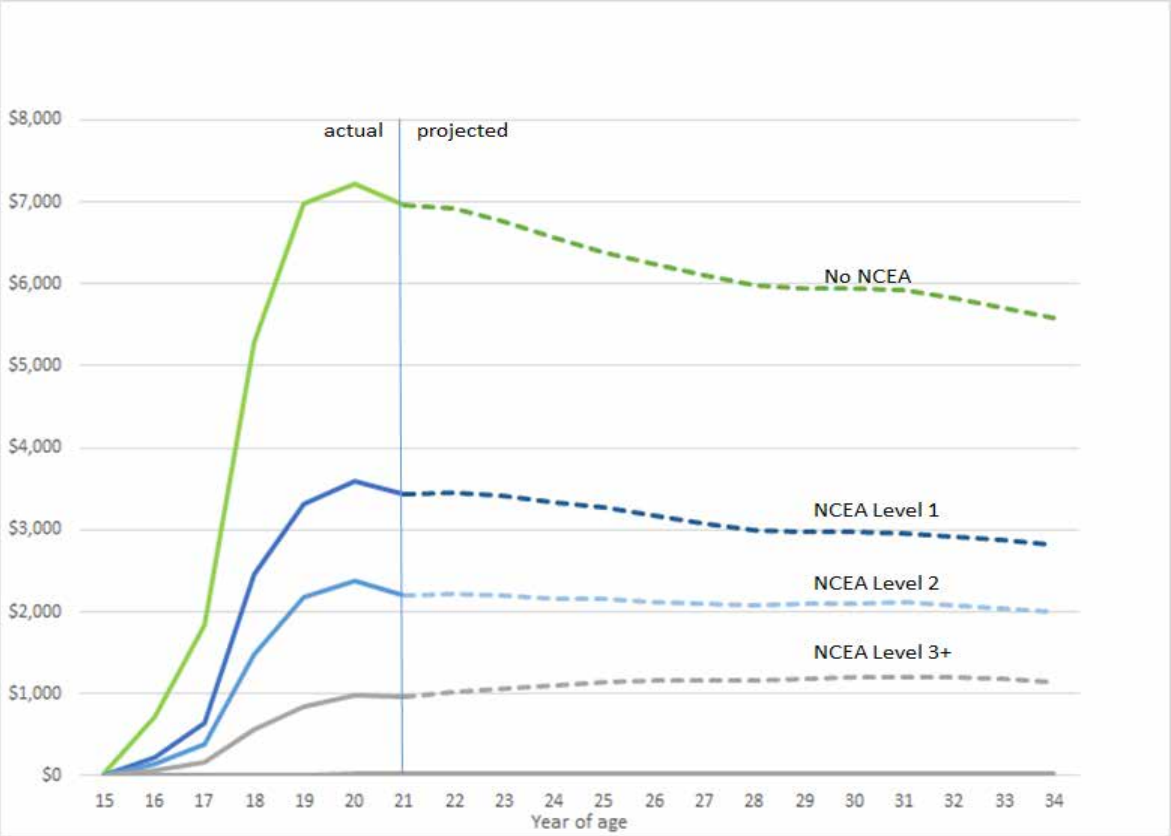
4.1 Calculating projected fiscal costs at a micro-data level

In Treasury's earlier analytical work statistical record linkage was used to help estimate the likely longer-term outcomes of the study population. The approach involved linking data for an older birth cohort (specifically the July 1978 to June 1979 birth cohort) to the data for the 1993 birth cohort, to simulate the likely outcomes for this latter population. Records were linked on the basis of benefit receipt and corrections sentencing rates and patterns when aged 16 to 21 years inclusive, as well as on the basis of gender and ethnicity. Observed outcomes and costs experienced by the 1978/79 cohort were then used to estimate the outcomes and costs of the 1993 cohort up to age 35.

Using a similar matching technique the outcomes of the current population of children aged 0 to 14 years are estimated by linking each of them to an individual from the 1993 cohort. Records are linked on the basis of the child's contact with child and protection services, caregivers' benefit receipt, caregivers' corrections sentencing history, and some early secondary school enrolment data (for the 13 and 14 year olds) as well as gender and ethnicity. The link through to the 1978/79 birth cohort provides outcome and cost projections to age 35 for all children aged 0 to 14 years.

Matching individuals rather than population groups gives us the flexibility to estimate costs for very different subsets of the population. This is particularly important when we are looking to identify specific target populations for investment decisions. The statistical matching method uses real patterns for individuals over time with very similar observed characteristics up to a certain age.

Figure 7: Projected fiscal cost trajectories by NCEA attainment (1993 cohort)



The approach assumes longitudinal patterns of benefit receipt and corrections sentences can be moved around in time from one cohort to another, and that, conditional on a set of "early indicator" matching variables, these patterns remain relevant to later cohorts. The success of this approach depends on how well we establish good matching criteria and on how relevant these are for forecasting future outcomes. We have also not accounted for differences in macro-economic conditions experienced by the two cohorts. As a result, future outcome estimates will in part reflect the particular patterns of labour demand and unemployment that have occurred over the last 20 years. Ideally we would like to remove the effects of these macro-economic fluctuations and have a more constant underlying macro-economic picture underpinning the analysis. This remains an issue for further investigation.

Long-run shifts in New Zealand's social assistance policies could also influence the success of the cohort matching if they have affected the outcomes of different birth cohorts very differently. Ideally, we would adjust individuals' outcomes to remove the effects any secular trends that are external to the individual but affect the outcomes of the cohort as a whole. In practice, however, it may be difficult to do so in an objective way using the data currently available.

4.2 Understanding the association between risk and future fiscal costs

The resulting large micro-level datasets (approx. 60,000 records in the 1993 cohort, and approx. 800,000 records in the current population of children aged less than 15) allow us to examine the association between risk and future fiscal costs and to do this for different sub-populations.

Figure 8 shows the level of projected welfare costs (per person, up to age 35), how this varies from high to low risk, and the estimated impact under each scenario.

Figure 8: Risk and future welfare costs

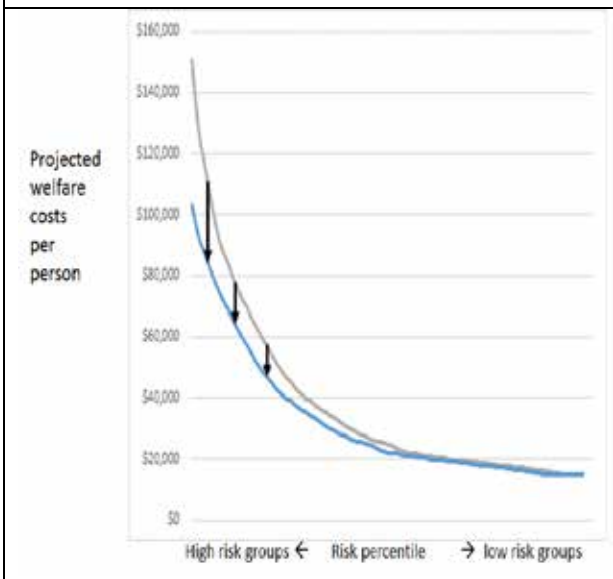
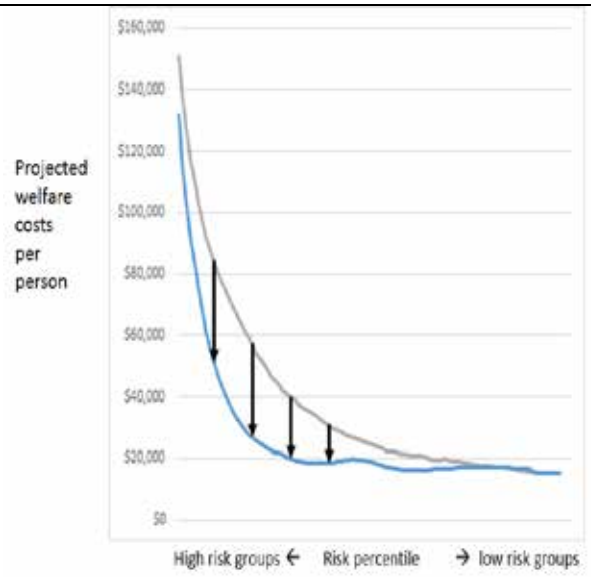
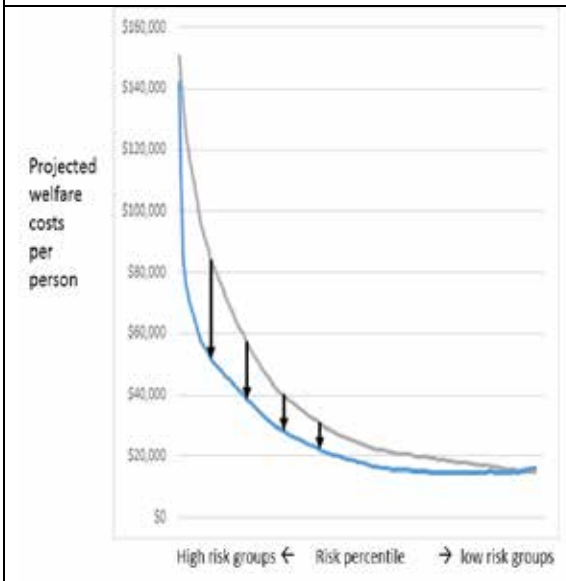
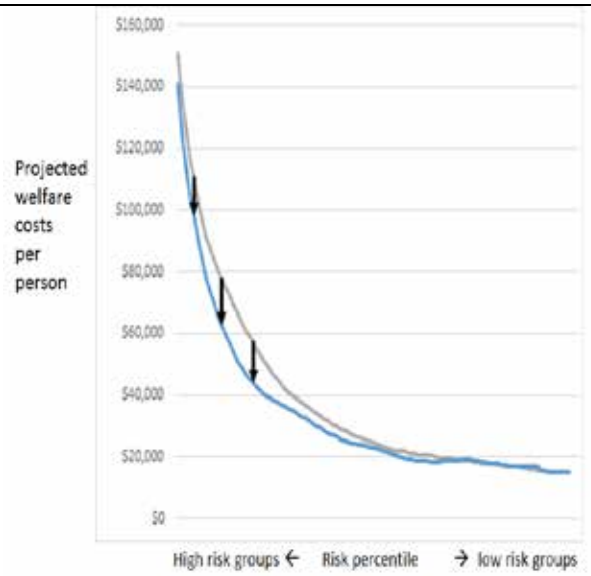
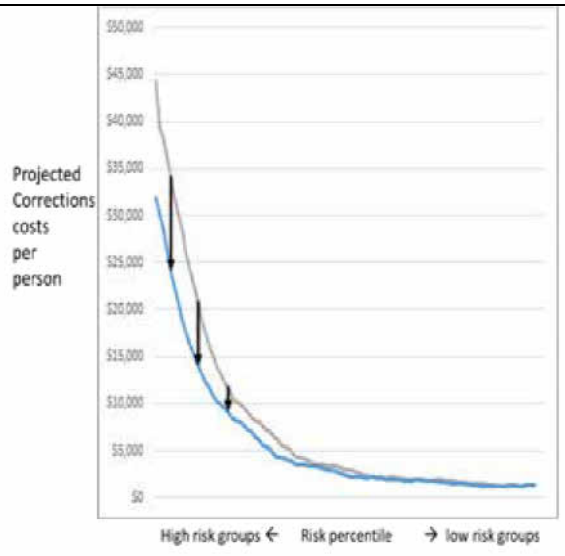
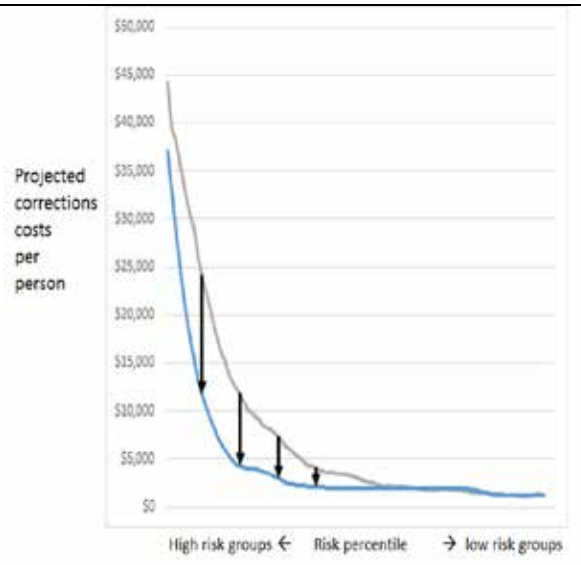
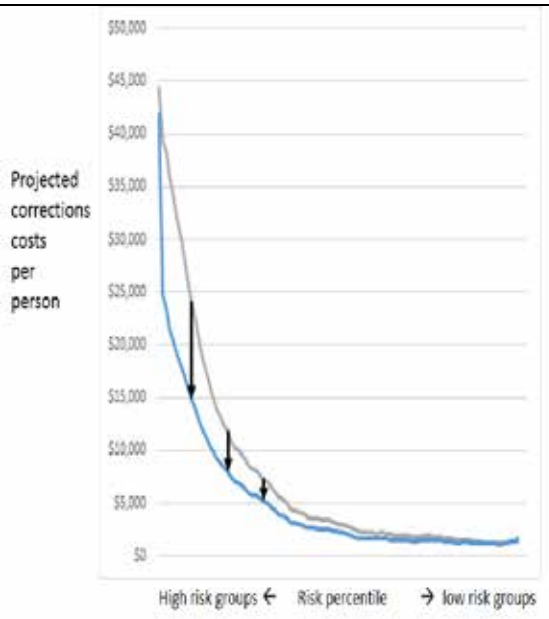
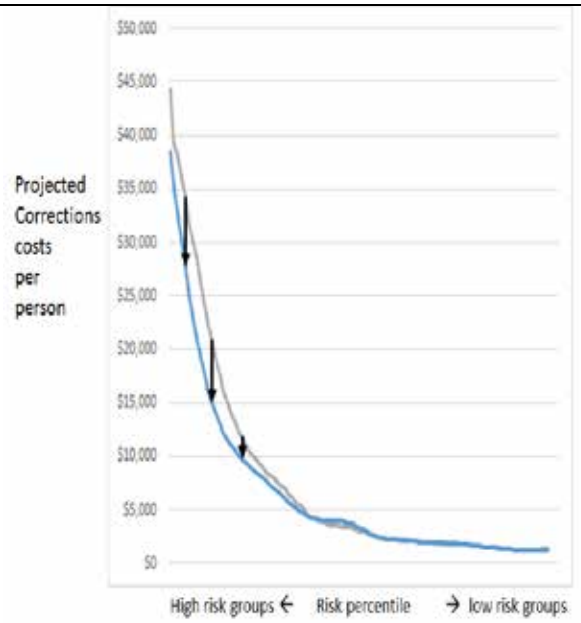
<table>
<tr><td>Minimise childhood vulnerability scenario</td><td>Equitable Māori outcomes scenario</td></tr>
<tr><td> 
</td><td> 
</td></tr>
<tr><td>Broader human capital investment scenario</td><td>Regional convergence scenario</td></tr>
<tr><td> 
</td><td> 
</td></tr>
</table>

Figure 9 shows the level of projected corrections costs (per person, up to age 35), how this varies from high to low risk, and the estimated impact under each scenario.

Figure 9: Risk and future corrections costs

<table>
<tr><td>Minimise childhood vulnerability scenario</td><td>Equitable Māori outcomes scenario</td></tr>
<tr><td> 
</td><td> 
</td></tr>
<tr><td>Broader human capital investment scenario</td><td>Regional convergence scenario</td></tr>
<tr><td> 
</td><td> 
</td></tr>
</table>

4.3 Calculating the cost ratios for each scenario

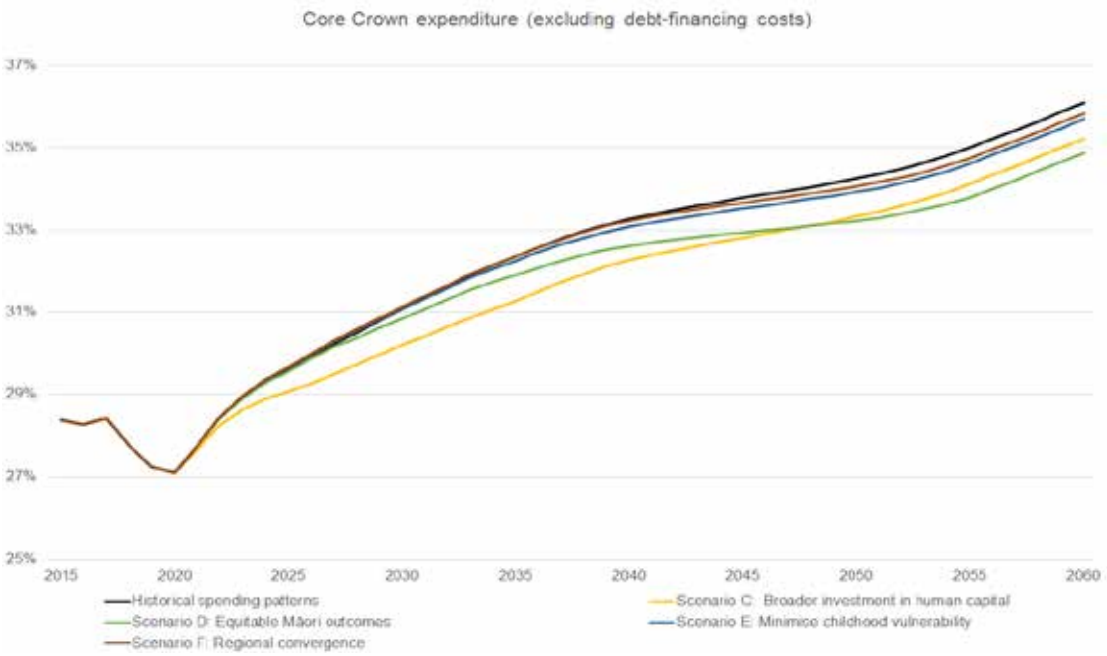
The parameters needed for the long term fiscal model representing each of the scenarios likely impact on fiscal costs (as far as we could cover in our IDI data) were calculated for welfare and corrections costs separately. The parameters are cost ratios, which are the ratio of the area under the blue curve to the area under the grey curve in each of the above 8 graphs from Figures 8 and 9. Equivalently they can expressed as the ratio of mean welfare (or corrections) costs for the population (under the modified risk profile) to the mean welfare (or corrections) costs observed in the current population.

Figure 10: Cost ratios for each scenario

		Minimise Childhood vulnerability	Equitable Māori Outcomes	Human Capital Investment	Regional Convergence
Corrections	Ratio	0.79	0.52	0.75	0.82
Welfare	Ratio	0.88	0.65	0.84	0.87

In the long term fiscal modelling we apply these ratios to current levels of spending to create new target benchmarks. We transition to these new benchmark levels in a steady linear fashion over a pre-determined span of time. Figure 11 is an illustration of the results of long term fiscal modelling of these scenarios. These later stages of the long term fiscal modelling are described in Burton et al (2016).

Figure 11: Fiscal track under the social investment scenarios



Source: Burton et al (2016)

5 Summary

The note has used integrated administrative data from the IDI (Integrated Data Infrastructure at Statistics New Zealand) to provide measures of potential fiscal impacts of four aspirational social investment scenarios. It has used two large analytical datasets created from earlier Treasury projects that created predictive risk models and projected future fiscal costs at a microdata level. These datasets meant we could create a descriptive picture of potential non-fiscal outcomes and estimate fiscal impacts of these targeted social investment scenarios. The descriptions of the scenarios and the resulting estimated fiscal trajectories can be found in Burton et al (2016).

Appendix 1: Study population and definitions

The study uses data from Statistics New Zealand's Integrated Data Infrastructure (IDI), which combines and integrates administrative data from a range of government agencies, including Inland Revenue, the Ministry of Education, the Ministry of Health, the Ministry of Social Development and the Department of Internal Affairs.[1]

The full IDI population of identities is an 'ever in NZ' population covering any NZ born individual, anyone granted a permanent residency visa since 1998 and any one ever registered with the IRD. Depending on the research objectives this 'super-population' generally needs to be filtered to remove people not in NZ for very long, people who left NZ a long time ago.

The 1993 cohort population used to produce the "life charts" in this study comprises New Zealanders who were born in 1993 and satisfied the conditions listed in Table 1.

Appendix Table 1 - Study population criteria

	Selection criteria
1	Born in 1993
2	Enrolled as domestic students in New Zealand schools in 2008 or 2009 (ie, when they were aged 15 and 16 years
3	Living in NZ for most (6 months or more) of their 20th year
4	Link to an identity on the IDI 'spine' (this means they have an IRD number or are NZ born or were granted a permanent residency visa)

There are 56,300 individuals satisfying these conditions in the IDI.

Defining the 1993 cohort population in this way means that we exclude those for whom we have insufficient evidence to determine whether they are "on track". This means:

- We **do not** include New Zealand resident individuals born in 1993 who had left New Zealand before their secondary schooling years.
- We **do not** include individuals born in 1993 who attended secondary school in NZ but who subsequently have emigrated.
- We **do** include non-NZ born individuals who went to secondary school in NZ (as domestic students in 2008 or 2009) and who appear to be still in NZ at age 21.

Table 2 describes the various characteristics and life events used to describe the interactions the different groups of children in the cohort have had with government agencies throughout their childhood.

[1] See http://www.stats.govt.nz/browse_for_stats/snapshots-of-nz/integrated-data-infrastructure.aspx for more information about IDI.

Appendix Table 2 - Characteristics and life events – notes on definitions

Characteristic	Notes:	Number in 1993 Cohort	Prevalence in 1993 Cohort
Older sibling had CYF notification prior to child's birth	For this analysis only siblings that are present in welfare data are included	1,098	1.9%
Caregiver had served corrections sentence prior to child's birth	Only caregivers identified through the welfare data are included (it is not until 1998 data that birth parents are fully identified in IDI data)	4,863	8.6%
Family was supported by welfare at child's birth		11,571	20.5%
Child had CYF notification prior to age 5	Note that CYF data lacks completeness pre-2000 so affects the early age data for this cohort	4,467	7.9%
Caregiver with corrections sentence whilst child less than 5	Corrections sentences include community or custodial sentences	6,720	11.9%
Family supported by welfare whilst child less than 5	This refers to where family was supported more than 25% of the days child was less than 5	17,637	31.3%
Child spent at least 1 night in hospital aged less than 5		15,081	26.8%
Child had CYF notification whilst aged 5 to 12	Note that CYF data lacks completeness pre-2000 so affects the early age data for this cohort	6,078	10.8%
Caregiver with corrections sentence whilst child aged 5 to 12		5,316	9.4%
Family supported by welfare whilst child aged 5 to 12	This refers to where family was supported more than 25% of the days child was between 5 and 12	16,473	29.3%
Child spent at least 1 night in hospitalwhilst aged 5 to 12		10,680	19.0%
Child had CYF notification whilst aged 13 to 17		6,942	12.3%
Caregiver with corrections sentence whilst child aged 13 to 17		4,110	7.3%
Family supported by welfare whilst child aged 13 to 17	This refers to where family was supported more than 25% of the days child was between 13 and 17	10,524	18.7%
Child spent at least 1 night in hospital aged 13 to 17		9,084	16.1%
Changed high school at least once		16,530	29.4%
Left school before turning 18		28,359	50.4%
Truant, suspended or stood down from school		12,537	22.3%
Did not achieve NCEA level 2 equivalent		14,337	25.5%
Referred to CYF Youth Justice services	Note: this indicator refers only to CYF Youth Justice services	2,571	4.6%
Parent before aged 20		4,161	7.4%
Used drug or alcohol addiction services before 20	Mental health service use data complete only up to 2012 when this analysis completed	1,680	3.0%
Served Corrections sentence before age 21	This refers to community sentences served as an adult	3,555	6.3%
Served custodial sentence before aged 21	This refers to custodial sentences served as an adult	1,005	1.8%
Received welfare benefit for more than 2 years before 21	This refers to main welfare benefits received as an adult	4,026	7.1%
Child ever had CYF notification	(these indicators were derived for use in the main report: He Tirohanga Mokopuna (the Statement of the Long Term Fiscal Position 2016).	11,538	20.5%
Family supported by welfare 25% of time aged 0-4 or 5-12 or 13-17		22,485	39.9%
Child spent at least 1 night in hospital whilst aged 0-17		26,046	46.3%

Appendix 2: Non-fiscal indicators

Table 3 shows the non-fiscal outcome assumptions for each of the scenarios modelled. The data in the table is presented in Figures 1, 3, 4 and 5 in the body of the note.

Appendix Table 3 - Comparisons for each scenario sub-population

Characteristic	Minimise childrens vulnerability		Equitable Maori Outcomes		Regional Convergence		Human Capital Investment	
	High Risk	Others	Māori	Non-Māori	Other Regions	Auck/Well /Chch	NCEA < L2	NCEA l2+
Older sibling had CYF notification prior to child's birth	12.4%	0.8%	4.7%	1.4%	2.6%	1.5%	4.5%	1.1%
Caregiver had served corrections sentence prior to child's birth	40.3%	5.2%	26.3%	4.9%	11.5%	6.3%	18.5%	5.3%
Family was supported by welfare at child's birth	81.5%	13.9%	48.5%	14.7%	25.2%	16.8%	35.4%	15.5%
Child had CYF notification prior to age 5	47.2%	3.7%	18.8%	5.7%	10.4%	5.9%	16.2%	5.1%
Caregiver with corrections sentence whilst child less than 5	40.3%	5.2%	26.3%	4.9%	11.5%	6.3%	18.5%	5.3%
Family supported by welfare whilst child less than 5	92.8%	24.7%	70.4%	23.1%	37.2%	26.6%	53.3%	23.8%
Child spent at least 1 night in hospital aged less than 5	46.3%	24.7%	40.8%	23.9%	31.3%	23.2%	34.0%	24.3%
Child had CYF notification whilst aged 5 to 12	45.6%	7.0%	24.0%	8.0%	12.4%	9.5%	22.9%	6.7%
Caregiver with corrections sentence whilst child aged 5 to 12	44.4%	5.7%	29.1%	5.3%	12.3%	7.2%	20.4%	5.7%
Family supported by welfare whilst child aged 5 to 12	81.7%	23.6%	62.0%	22.4%	33.5%	25.9%	50.1%	22.2%
Child spent at least 1 night in hospitalwhilst aged 5 to 12	27.9%	18.0%	26.5%	17.4%	20.9%	17.4%	23.4%	17.5%
Child had CYF notification whilst aged 13 to 17	43.1%	9.0%	27.3%	9.2%	13.4%	11.5%	27.1%	7.3%
Caregiver with corrections sentence whilst child aged 13 to 17	35.0%	4.3%	22.6%	4.1%	9.5%	5.5%	16.0%	4.3%
Family supported by welfare whilst child aged 13 to 17	54.6%	14.8%	39.8%	14.3%	20.0%	17.7%	34.1%	13.4%
Child spent at least 1 night in hospital aged 13 to 17	27.6%	14.9%	24.7%	14.3%	17.2%	15.3%	24.2%	13.4%
Changed high school at least once	50.6%	27.1%	43.3%	26.4%	30.6%	28.4%	46.2%	23.6%
Left school before turning 18	69.7%	48.3%	63.4%	47.6%	53.2%	48.1%	73.1%	42.6%
Truant, suspended or stood down from school	56.6%	18.5%	46.1%	17.3%	23.9%	21.0%	46.0%	14.2%
Did not achieve NCEA level 2 equivalent	56.8%	22.1%	45.2%	21.3%	26.3%	24.8%		
Referred to CYF Youth Justice services	19.3%	3.0%	12.7%	2.9%	5.9%	3.5%	13.1%	1.7%
Parent before aged 20	24.1%	5.6%	20.3%	4.7%	8.8%	6.2%	17.8%	3.8%
Used drug or alcohol addiction services before 20	10.0%	2.2%	6.8%	2.2%	3.7%	2.4%	8.0%	1.3%
Served Corrections sentence before age 21	23.4%	4.5%	16.1%	4.3%	8.3%	4.7%	17.1%	2.6%
Served custodial sentence before aged 21	9.0%	1.0%	5.4%	1.0%	2.4%	1.3%	5.6%	0.5%
Received welfare benefit for more than 2 years before 21	25.1%	5.2%	17.6%	5.0%	8.6%	6.0%	20.2%	2.7%
Child ever had CYF notification	72.4%	14.9%	42.9%	15.8%	18.2%	23.4%	40.3%	13.7%
Family supported by welfare 25% of time aged 0-4 or 5-12 or 13-17	95.6%	33.9%	76.8%	32.2%	36.2%	44.6%	62.9%	32.1%
Child spent at least 1 night in hospital whilst aged 0-17	68.2%	43.9%	64.1%	42.5%	42.4%	51.1%	56.9%	42.6%
On track at 21	40.7%	77.4%	50.6%	78.7%	70.4%	76.6%	45.1%	83.6%
Cohort numbers	5,508	50,799	9,747	46,563	25,029	31,278	14,337	41,970

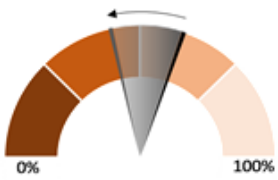
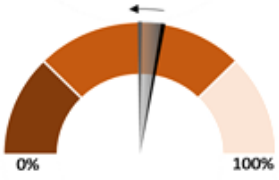
Appendix 3: Dial graphs

Many children appear to have what we would call “poor starts” in life, but proceed to be at good levels of the education system or in relatively well-paid employment by their early twenties. We were interested in how much the linked administrative data can tell us about the differences between these children and those that end up less educated, not employed and/or in less well-paying employment.

To begin to address this we separated out the 1993 cohort into those whose family was supported by welfare at their birth and those whose families were not supported by welfare at their birth. Then we looked at “on track” rates for these children when they had also experienced each of a series of life events later in their life.

A selection of these rates (see Appendix Table 4) are presented in a dial graph format in Burton et al (2016). The dials present the proportion of children who are “on track at 21” and shows this for children who experience a particular life event (or change in circumstance) and those who did not. The difference between these proportions (the grey area), represents the size of the association we observe between the ‘on track’ outcome and the adverse childhood event.

Appendix Figure 1 - ‘On track at 21’ dial graphs

<table>
<tr><td></td><td colspan="3">Proportion who are on track at 21</td></tr>
<tr><td></td><td colspan="3">For children where family was supported by benefit at birth</td></tr>
<tr><td>Event: Caregiver corrections sentence (whilst child aged 0 - 4 years)</td><td> 
</td><td>Event: Changed high school at least once</td><td> 
</td></tr>
<tr><td>Event: Child spent 1 night in hospital aged 5 to 12</td><td> 
</td><td>Event: Parent before aged 20</td><td> 
</td></tr>
</table>

A more comprehensive list of life events and their impact on the likelihood of a person being “on track” is set out in Appendix Table 4. The events in bold correspond to the four events pictures in the dials in Appendix Figure 1.

Appendix Table 4 - Proportion 'on track at 21' and experience of other childhood events (1993 cohort)

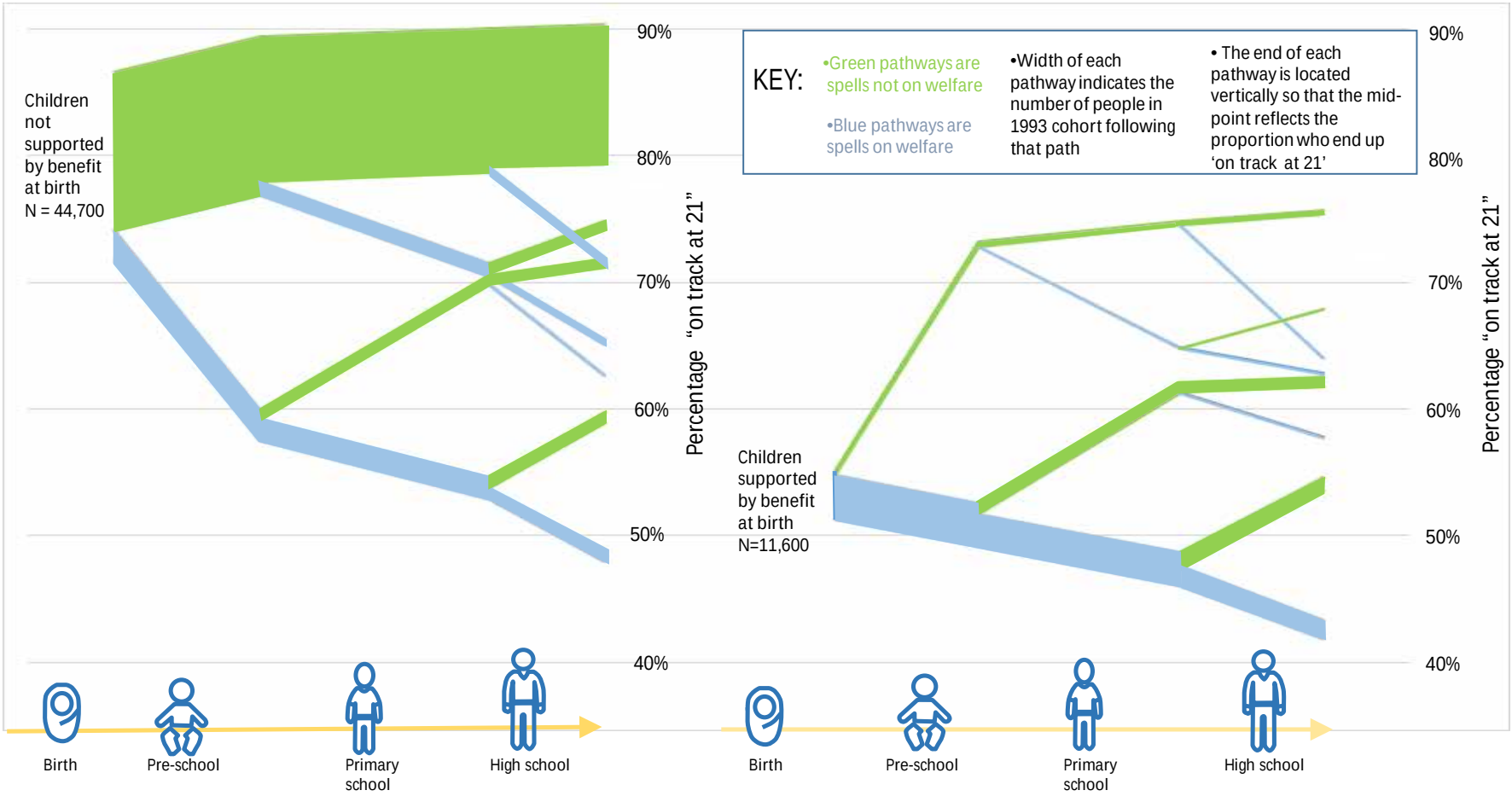
Event	Supported by benefit at birth		Not supported by benefit at birth	
	Experienced the event	Did not experience the event	Experienced the event	Did not experience the event
Older sibling had CYF notification prior to child's birth	38.0%	54.3%	45.2%	79.4%
Caregiver had served corrections sentence prior to child's birth	41.6%	57.8%	47.3%	80.3%
Child had CYF notification prior to age 5	39.6%	57.3%	50.9%	80.3%
Family was supported by welfare at child's birth	50.9%	73.4%	58.8%	83.1%
Caregiver with corrections sentence whilst child less than 5	42.1%	60.0%	48.8%	80.8%
Child spent at least 1 night in hospital aged less than 5	48.7%	56.3%	74.3%	80.6%
Family supported by welfare whilst child aged 5 to 12	47.8%	65.6%	60.3%	83.5%
Child had CYF notification whilst aged 5 to 12	38.6%	58.7%	52.9%	81.0%
Caregiver with corrections sentence whilst child aged 5 to 12	39.6%	58.8%	47.0%	80.6%
Child spent at least 1 night in hospital whilst aged 5 to 12	49.6%	54.5%	74.4%	80.1%
Caregiver with corrections sentence whilst child aged 13 to 17	39.3%	57.2%	47.1%	80.3%
Child had CYF notification whilst aged 13 to 17	35.9%	60.1%	49.7%	81.8%
Family supported by welfare whilst child aged 13 to 17	44.2%	60.2%	58.6%	82.0%
Child spent at least 1 night in hospital aged 13 to 17	41.8%	56.5%	69.2%	80.8%
Changed high school at least once	42.6%	60.4%	69.2%	82.7%
Left school before turning 18	47.0%	63.6%	73.6%	84.1%
Truant, suspended or stood down from school	38.5%	63.4%	57.9%	83.6%
Did not achieve NCEA level 2 equivalent	32.0%	69.9%	52.2%	86.1%
Referred to CYF Youth Justice services	30.0%	56.2%	42.6%	80.2%
Parent before aged 20	28.0%	58.2%	36.5%	81.4%
Used drug or alcohol addiction services before 20	29.6%	54.8%	46.3%	79.9%
Served Corrections sentence before age 21	28.5%	57.4%	40.5%	80.8%
Served custodial sentence before aged 21	10.0%	55.4%	16.0%	79.8%
Received welfare benefit for more than 2 years before 21	13.7%	61.2%	19.7%	82.1%

Appendix 4: Family welfare history diagram

The diagram in Appendix Figure 2 separates out those children (from the 1993 cohort) whose families were not supported by welfare at the child's birth (left hand side of graph) and those who were supported by welfare at the child's birth (right hand side of graph.)

Each of these groups is then split successively into separate branches depending on whether their family was on welfare for at least 25 per cent of the time while they were aged 0 to 4 years old, 5 to 12 years old and then 13 to 17 years old. The location of each grouping illustrates the percentage who are "on track" at 21. Hence the figure shows the likelihood of being on track at 21 given a family's particular pattern of welfare receipt.

Appendix Figure 2 - Family welfare history and adult outcomes



Appendix Figure 2 shows:

- Higher "on track at 21" rates for those with no history of being supported by welfare at birth or during their childhood and that this is by far the largest sub-group.
- Relatively even negative slopes for those who experience welfare at all ages from pre-school to high school, regardless of the extent of welfare in their lives up to that point (i.e. the slopes of the branches are reasonably similar when comparing them with others above or below them in the diagram).
- Smaller but (generally) consistent negative slopes for those with periods of welfare later in childhood (ie, the slopes of the branches get progressively smaller when comparing them with others to the right of them in the diagram).
- Those with two or more terms of welfare support have lower "on track" rates compared to those without multiple terms of welfare support.

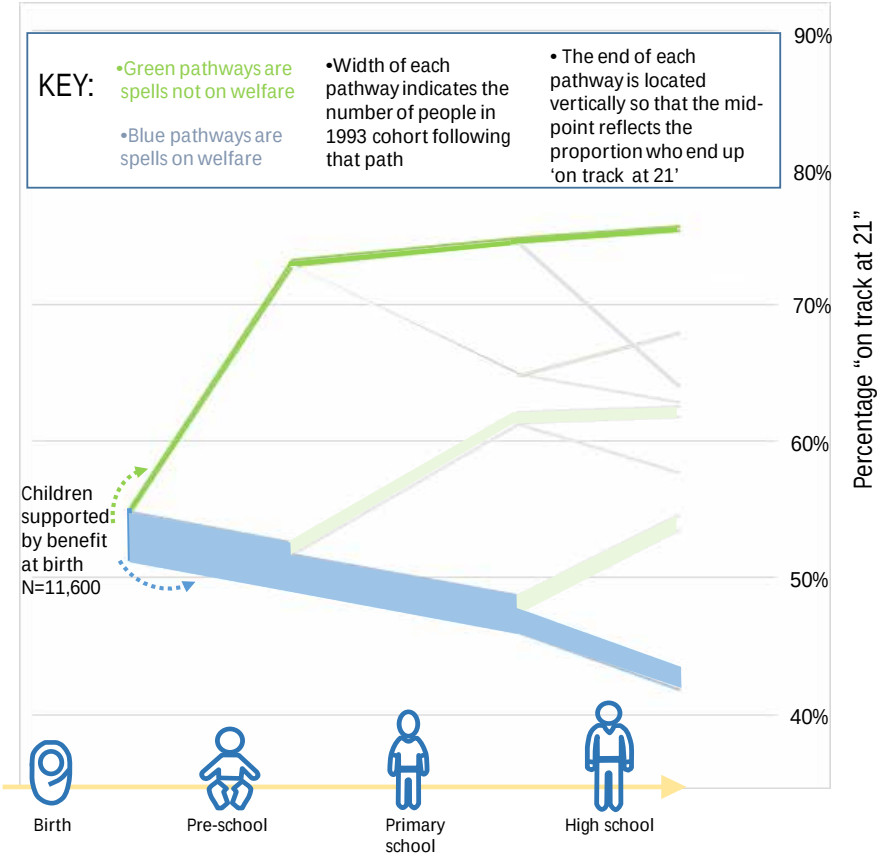
At this point it is important to emphasise that in using welfare support as a "risk factor" we are pointing to its interpretation as a proxy for adverse events that may have led the family to need welfare support from the state. However the intention of welfare support will have been to help buffer the family from the worst impacts of these events and their consequent exposure to periods of very low income. This (presumably) beneficial effect is also reflected in moderating the sizes of the slopes of the branches in this graph from what they would have otherwise been.

Appendix Table 5 contains the data that is used to create the graph.

Appendix Table 5 - Proportion "on track at 21" within each family welfare pathway

	Proportion on track at 21						
		Supported by benefit 25% of time (age 0-4)?	Supported by benefit 25% of time (age 5-12)?	Supported by benefit 25% of time (age 13-17)?		Number on track	Number in cohort
Not supported by benefit at birth	79.1%	No 83.1%	No 84.3%	No	84.8%	27867	32859
				Yes	71.8%	900	1254
			Yes 70.5%	No	74.5%	1392	1869
				Yes	65.3%	942	1443
		Yes 58.7%	No 70.7%	No	71.5%	1401	1959
				Yes	62.7%	126	201
			Yes 53.7%	No	59.3%	1497	2526
				Yes	48.4%	1272	2628
Supported by benefit at birth	53.3%	No 73.0%	No 74.6%	No	75.5%	729	966
				Yes	64.0%	48	75
			Yes 64.8%	No	62.8%	81	129
				Yes	67.9%	57	84
		Yes 50.9%	No 61.6%	No	62.1%	1389	2235
				Yes	57.7%	168	291
			Yes 47.4%	No	54.1%	1755	3246
				Yes	42.6%	1941	4557

Appendix Figure 3: Subset of family welfare diagram - Children supported by benefit at birth



Appendix Figure 3 is an extract from Appendix Figure 2 which illustrates family welfare pathways for the 1993 cohort. Focussing on the group of children supported by benefit at birth, but whose families subsequently have less time supported by benefit (green pathway), we see higher rates of being 'on track at 21' (76%), in fact quite close to the cohort's overall average of 77%. Those children whose families have significant time on benefit consistently through the child's life, have much lower 'on track at 21' rates (43%).

Unpicking the data in this way can help us see the potential of policies that target at different times in children's lives. Always, of course, bearing mind that we are not inferring that the spells on benefit caused poorer outcomes, but rather highlighting the potential of identifying possible groups to target (and when in their lives) for the provision of better social services.

References

Christopher Ball, Sarah Crichton, Robert Templeton and Sarah Tumen (the Treasury), Rissa Ota and Conrad MacCormick (Ministry of Social Development) (2016). Analytical Paper 16/01: *Characteristics of Children at Greater Risk of Poor Outcomes as Adults*. Wellington: the Treasury.

New Zealand Treasury (2016c) He Tirohanga Mokopuna: Statement on the Long-Term Fiscal Position 2016.

Tony Burton, Louise Butler, Melissa Piscetek, Paul Rodway, Caroline Reid and Robert Templeton (2016) “The benefits of improved social sector performance,” the Treasury.