Te Mahau

22/12/2022

James Dulcich
fyi-request-21253-c6534f38@requests.fyi.org.nz

Tēnā koe James

**OIA: 1301098 – Equity Index formulae**

Thank you for your email of 26 November 2022 to the Ministry of Education requesting the following information:

> *I would like to request under the Official Information Act the exact formula used to compute the Equity Index from the 37 data points outlined in the Equity Index Varables Facts Sheet (2022). I have been looking for hours and have been unable to locate it anywhere on the Ministry of Education's website. For each year possible, I would like to see this equation for the years 2019, 2020, 2021, 2022, and 2023.*

Your request has been considered under the Official Information Act 1982 (the Act).

There are two sets of formulae that are required to answer your request in full. We provide separate documentation for the formulae for the 2019 year and 2023 years. The Equity Index model changed between 2019 and 2021, and the 2019 formula responds to your request for 2019 and 2020, whereas the 2023 formula (finalised in 2022 for use in 2023 school funding) responds to your request for 2021 to 2023.

The 2019 formula (covering 2019 and 2020) was previously described in a technical report published in September 2019 on the Education Counts website. The formula is described in Section 2 of the Technical Report available on that page. Please see the following links for this information:

- https://www.educationcounts.govt.nz/publications/schooling/he-whakaaro-accounting-for-educational-disadvantage
- https://www.educationcounts.govt.nz/__data/assets/pdf_file/0003/196005/Equity-Index-Technical-Report-Final.pdf

As this information is publicly available at the above links, I am refusing this aspect of your request under section 18(d) of the Act.

The 2023 formula (covering 2021 to 2023) is described in **Appendix A** (attached). A single set of formulae is provided for this year range, because the overall formula has remained unchanged throughout this development period. The main change to the Index across this period was to validate and include additional variables—nonetheless, the formula described in the technical appendix that we are providing was used irrespective of the basket of socio-economic variables that were included.

Please note that Appendix A is based on a draft Technical Report that is planned for release in 2023, and some details might change between this draft and the final published report.

We also considered whether to include individual model coefficients associated with each factor within the model; however, this information is not held by the Ministry. This is because the Equity Index has been developed within the Statistics New Zealand's Secure Data Laboratory, using the Integrated Data Infrastructure (IDI), and the coefficient or weight of each independent variable within the model has not been approved for release from the IDI. Ordinarily, the Ministry would, therefore, partially transfer this aspect of the request to Statistics New Zealand. However, Statistics New Zealand cannot be considered to 'hold' this information either, as they would need to extract and create this information. For either agency to provide this information it would require the creation of new analytical code to calculate these coefficients for each version of the model and then a process to confidentialise the output to ensure the protection of personal information would need to be undertaken. This work would require complex skill, judgment, analysis, interpretation, and quality assurance of the information, which in-keeping with guidance from the Office of the Ombudsman, would constitute the creation of information. For this reason, I am refusing this information under section 18(g) of the Act, as the information is not held by the Ministry, and I have no grounds to believe it is either held by, or more closely related to the functions of, another department or organisation subject to the Act.

We consider that providing these model coefficients requires complex calculation because of the complexity of the models and because it requires a process to confidentialise output. Each Equity Index is based on 3 scoring years; each scoring year is calculated using 28 Poisson regression trained models, with each containing approximately 175 beta coefficients for the parental and 55 beta coefficients for the non-parental model. In total, this expands to approximately 9,650 coefficients per Equity Index (based on the 2023 model). These coefficients will vary slightly for each of the 5 years that you have requested because the training and scoring data changes each year. Although we are refusing to provide model coefficients as part of this request, we are providing a comprehensive explanation of how the model works and some of the design decisions relating to the development of the Index.

It is important to note the coefficients on their own would not provide any reliable or accurate insight into the model. There are not, strictly speaking, weights involved in creating the Equity Index, as would be the case of arithmetically creating a compound index (e.g. 0.5 to factor 1 and 0.5 to factor 2). The Equity Index is based on a statistical model that examines the relationship between the whole basket of socio-economic factors and educational outcomes. Because all of these factors are related to each other (what statisticians call 'multicollinearity'), it would be incorrect to interpret any parameters in this model as the 'weight' put on a particular factor.

The statistical model is estimated using hundreds of thousands of data points. Experienced researchers have determined the final specification based on a combination of previous research and analysis, an external technical reference group, and stakeholders from across the sector.

Please note, the Ministry now proactively publishes OIA responses on our website. As such, we may publish this response on our website after five working days. Your name and contact details will be removed.

Thank you again for your email. You have the right to ask an Ombudsman to review my decision on your request, in accordance with section 28 of the Act. You can do this by writing to info@ombudsman.parliament.nz or to Office of the Ombudsman, PO Box 10152, Wellington 6143.

Nāku noa, nā

Helen Hurst
**Hautū Taupua | Acting Deputy Secretary**
**Te Mahau | Te Pae Aronui (Operations and Integration)**

MINISTRY OF EDUCATION
TE TĀHUHU O TE MĀTAURANGA

# 2023 EQUITY INDEX MODEL FORMULA

Insights and Analysis, Te Pae Aronui



**Disclaimer**

These results are not official statistics. They have been created for research purposes from the Integrated Data Infrastructure (IDI) which is carefully managed by Stats NZ. For more information about the IDI and please visit https://www.stats.govt.nz/integrated-data/.

The results are based in part on tax data supplied by Inland Revenue to Stats NZ under the Tax Administration Act 1994 for statistical purposes. Any discussion of data limitations or weaknesses is in the context of using the IDI for statistical purposes, and is not related to the data's ability to support Inland Revenue's core operational requirements.

OIA: 1301098

National Office, Mātauranga House, 33 Bowen Street, Wellington 6011
PO Box 1666, Wellington 6140. Phone: +64 4 463 8000 Fax: +64 4 463 8001

The modelling of the Equity Index consists of two main stages (i.e., estimation and scoring). The first stage of the Equity Index Funding model process is undertaken within Statistics NZ's secure data laboratory using anonymised records from the Integrated Data Infrastructure. It establishes the relationship between a student's socio-economic characteristics and their educational achievement. This relationship is represented by a set of parameters for each year of school, which can be used to predict the achievement of current students in Stage 2.

# 1 Stage 1 – Equity Index Model Estimation

## 1.1 Defining the Training Population

The first stage in the development of Equity Index numbers involves establishing the relationship between the socio-economic characteristics of past students at various school ages and their NCEA achievement. This is used to assess the likely impact of socioeconomic barriers on the educational achievement of today's students, based on their characteristics. The first step in this process is to identify the population of students we will use in modelling this relationship, what we call the 'training population'.

The training population is selected based on year of birth, with birth cohorts selected on the basis that sufficient time must have elapsed since birth to observe most people having completed their school education. On the other hand, births should have been sufficiently recent to take advantage of new data developments, to both maximise the chance that the established relationship between socio-economic characteristics and educational achievement has not changed, and to minimise the broader impacts of changes in society or the educational environment. The training population also needs to be big enough to ensure the model works well for all subgroups of students.

In weighing up these considerations, the training population was chosen as the population of children born in the three years ending 20 years before the year which the Equity Index (EQI) is being calculated for. For example, the 2023 EQI model (i.e., the published EQI for Funding Year 2023) training population is based on the population of people born between the start of 1999 and end of 2001. It is worth noting that the training population will step forward in time as the scoring population does. Therefore, for the 2024 EQI model, the training population are people born between 2000 and 2002. A set of additional criteria were also applied to the training population to ensure the students selected were representative of the population of students the EQI funding would apply to, and for whom sufficient socio-economic data should be available.

- The person must be linked to the IDI spine[1], something that is necessary for linking de-identified IDI datasets. Without this, the socio-economic characteristics of the person are not able to be identified.
- The person must have been enrolled in a New Zealand school as at the IDI extract date[2] and have been enrolled as a student at some stage between age 14 and 17, approximately the ages they would have been in school years 11 to 13 and therefore expected to be studying for NCEA.
- They must not have been overseas for more than 180 days while aged 15-17 and must not have been enrolled at a private school in any year during this period.

[1] For more information about the IDI spine, see: https://vhin.co.nz/guides/idi-spine/.

[2] Domestic student status is only able to be determined as at the date the Ministry of Education data was extracted for the IDI refresh being used for training. The IDI is refreshed with updated data periodically by Statistics NZ, generally on a quarterly basis.

- They must not have been enrolled at a school offering qualifications that are not on the New Zealand Qualifications Framework (NZQF)[3] between ages 15 and 19.

## 1.2 Model structure

As discussed above, the first stage in calculating Equity Index numbers involves establishing a relationship between indicators of socio-economic barriers, identified at different ages, and later educational achievement. This is done through a regression model. In early iterations of the Equity Index (Ministry of Education, 2019a), and in the Treasury work that preceded it (Crichton, Dixon, & Ball, 2018) (Crichton, Dixon, & Ball, Unpublished), a logistic regression model was utilised. At that time the outcome variable of interest was a simple indicator of NCEA Level 2 achievement, described by a binary variable, set to one if a student achieved NCEA level 2, and zero if they did not.

Since the specification of these early models, a new and more nuanced measure of achievement was adopted. This measure involves adding up the number of credits a student has achieved across NCEA levels 1 and 2, weighted by both whether the achievement was a unit standard or an achievement standard, and by the level of endorsement achieved. With the change from a binary measure a logistic regression model was no longer appropriate.

A Poisson model was adopted for the Equity Index.[4] Poisson models are commonly used where the dependent variable is a 'count' variable, for example in modelling the number of discrete events that occur. In our case the dependent variable for educational achievement described below is not, strictly speaking, a count variable, however it is not unreasonable to view it as being analogous to a count.

The model was therefore specified as:

$$y_i = e^{(\alpha_a + \beta_{1a}x_{1ia} + \beta_{2a}x_{2ia} + \cdots + \beta_{pa}x_{pia} + \varepsilon_{ia})}$$

Where $y_i$ represents the educational achievement of student i and $x_{1ia}$, $x_{2ia}$, … $x_{pia}$ represent measures of the level of socio-economic disadvantage experienced by the student at age *a*. The model assumes that the achievement measure $y_i$ is a random variable from a Poisson distribution with equal mean and variance. The parameters $\beta_{1a}$, …, $\beta_{pa}$ represent the estimated relationship between the $p$ socio-economic indicators and educational achievement at age $a$.

The model is run separately for each year of age a person would normally be at school, i.e., $a = 5, 6, \ldots, 17, 18$.

## 1.3 Measuring educational achievement

Unlike earlier iterations of the Equity Index model, which used a simple indicator of whether a student attained NCEA Level 2 as the dependent variable measuring educational achievement, the current iteration of the model uses a more complex mix of NCEA level 1 and 2 results, which takes into account the number of credits achieved, the level at which

[3] Schools are considered to be offering non-NZQF qualifications if more than ten percent of qualifications gained by students are not on the New Zealand Qualifications Framework.

[4] We also tested the use of a negative binomial model due to concerns about over-dispersion in the underlying achievement distribution. Over-dispersion occurs when the Poisson equidispersion assumption, that the distribution of the dependent variable is equal to its variance, is violated. While over-dispersion would not affect our model parameter estimates or predictions, it could impact on the estimated standard errors and measures of fit. The Poisson model was found to provide better predictive accuracy and stability over time, however, and was preferred for the final calculation of the Equity Index.

they are achieved, the credit endorsement (achieved, merit or excellence), and whether the credit was a unit or achievement standard.

### 1.3.1 Measure criteria

Following the development of the initial Equity Index model, technical reviewer feedback was received suggesting the use of the simple NCEA Level 2 indicator be reviewed. Similarly, the Ministry's Sector Reference Group, made up of 30 sector representatives and principals, also raised this as an area for improvement in the model. Concern centred around the idea that using a binary measure of achievement on NCEA Level 2 (i.e., pass/not pass) as the dependent variable did not reflect the granularity of NCEA achievement. With high and rising NCEA Level 2 achievement rates over time, there was a further concern the indicator would lose power as a predictor of future success over time.

The Ministry has continued to explore alternative options since the first Equity Index model was developed. These options focussed on answering a number of questions.

- Which levels of NCEA should be included in the new measure?
- How should different endorsement levels be treated?
- Should all subjects be weighted equally, or should more prominence be given to certain subjects, such as English and Maths?
- Should standards be weighted differently if they are more difficult (e.g., according to their pass rate)?
- Should unit standards be treated differently to achievement standards?

In answering these questions, we sought to balance several factors. Firstly, the measure should have a sound theoretical basis, preferably with some precedence in its use. Examples of measures meeting this requirement are the Weighted Relative Performance Index (Crampton & Udahemuka, 2018) and the Expected Percentile (Ussher, 2008).[5] Rank scores, based on a student's best credits at NCEA Level 3, also have precedence in the education sector, being used by many universities for entry to some courses.[6]

Secondly, the measure should provide a sensible measure of educational success, that reflects both the quantity and quality of NCEA credits achieved in a way that reflects the diversity of student achievement and provides an equitable representation of students.

In order to meet these requirements, the following criteria were used to assess candidate measures.

- The measure should incorporate some indicator of the quality of the credits achieved by a student.
- The measure should incorporate achievement and unit standards, in order to reflect a wide range of course choices.
- The measure should use a combination of Level 1 and 2 credits, but should exclude Level 3 credits. Participation is much higher at Levels 1 and 2 than at Level 3, giving good coverage of the student population at these levels. In addition, Level 2 is considered to be a 'minimum' qualification to enable further education, training or employment.

[5] Unpublished.

[6] See https://www.nzqa.govt.nz/assets/About-us/Publications/Brochures/NCEA-Factsheet-5-July-2017-FINAL.pdf. Rank scores are often referred to by universities as Preferential Entry Scores or Guaranteed Entry Scores.

### 1.3.2 Selecting the measure to use

In addition to meeting the criteria above, the selected measure should also perform well in a modelling context. In total, 70 measures were tested. This was made up of seven measures representing all combinations of NCEA levels[7] for each of the following:

- Eight measures constructed as the sum of achievement standard credits only, or both unit and achievement standards; and unweighted, weighted by score, weighted by expected percentile (EXP), and weighted by both score and EXP.
- Two additional measures based on the Weighted Relative Performance Index (WRPI) and Average EXP.

While some of these models did not meet all three criteria discussed above, it was considered important to compare these against the models which did meet the criteria. This ensures that the impact of applying the criteria on measure performance is well understood.

Each measure was tested to see the degree to which it could be predicted by the socioeconomic factors selected for the Equity Index model, discussed in Section 1.4. This was assessed by running each of the candidate measures in a Poisson regression model and measuring the model goodness-of-fit. Candidates were selected based on their performance across multiple goodness-of-fit measures.[8]

The two measures that performed consistently better than others were the sum of NCEA achievement credits, and the sum of NCEA score-weighted achievement credits. For both measures, using NCEA levels 1,2 and 3 usually provided the best performance, followed by NCEA levels 1 and 2. Including unit standard credits generally resulted in a loss of model performance, however the score-weighted unit standard and achievement standard measure performed almost as well as the score-weighted achievement standard measure, and has the advantage of meeting all three criteria set out in 1.3.1 above.

The final measure selected for the Equity Index was the sum of NCEA level 1 and 2 endorsement-weighted achievement and unit standard credits.

### 1.3.3 Optimising the weightings

Given the decision to use level 1 and 2 score-weighted credits, a set of weights needed to be selected. The option used in the measure selection in Section 1.3.1above was to use the weighting set adopted by some NZ universities in their calculation of ‘rank scores’ for university entrance.[9] This weighting scheme assigns weights of 2, 3 and 4 for achievement standards with achieved, merit and excellence levels of achievement respectively. Different universities treat unit standards differently, however. While some do not give any weight to unit standards in calculating rank scores, others accept a list of approved unit standards.

We tested whether there was a combination of weights that would perform better than the standard rank score weightings, in a similar way to that applied in the selection of measures. While a preference was given to the standard 2, 3, 4 weighting due to the precedence set by universities, we wanted to assess how well alternative weightings would work relative to this set. Measures were constructed for dependent variable $D_c$ as:

$$D_c = W_U N_U + W_A N_A + W_M N_M + W_E N_E$$

---

[7] Level 1, Level 2, Level 3, Levels 1&2, Levels 1&3, Levels 2&3, and Levels 1&2&3.

[8] Measures used were pseudo-$R^2$, mean-normalised root mean squared deviation, IQR-normalised root mean square deviation, Spearman’s coefficient, and to a lesser degree, Pearson’s coefficient.

[9] Unit standards were weighted in the same way as achievement standards, however most unit standards only have achieved and not achieved grades.

Where $W_I$ is the weight for credit achievement level $I$, where $I$ is $U$, $A$, $M$ or $E$ (unit, achieved, merit and excellence, respectively) and $N_I$ is the number of credits at that achievement level. Weights are set as integers from 0 to 10, and all combinations of these weights were tested for $W_U, W_A, W_M$, and $W_E$. This gave a total of 14,641 ($11^4$) different combinations of weights and hence dependent variable weightings. Once we exclude duplicate weight combinations[10] and the case of all weights being set to zero, we are left with 13,025 unique sets of weights.

We ran regressions for each of these sets of weights, and calculated goodness-of-fit statistics as in Section 1.3.2above. We then estimated the local maxima of different regression performance measures on this 11 x 11 x 11 x 11 space. The results of this analysis show that the maximum regression performance is achieved at around the weighting set $W_U$ = 0.45, $W_A$ = 2.05, $W_M$ = 3.35, and $W_E$ = 4.15.

This optimal weighting was simplified to $W_U$ = 1, $W_A$ = 4, $W_M$ = 6, and $W_E$ = 8 in the final Equity Index model.[11] This was almost identical to the rank score weighting set we started with, with the exception that unit standards were only given a weight that was a quarter of the weight given to the achieved achievement standards. Apart from being a weighting set that performs well in a modelling context, the choice of a smaller weight for unit standards is justifiable on the basis that unit standards are disproportionately represented among non-university entrance subjects, a distinction we don't otherwise make in the construction of our dependent variable. These subjects do not contribute to rank scores used by universities.

## 1.4 Indicators of socio-economic status

The intent of Equity Index funding, as with the Decile funding system that preceded it, is to provide schools with additional funding to help provide more equitable outcomes for students facing greater socio-economic barriers to educational achievement. In order to do this, it is important to choose a set of measures of socio-economic status which meet two key criteria.

Firstly, measures should have a strong theoretical basis as a measure of socio-economic barriers or should be a good proxy for it. An example of the former is income, which gives direct information about the financial resources available to a family. On the other hand, the model also includes other proxies for socio-economic status such as unstructured school changes. Although it does not provide a direct measure of socio-economic status, it is acknowledged that students who experience a higher amount of school transience are more likely to experience poorer educational outcomes (Webber & Loader, 2020; Dixon, 2018; Hutchings, et al., 2013). Hence, it may provide an important proxy indicator of socio-economic barriers.

Secondly, measures should be related to educational achievement. The extent that different measures predict educational achievement will help determine the level of funding that students with those characteristics will contribute to the school. If a measure is unrelated to educational achievement, it should not attract Equity Index funding for the school.

In general, the regression model should be parsimonious. This means we want to include the minimum number of variables to achieve a good fit with the data. Adding additional variables beyond this point will tend to improve the fit of the model but runs the risk that the model will not perform well at predicting when applied to other data. For example, the model

---

[10] For example, the set of weights $W_U$=1, $W_A$=2, $W_M$=3, and $W_E$=4 is the same as the set $W_U$=2, $W_A$=4, $W_M$=6, and $W_E$=8 in practice.

[11] Note that this is equivalent to the set $W_U$=0.5, $W_A$=2, $W_M$=3, and $W_E$=4, and is approximately the same as the empirically determined optimal set.

might predict achievement well for the birth cohorts used in constructing the model but predict achievement of current school-age cohorts relatively poorly.

In some cases, a measure might meet one of these criteria well, but fail to meet the other, hence they are excluded from the model. For example, sex has been shown to be a significant predictor of educational achievement, with girls performing better than boys on average, however sex is not an indicator of socio-economic status, nor does it provide a proxy for underlying socio-economic barriers.

We define a socio-economic indicator as a measure that represents an individual's social and economic standing in society, often described by an individual's achievements in education; employment and occupational status; or income and wealth (Winkleby, Jatulis, Frank, & Fortmann, 1992). Studies of children's educational achievement have shown that social background remains one of the major sources of educational inequality (Ainley, Graetz, Long, & Batten, 1995).

### 1.4.1 Selected indicators

A range of approaches were used to select variables for inclusion. Candidates were identified based on both practical and theoretical considerations. A starting point for variable selection was the set of variables used in the early development of the Equity Index model developed by Crichton et al (2018) for the Treasury. Apart from meeting the criteria discussed above, measures had to have robust data available in the IDI that was comprehensively available for children born in the late 1990s throughout their school years.

Apart from variables derived from administrative data, the Treasury model also tested the use of variables derived from the 2013 Census. The Census variables were not found to add much predictive power to the model. Furthermore, Census variables are only updated quinquennially. As such we focus on administrative variables in our model, only using Census data to help fill specific administrative data gaps in areas where frequent updates are unnecessary, such as in identifying parents or caregivers,[12] and measuring parental education.

The model includes the following factors that have been established in the literature as being representative of a student's socio-economic barriers and associated with educational success. These are summarised briefly below and described in more detail in Table 1.

- **Parental socio-economic indicators:** Studies have shown that educational success depends very strongly on the socio-economic status of a student's parents (Ainley, Graetz, Long, & Batten, 1995). We include measures of parental education, parental income and benefit receipt, sole parent status, the age of parents at the student's birth, the number of older siblings and parental contact with the justice system.
- **Child socio-economic indicators:** Children who have experienced poverty, abuse or neglect are more at risk of poor educational achievement. Variables capturing Care and Protection and Youth Justice contact with Oranga Tamariki, and time being supported by benefit are included in the model.
- **National background:** The national background or immigration status of parents has also been found to be an important mediating variable on the effect of SES on children's educational achievement (Portes & MacLeod, 1996). Variables used to capture a student's national background and immigration status are country of birth, migrant category, time spent overseas, age at first arrival in New Zealand, and ethnicity.

[12] A derived Stats NZ table is used to identify parents and caregivers. This is based on a mix of Census data and data from other sources.

- **Transience**: Research suggests that students who move home or school frequently are more likely to underperform in formal education when compared with students that have a more stable school life (Webber & Loader, 2020; Dixon, 2018; Hutchings, et al., 2013). Variables have been constructed from IDI administrative data estimating the number of home and school changes, occurring both separately and concurrently.

**Table 1** Socio-economic variables used in the equity index model

| Type of measure | Measure | Categories | Parental Model | Non-Parental Model |
|---|---|---|---|---|
| Parental socio-economic indicators | Father's/mother's wage and salary income[13] | 9/8 categories* | ✓ | × |
| | Father's/mother's self-employment income | 8/7 categories* | ✓ | × |
| | Father's/mother's main benefit income[14] | 7/8 categories* | ✓ | × |
| | Father's/mother's second tier benefit income[15] | 6/7 categories* | ✓ | × |
| | Father's/mother's highest qualification level | No qualification, NQF level 1,2,3, …, 9, 10, Missing qualification | ✓ | × |
| | Father's/mother's age at child's birth | Under 18, 18-19, 20-24, 25-29, 30+, Missing | ✓ | × |
| | Mother's age at birth of first child | | ✓ | × |
| | Mother's number of children at child's birth | 0-1, 2-3, 4-5, 6+, Missing | ✓ | × |
| | Father/mother with a community sentence | Yes/No/Missing | ✓ | × |
| | Father/mother with a custodial sentence | Yes/No/Missing | ✓ | × |
| | Father/mother with a proven charge | Yes/No/Missing | ✓ | × |
| Child socio-economic indicators | Care and protection family group conference | Yes/No | ✓ | ✓ |
| | Youth justice family group conference | Yes/No | ✓ | ✓ |
| | Care and protection investigation | Yes/No | ✓ | ✓ |
| | Youth justice investigation | Yes/No | ✓ | ✓ |
| | Care and protection notification | Yes/No | ✓ | ✓ |
| | Youth justice notification | Yes/No | ✓ | ✓ |
| | Care and protection placement | Yes/No | ✓ | ✓ |
| | Youth justice placement | Yes/No | ✓ | ✓ |
| | Proportion of lifetime supported by benefit | 7 categories* | ✓ | ✓ |
| National background | Ethnicity | Māori, Pacific, Asian, MELAA,[16] other, European, missing[17] | ✓ | ✓ |
| | Age at visa approval | Less than 5, 5-11, 12-13, 14+, Missing | × | ✓ |
| | Region of birth | Europe excl. U.K.; Latin America and the Caribbean; Mainland South-East Asia; Maritime South-East Asia; Middle East and North Africa; New Zealand; North-East | × | ✓ |

[13] All income variables are measured as average annual income over the child's lifetime.
[14] Working-aged income-tested benefits such as Job-Seeker Support, Supported Living Payment, and Sole Parent Support.
[15] Second tier benefits include accommodation supplement, disability allowance, and family tax credit. Some payments are only available to people receiving a main benefit, while others are available to a wider population.
[16] Middle Eastern, Latin American, or African.
[17] Ethnic groups are prioritised in the order given. For example, if a person identifies as Māori and European, they are recorded as Māori.

| | | | | |
|---|---|---|---|---|
| | | and Central Asia; Northern America; Polynesia (excl. Hawaii), Melanesia and Micronesia; Southern Asia; Sub-Saharan Africa; United Kingdom; Missing | | |
| | Migrant status | NZ-born citizen, Skilled/Business, Family, Refugee, International Categories, Returning resident,[18] Temporary, Non-visa[19] | × | ✓ |
| | Proportion of lifetime spent overseas | 4 categories* | ✓ | ✓ |
| Transience | Number of home changes | 0, 1, 2, 3, …,19, 20+ | ✓ | ✓ |
| | Number of concurrent home & school changes | 0, 1, 2, 3, …, 9, 10+ | ✓ | ✓ |
| | Number of non-structural school changes | 0, 1, 2, 3, …, 9, 10+ | ✓ | ✓ |

* The number of categories is as at the 2023 Equity Index and will be reviewed periodically.

### 1.4.2 Identification of parental relationships

In the calculation of the Equity Index, parental relationships are identified through birth records from the Department of Internal Affairs (DIA), when available. Where there is no link to a birth record in the IDI, as is the case for most migrants, we do not have the full range of parental factors populated. This missing information is problematic for our modelling, as without any correction, it would be likely to result in biased results for students from migrant backgrounds. For this reason, we run two separate models at each of year of age, a 'parental model' using data for students who we have a parent or parents identified in the birth record data, and a 'non-parental model' using data for students for whom a birth parent is not identified.

This also partly addresses another issue affecting students from migrant backgrounds. The experiences of migrant children, as measured in the administrative data in the IDI, will be missing to varying degrees, depending on the amount of their life they have lived in New Zealand.[20] This means that we are limited in the information we have available to us in the model for those students, which might cause an unintentional bias on model performance if not accounted for. If we included migrants in the same model as non-migrants, careful adjustment would need to be made for time spent in the country at different ages. Table 1 identifies which variables are included in the parental and non-parental models.

### 1.4.3 Categorising variables

Following the selection of variables, the next step is to consider the way variables are categorised. Variables are collected in different ways and might be continuous, such as income, or categorical, such as country of birth.

#### *1.4.3.1 Continuous variables*

In the case of continuous variables, a decision was required about whether to retain the variables as continuous or transform them into categories. The former approach has a number of advantages, doing as little to the raw data as possible and retaining the ability to distinguish between small differences in the measure. However, it relies on identifying a

[18] If residence was approved prior to 1997 we are not able to identify the residence category but may identify a person as holding a returning residence visa.
[19] People without a visa who were not born in New Zealand may be the children of New Zealanders born overseas, Australians, or other visa holders with a visa issued prior to 1997.
[20] This issue will also affect some non-migrant children, but to a far lesser extent.

functional form of the relationship between the measure and educational achievement that is consistent across the distribution of the measure and across time.

The latter approach requires us to form sensible categories or *bins* in which to divide the continuous distribution. This introduces a trade-off between using categories that are detailed enough to be sufficiently predictive but are not so detailed that inferences are drawn from random variation in the data. Some advantages of this approach are that missing data can be treated as a separate category without the need for imputation. Moreover, this approach has the potential to accommodate meaningful breaks in the data, such as with school ages or tax brackets, to extract maximum information around these cut-offs.

On the other hand, the model might be sensitive to the choice of "binning boundaries". Differences in variable distributions between the training data and the scoring data is an inherent problem in the Equity Index model but could be even more concerning with categorical binning if effect sizes differ greatly between neighbouring bins. Finally, small cells may be of concern with categorical binning.

Given that the focus is on producing a predictive model to estimate the impact of socio-economic barriers on educational success for current students based on outcomes associated with socio-economic barriers for past students, that is as automated as possible. Based on this requirement we transformed all continuous variables into categorical variables, with bins determined either to reflect important boundaries within the data (e.g. income tax brackets) or through an automatic data-driven process (described below). This could be re-assessed in future work.

Categorisation of continuous variables was done using a mix of theory-driven and data-driven approaches. For example, it made sense to look at the age migrant children received their first visa in relation to school years, as this could impact on transitions between schools.

Where there was no obvious theoretical basis for choosing categories, a data-driven approach was adopted. This approach involved identifying $n-1$ breakpoints, resulting in the creation of $n$ categories. Each breakpoint had to conform to two rules: it must be bigger than the previous breakpoint, and the category resulting from the breakpoint must have a pre-defined number of students ($x$) falling within it.[21]

The procedure to find optimal breakpoints is as follows:

1. Set $n = 2$, giving 2 categories, based on 1 $(n-1)$ breakpoint.
2. Create a random breakpoint between the minimum and maximum values of the continuous variable. Repeat this 100,000 times, testing whether each set of breakpoints has at least $x$ students in each group.
3. A Poisson regression is run using each set of breakpoints, and the model performance is assessed using the same goodness-of-fit statistics discussed in Section 1.3.2.
4. The 100 breakpoints with the best model performance are selected, and each breakpoint from 1 to $n-1$ is set taking the median value of this set.
5. If the increase in performance since the previous iteration is less than 0.01%, we stop, and these breakpoints are adopted for our modelling. Otherwise, we increase $n$ by 1 and return to step 2.

To ensure the increase in performance was not the result of over-fitting, the model was run on the cohort of students born in 1998 and tested against the 1999 and 2000 birth cohorts.

[21] The current model uses a threshold of x=100 students.

The expected regression performance loss was no greater than 1 percent in any case, indicating that the categories fitted well across years.

#### *1.4.3.2 Categorical variables*

In the case of categorical variables (both ordinal, such as qualification, and nominal, such as country of birth), categories are determined based on meeting a minimum threshold[22] for the number of students in each category in each model. The process begins with the most detailed classification of categories, as shown in Table 1. Small categories are then combined with adjacent categories in an iterative fashion. Exceptions to this rule are that ethnic groups are only merged with the 'other ethnicity' category, not with individual ethnic groups; no categories are merged with a 'missing' category; temporary migration categories are only merged with other temporary categories; and permanent migration categories are only merged with other permanent categories. In addition, there are specific rules for country of birth, whereby:

- NZ, Australia and missing can be combined.
- Central Asia and North-East Asia can be combined.
- Europe excluding UK and UK can be combined.
- Latin America and the Caribbean, and North America can be combined.
- Mainland South-East Asia and Maritime South-East Asia can be combined.
- Melanesia, Micronesia, and Polynesia (excludes Hawaii) can be combined.
- North Africa and the Middle East, and Sub-Saharan Africa can be combined.
- Southern Asia stands by itself.

Once these prioritised mergers are completed, if the threshold criterion is still not met for a particular category, that category will be merged with the category with the smallest count.

# 2 Stage 2 ‑ Scoring students and schools

As with the first stage, Stage 2 is undertaken in the secure data laboratory environment, using the IDI. The model parameters from Stage 1 are used as inputs in the prediction of educational achievement for students currently enrolled in New Zealand schools. These predictions, or the de-identified student-level equity index endorsement weighted credits are then averaged at the school-level to obtain an Equity Index for schools. The results are confidentialised in accordance with Statistics New Zealand's confidentiality rules to ensure no students can be identified or their information revealed. The final index is then checked by Statistics New Zealand before being released to the Ministry of Education from the secure data environment.

## 2.1 Defining the scoring population

While the training population is constructed based on a cohort of people born over a three-year period, and observed over subsequent years, the scoring population consists of students recently enrolled in New Zealand schools. The scoring population for a particular year is defined according to the criteria that the person must be enrolled as a domestic student for more than one day at some point in July of that year and that they must be on the IDI spine. Using the example of EQI 2023, the training models of 1999, 2000, and 2001 cohorts are applied in the scoring cohort of 2019, 2020, and 2021, respectively. Specifically, in the training stage (i.e., Stage 1), we identify the relationships between socio-economic factors (see Section 1.4) and educational outcomes (see Section 1.3) and this information

[22] As with continuous variables, this is currently set at 100 students.

(i.e., model parameters) is then applied to the schooling population to estimate educational outcomes based on socio-economic factors for recent learners. The student-level numbers are then grouped and averaged to form the basis of the Equity Index at the school level.

## 2.2 Calculating Student-level Equity Index Numbers

Student-level equity index numbers are predicted educational achievement calculated for each student in the scoring population using their socio-economic factors as identified in the IDI. Predictions are made using the parameters estimated from the training population, using the models described in Section 1. The predicted achievement of student j ($\hat{y}_j$) of age *a* is calculated using the following equation.

$$\hat{y}_j = e^{(\alpha_a + \beta_{1a}x_{1j} + \beta_{2a}x_{2j} + \cdots + \beta_{pa}x_{pj})}$$

The parameters $\alpha_a$, $\beta_{1a}$, …, $\beta_{pa}$ come from the Poisson model run on the training population for age $a$, and the variables $x_{1j}$, …, $x_{pj}$ are the $p$ measured socio-economic characteristics of student $j$.

Student-level equity index numbers represent the predicted weighted NCEA level 1 and 2 achievement of the student, based solely on the socio-economic factors measured for the student and their parents. While these numbers could be used to directly calculate school scores, and therefore school funding, NCEA achievement has been shown to shift over time, and it is desirable that equity funding does not change due to these secular shifts. As such, each year's student equity scores are normalised to a 0 to 1 range before school scores are calculated. The direction of the scale is also reversed, with higher values of the normalised score representing higher levels of socioeconomic barriers.

## 2.3 Calculating School Equity Index

In earlier iterations of the Equity Index, and in the Treasury work that preceded it, students were considered as being 'socio-economically disadvantaged' if they had a predicted achievement (in this case simply the likelihood of achieving NCEA Level 2) below a specified threshold.[23] School scores were then calculated as the percentage of students who were predicted to be disadvantaged based on the threshold. The Treasury report also looked at taking the mean predicted risk of not achieving NCEA Level 2 across each school's students as a school-based measure and compared this with the measure based on percentage of disadvantage. While both measures had a similar distribution, the disadvantage-based measure had a closer relationship to school Deciles.

In moving to a continuous measure of achievement, the decision was made to also calculate school scores based on average risk, or in this case mean predicted achievement across NCEA Levels 1 and 2, as was initially looked at by The Treasury. This change was made to avoid using an arbitrary cut-off point which failed to take into account the circumstances of all students at a school. It also moves funding away from a risk-focussed approach to more of an overall equity focus.

The implication of this is that schools with a large proportion of moderately disadvantaged students would now be funded similarly to a school with a higher percentage of more heavily disadvantaged students, but also a higher proportion of more advantaged students. The school equity scores therefore represent the average of the predicted weighted NCEA level

[23] In the case of the earlier EQI model (Ministry of Education, 2019a) and the Treasury report (The Treasury, Predicting Students' Future NCEA Attainment Using Census and Administrative Data, 2018), the 25 percent of students with the highest risk of not achieving NCEA level 2 were considered to be disadvantaged.

1 and 2 achievement of the school's students, based on the basket of socio-economic barriers faced by students and their families, adjusted for changes in year-to-year student achievement.

## 2.4 Converting School Equity Scores to the Equity Index

Each scoring year has a different cohort of students resulting in different student-level and school-level index distributions. Further standardisation and scaling are conducted to ensure that the final Equity Index distribution will have similar distributions across the three scoring years, and consequently ensure that the Equity Index is as stable as possible on a year-to-year basis. The procedure is described as followed:

Before being converted into the Equity Index, raw school equity scores are standardised such that:

1. The equity scores should have the same mean across schools (weighted by school roll) each year i.e.

$$\bar{m} = constant \tag{1}$$

2. The distribution of schools should have the same spread, or standard deviation i.e.

$$\sigma = constant \tag{2}$$

Where $\bar{m}$ is the weighted mean of equity scores for a given year, and $\sigma$ is the standard deviation of the scores. While it is straight-forward to scale a distribution to a given mean, scaling a distribution to a given weighted mean is not as straight-forward. Firstly, we need to work out the mean and standard deviation and to scale the distribution in order for the weighted mean to be constant. The following formulae illustrate how this is done.

By expanding equation (1) we have:

$$\frac{\sum_{i=1}^{N} I_i c_i}{\sum_{i=1}^{N} c_i} = \frac{\sum_{i=1}^{N}\left[m + (\hat{I}_i - \hat{m})(\sigma/\hat{\sigma})\right]}{\sum_{i=1}^{N} c_i}$$
$$= constant \tag{3}$$

Where $I_i$ and $c_i$ denote the scaled index value and the roll of school I, and N represents the number of schools. The expression in square brackets on the right-hand side re-writes the scaled index $I_i$ in terms of the unscaled index values, $\hat{I}_i$, the mean m and standard deviation $\sigma$ of the scaled index and mean $\hat{m}$ and standard deviation $\hat{\sigma}$ of the unscaled index.

If we set the constant values as $\sigma$=1 and $\bar{m}$=0, we can solve equation (3) for *m*.

$$\sum_{i=1}^{N}\left[m + (\hat{I}_i - \hat{m})(\sigma/\hat{\sigma})\right] \times c_i = 0$$
$$\rightarrow m = \frac{1}{\sigma}\left(\hat{m} - \overline{\hat{m}}\right) \tag{4}$$

Where in the last implication $\sigma$=1 and $\overline{\hat{m}}$ is the weighted mean of the unscaled index.

$$\overline{\hat{m}} = \frac{\sum_{i=1}^{N} \hat{I}_i c_i}{\sum_{i=1}^{N} c_i} \tag{5}$$

This gives the mean to which the equity scores should be scaled in order for the weighted mean to be zero. Then, the three scoring years are scaled using the following process:

1. Scaling each of the three years' equity scores to the means and standard deviations given above. After this, all three years' scores have a weighted mean of zero and standard deviation of one, with higher values indicating higher levels of socioeconomic disadvantage.
2. These scores are combined as a weighted average for each school, based on the roll in the three years.
3. The weighted averages from 2. are scaled again as in 1. above, however, schools with extreme values are excluded before applying the scaling process. These outlier schools then have their scores transformed to the same scale as the other schools.
4. The scaled weighted averages from step 3. are finally linearly transformed into a scale from 344 to 569 inclusive, where the value of -2.075 is assigned to 344 and the value of 2.399 is assigned to 569. Any value of less than -2.075 is also assigned the value 344, and any value of greater than 2.399 is assigned the value 569. Note that the capped untransformed numbers are used for EQI 2023 and will be recalibrated periodically.

The success of the approach relies on the assumption that the shape of the equity score distribution remains relatively unchanged over time. This will be tested each year when the Equity Index is updated.

Scaling the scores and averaging them across years guards against large fluctuations in scores, and therefore funding, over time. This is particularly important for schools with small rolls, which could otherwise see large changes in funding as children come and go from the school.

# References

Ainley, J., Graetz, B., Long, M., & Batten, M. (1995). Socioeconomic status and school education. J. Pediatr. Psychol, 20(2), 205-216.

Crichton, S., Dixon, S., & Ball, C. (Unpublished). Predicting Students' Future NCEA Attainment Using Census and Administrative Data. Wellington: The Treasury.

Crampton, E., & Udahemuka, M. (2018). Score! Transforming NCEA Data. Wellington: The New Zealand Initiative.

Dixon, S. (2018). Student Mobility Across Schools and its Links to Under-achievement. Wellington: NZ

Hutchings, H. A., Evans, A., Barnes, P., Demmler, J., Heaven, M., Hyatt, M. A., & Rodgers, S. E. (2013). Do children who move home and school frequently have poorer educational outcomes in their early years at school? An anonymised cohort study. PloS one, 8(8), e70601.

Ministry of Education. (2019a). Equity Index Technical Report.

Ministry of Education. (2019b). Understanding student progress and achievement. He Whakaaro.

Portes, A., & MacLeod, D. (1996). Educational progress of children of immigrants: The roles of class, ethnicity, and school context. Sociology of education, 255-275.

The Treasury. (Unpublished). Predicting Students' Future NCEA Attainment Using Census and Administrative Data. Wellington: The Treasury.

Ussher, S. (2008). Post-school choices: How well does academic achievement predict the tertiary education choices of school leavers? Wellington: Ministry of Education.

Webber, A. (2020). He Whakaaro: What is the relationship between attendance and attainment? Wellington: Ministry of Education.

Webber, A., & Loader, M. (2020). He Whakaaro: How does school type impact on student outcomes? Wellington: Ministry of Education.

We **shape** an **education** system that delivers **equitable** and **excellent outcomes**

---

He mea **tārai** e mātou te **mātauranga** kia **rangatira** ai, kia **mana taurite** ai ōna **huanga**





MINISTRY OF EDUCATION
TE TĀHUHU O TE MĀTAURANGA