

Training to earnings

A modelling approach to estimate the impact of training on lifetime earnings



Introduction

This document provides an overview of a modelling approach to provide an indication of the impact of training on lifetime earnings. The aim is to build a model that can demonstrate the impact of funding training to learners, employers, industry bodies and government. One potential use of this information is to prioritise what training is funded within the food and fibre sector.

The modelling approach we use in this work is a simple one – essentially a bet estimate cure fit informed by available data.

It is intended that the model will be implemented in as a part of the investment tool developed for the Raising Aspirations project. Here the model is used to evaluate the impact of different training programmes on lifetime earnings.

Overall framing

An overview of the model logic at the top level.

Before we get started...

- The maxim *all models are wrong, some models are useful* should be kept front of mind for this work! We are trying to describe a complex system in a mathematical model. It is necessary to simplify at almost every stage and it will be easy to find flaws with the approach. However, this does not necessarily stop the model being useful.
- In particular, we think the model can be useful (without being right) if it can help us compare:
 - One training option with another
 - Change in business profitability, due to change in training inputs, over time

In both of these cases, imperfections in the model should apply roughly equally to both sides of the comparison.

- While empirical data to calibrate the model are available, these are incomplete (see next slide). We also require the use of judgement to tune the model. In this case this includes the use of ChatGPT to estimate the value of some parameters.
- This model is a starting point and can be refined and improved over time. We include some ideas for future improvement in this document.

Overall approach

- On average, people with higher qualifications earn more over the course of their lifetime than those with lower qualifications. There is a body of research that seeks to quantify this relationship. We use some of this research to help calibrate this model.
- However, the research that we have been able to find does not provide all that we need to forecast the impact of a learner completing their training on lifetime earnings. In particular:
 - Data can link demonstrate the *correlation* between qualifications and income but not the *causality*. Some part of the correlation is likely to be explained by the higher propensity of people with innate ability, and/or prior success in education, to seek higher qualifications.
 - Most research relates the level of qualifications to income but does not distinguish between different training programme duration.
 - We have not found research that covers a full range of education programmes including high school, vocational training and degree-level training.

Overall approach (continued)

- The model we build combines the available data with three key assumptions to address these gaps.

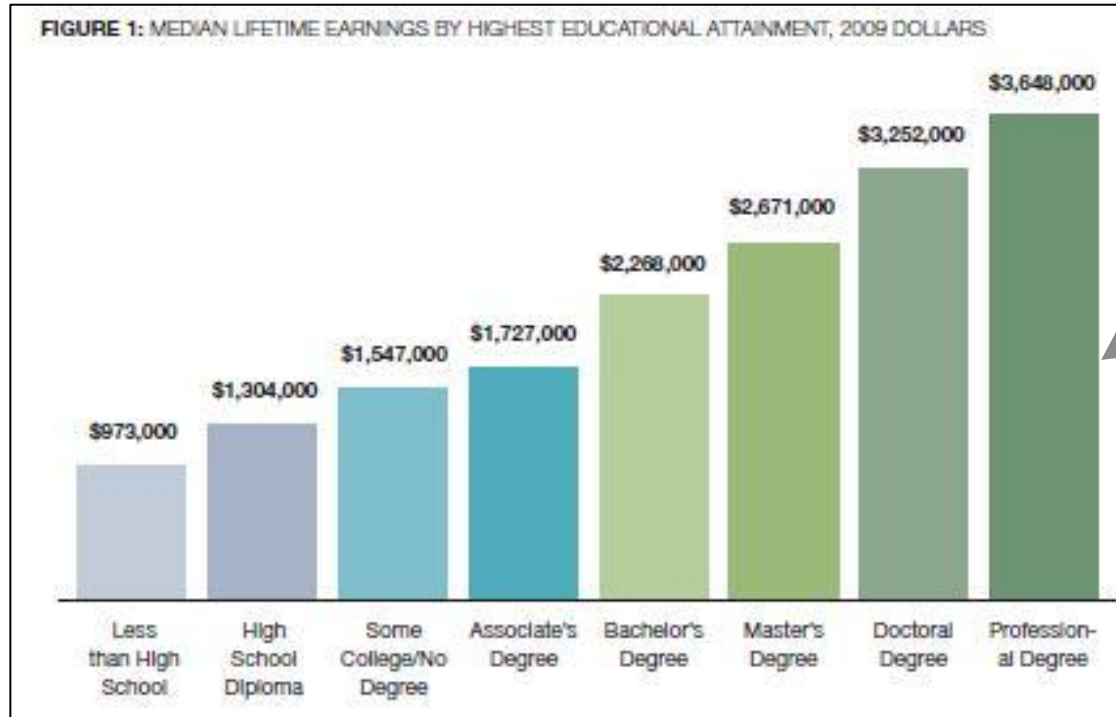
We make an assumption about:

1. The proportion of the qualification-to-income correlation that is causal i.e. extent to which the qualification brings about a higher income (whether through new skills gained or simply a signalling effect of gaining the qualification) c.f. being because people with higher innate ability are likely to both earn more and gain higher qualifications.
2. The returns from longer programmes of learning at any given level. We assume that there are diminishing returns.
3. The range of existing qualifications held by any learner training towards a new qualification.

Research linking qualifications and earnings

An overview of some of the data available to inform this model.

Georgetown University research

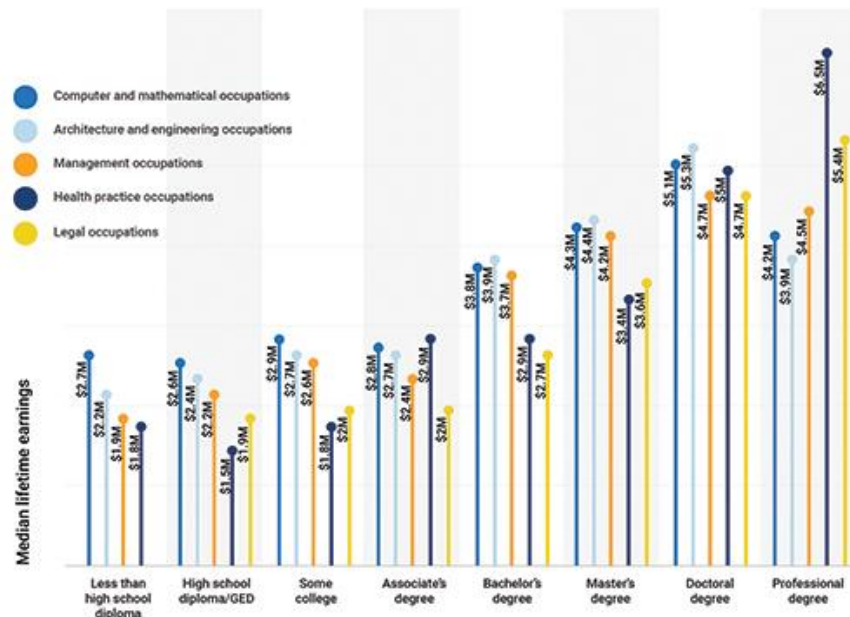


This figure provides a rough overall guide to the impact of higher qualifications on incomes. For example:

- A person with a bachelors degree earns ~2x a person with only high school qualifications
- A person with a professional degree earns ~3x a person with only high school qualifications

Georgetown University research

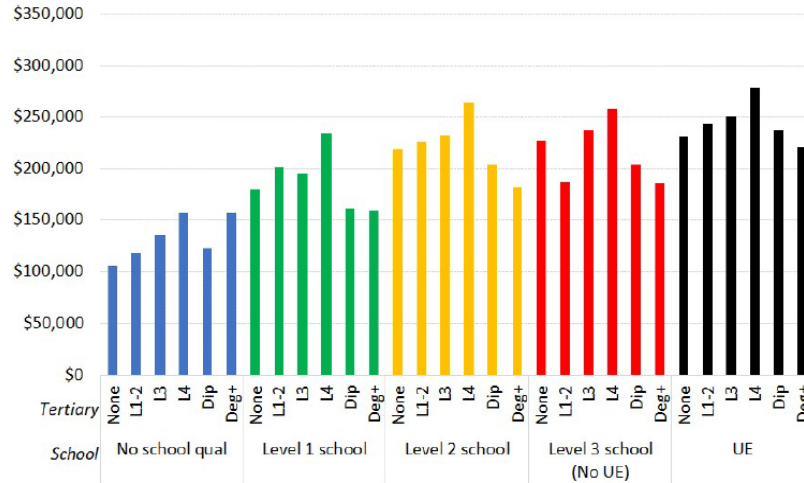
Figure 6. Computer and mathematical, health practice, and architecture and engineering are the top-paying occupations across education levels.



A similar figure dives a little deeper into the differences between occupations. This highlights that there are variables at play other than qualification level (albeit we will not attempt to address them in this simple model)

New Zealand data

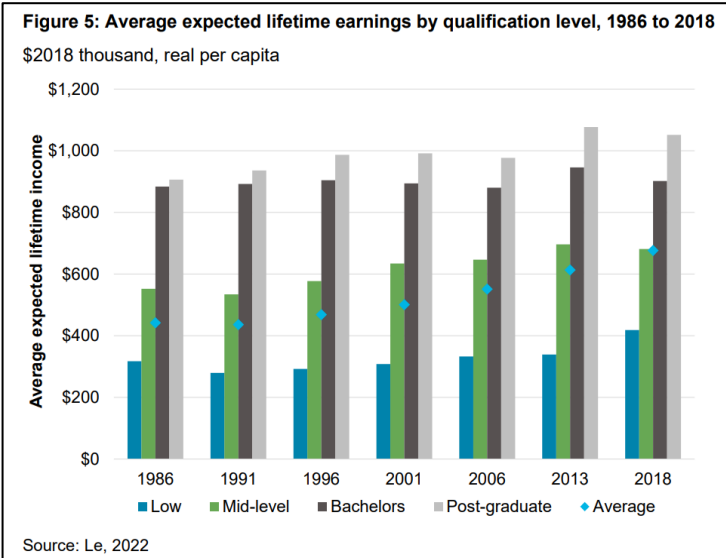
Average cumulative earnings by highest school and highest tertiary attainment nine years after leaving school



Data relates to the earnings in 2018 of people who left school in 2009 and who were living in New Zealand in 2018, whether working or not. Includes qualifications gained in New Zealand either in the workplace, or at tertiary education providers.

Some local by the Ministry of Education hints at the expected effect but is limited because it considers only the first nine years after leaving school rather than a full lifetime of earnings.

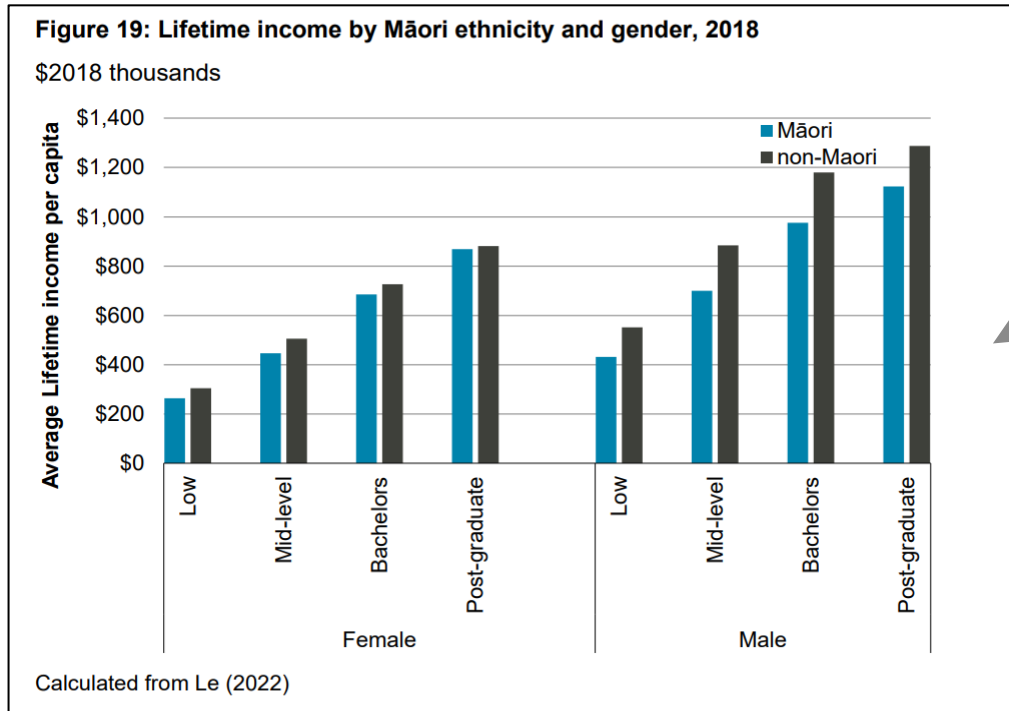
New Zealand data



Work published by the New Zealand Treasury in 2023 confirms a similar pattern to that of the Georgetown University group, that is, people with degrees earning 2-3 times more than those with 'low' qualifications.

Figure 5 shows the expected accumulated lifetime income for each qualification group adjusted for inflation to 2018 dollars. There is a significant difference by qualification level, with bachelors or higher degrees expected to earn at least twice, and in some years nearer three times, the lifetime income of a person with low qualifications. This pattern has been persistent over the last three decades, though there has been some real income growth for all the qualification groups except for those with bachelors' degrees. This growth has been largest and most persistent for those with mid-level and post-graduate qualifications. Despite these changes the difference in earnings between qualification levels remains low in New Zealand compared to many other countries.

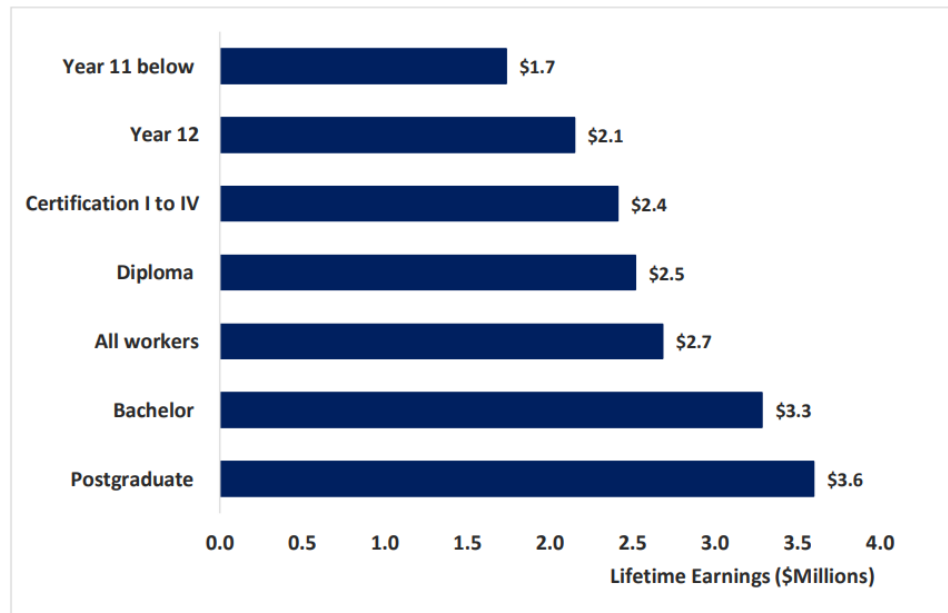
New Zealand data



The same paper confirms the intuition that ethnicity and sex are also contributing variables.

Australian data

Figure 5: Lifetime employee income by educational attainment, ages 25-29, 2021



Source: Author's calculations based on ABS (2021) *2021 Census – employment, income and education*

A table from an Australian paper helps to address a gap in the previous charts for the impact of vocational training

Future improvement ideas

- **Look for additional data** – Look for additional published work and / or undertake new primary research using IDI data. For the food and fibre sector this could include recent work by Gail Pacheco for Muka Tangata.
- **Consider other influences on income** – The model could be extended to consider other influences on income including occupations and demographic traits such as sex and ethnicity.

Qualifications to income model

A sub-model to characterise the
correlation between lifetime earnings
and qualifications

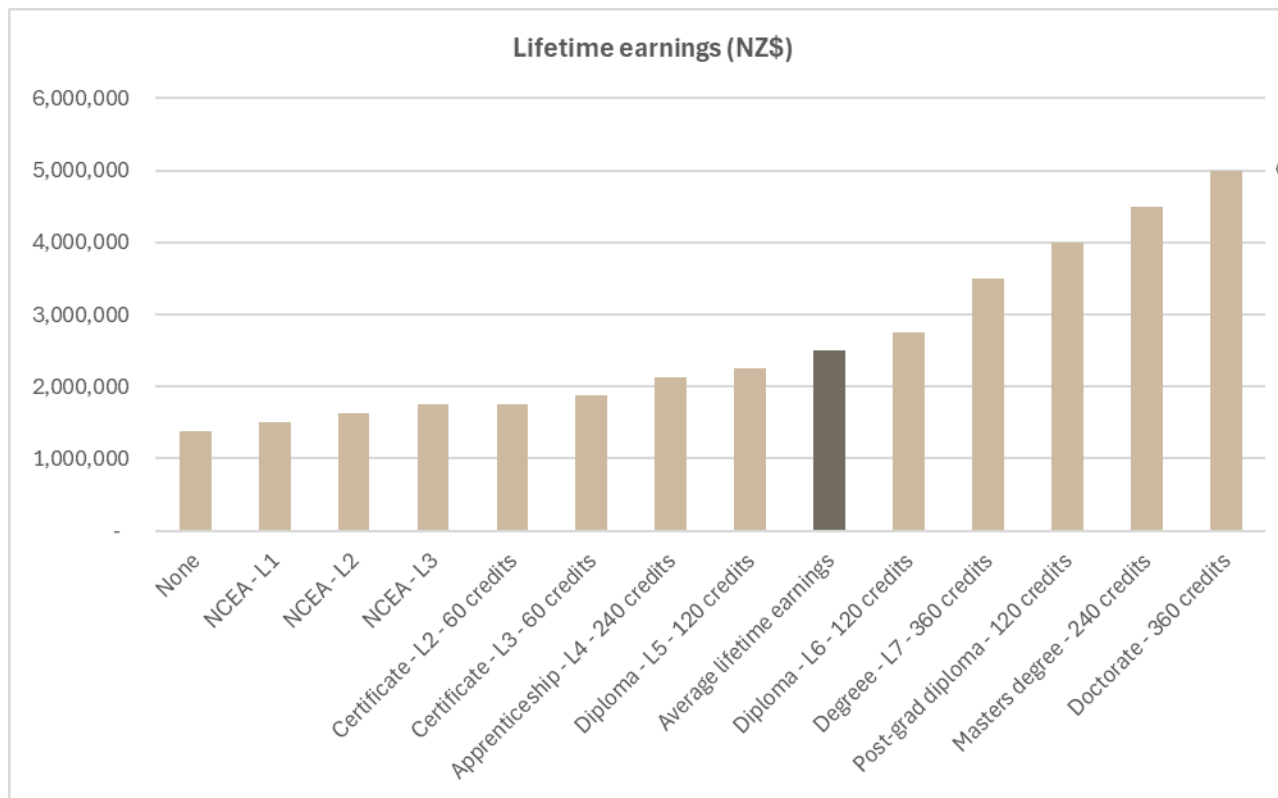
Average lifetime earnings

Method	Estimate
The Treasury 2023 paper – see slide 12 – estimates this at roughly \$700k in present value terms. This seems improbably low?	~\$700,000
A paper by the Taxpayer Union estimates closer to \$3.5m	~\$3,500,000
Assuming that over a career a person: • Works 35 years (e.g. from age 25 to 65, with 5 years off), • Is paid, on average, 48 weeks of the 52 • Works 35 hours per week • Earns the 2025 average wage (\$43.39 per hour)	~\$2,583,000

The baseline, or average, estimate of lifetime earnings is the starting point for the model. Existing estimates of this vary by method and assumption.

We use a value of \$2.5m in the model.

Overview – team capability sub-model



This chart shows the model used to 'fit' to the various data shown in the previous section to a range of benchmark training programmes (as defined by NZQF level and duration).

Correlation vs causality (ChatGPT's view)

There is a lot of research that links the impact of training to lifetime earnings. However, at least some of this correlation is because people with higher innate capability, or a better level of prior education achievement, are both more likely to earn more and more likely undertake further training (selection), rather than because the qualification has a causal effect on income. Where the qualification does have an effect on income it could be because it leads to higher productivity (productivity) or simply because the qualification signals high capability (signalling).

Estimate how much of the relationship between training and income is:

- 1. Causal changes due to gains in productivity*
- 2. Causal changes due to signaling*
- 3. Non-causal due to selection.*



Quick rule-of-thumb (all training, averaged) $\approx$ 45% productivity + 15% signalling + 40% selection

Training type	Causal: productivity	Causal: signalling	Non-causal: selection
Formal qualifications (cert/diploma/degree)	45–60%	10–20%	25–40%
Apprenticeships / long, work-based VET	55–70%	5–10%	20–35%
Short adult/vocational courses (micro-credentials, classroom-only)	25–40%	5–15%	45–65%

We use 50% causal
(whether due to
productivity or signalling)

Impact of training programme duration (ChatGPT's view)

I am interested in the earnings uplift from training. What evidence is there of a diminishing returns from more hours of training of a similar type and level? For example, if I study for a two years masters programme how much more benefit do I get than if I study for one year?

I would like to characterize this using a function with the form:

Earnings uplift(h) = ($h / href$) ^{b} where:

- h is the number of hours of learning

- $href$ is the typical learning hours in a year which in my case is 1200 hours.

Suggest the most suitable value of the exponent b .



Recommended default $b \approx 0.45$

- Implies **diminishing returns** (since $b < 1$).
- With $b = 0.45$:
 - **1 year (600 h)** delivers $(0.5)^{0.45} \approx \mathbf{0.73}$ of the 1200-hour uplift.
 - **2 years (2400 h)** delivers $(2)^{0.45} \approx \mathbf{1.37}$ times the 1200-hour uplift.
⇒ Moving from 1 year to 2 years gives **~37% more** uplift (ignoring signalling), while the **first year already captures ~73%** of the two-year uplift—consistent with concave “dose–response” findings from schooling and adult-training evaluations.

We use 45% causal
(whether due to
productivity or signalling)

Existing qualifications

- The range of existing qualifications held by any learner working towards a new qualification is modelled at the level of each training programme. For example, we assume that most people working towards a PhD have an existing masters degree.
- We consider the benefit of working towards any given qualification as being incremental to the weighted range of existing qualifications.

Future improvement ideas

- **Develop better estimates of the causality effect** – We may be able to better estimate how much of the impact of training is casual c.f. a selection effect. We could also adapt the model to use the different estimates of causality for different types of programmes suggested by ChatGPT.
- **Diminishing returns** – We may be able to find more evidence to justify the assumptions on diminishing returns.
- Do more to split out the signalling effect, productivity effect and the selection effect and how the balance of these changes by course level and course duration.
- Improve the logic of prior qualifications given the different sizes of the reference training programmes.

SCARLATTI

Turning research into action

scarlatti.co.nz