Australian Bureau of Statistics

CAPACITY TO CONTRIBUTE: ALTERNATIVE STATISTICAL MEASURES FOR USE IN THE DMI METHODOLOGY – PRELIMINARY RESULTS, PART 2

Direct Measure of Income Refinement Working Group Paper

March 2021

CAPACITY TO CONTRIBUTE: ALTERNATIVE STATISTICAL MEASURES FOR USE IN THE DMI METHODOLOGY – PRELIMINARY RESULTS, PART 2

Executive Summary

The Australian Bureau of Statistics (ABS) is evaluating the fitness-for-purpose of six alternative statistical summary measures that could be used in the Direct Measure of Income (DMI) methodology for calculating Capacity to Contribute (CTC). These summary measures are:

- the first quartile;
- the mid-hinge;
- the tri-mean;
- the mean;
- the trimmed mean; and
- the winsorised mean.

Definitions of these measures, as well as the median, are provided in Appendix 1. A detailed introduction to these measures was provided in the January 2021 DMI Refinement Working Group paper, 'Capacity to Contribute: Introduction to alternative statistical measures for use in the DMI methodology', and is available at the DESE website.

At the February 2021 DMI Refinement Working Group meeting, stakeholders considered the principles which could be used to assess the alternative statistical summary measures. Six principles were proposed by the ABS, and stakeholders suggested a seventh principle – simplicity. The resulting conceptual framework consists of the following seven principles:

1. relative orientation to the DMI score;
2. volatility;
3. confidentiality;
4. accuracy;
5. robustness to extreme income values;
6. sensitivity to distributional differences; and
7. simplicity.

Preliminary analysis of the first three principles was presented to the DMI Refinement Working Group in February 2021. This paper contains preliminary analysis of the remaining four principles.

Table 1 summarises the results of preliminary assessment against the seven principles.

Table 1: Summary of preliminary assessment of alternative statistical summary measures against the seven principles in the conceptual framework.

Statistical summary measure	Relative orientation to the DMI score	Volatility	Confidentiality	Accuracy	Robustness to extreme income values	Sensitivity to distributional differences	Simplicity
Legend	Direction • Scores increase: ↑ • Scores decrease: ↓ • Scores are similar: ↔ Magnitude • 95% of scores are within 2 points of DMI score: ●○○ • 95% of scores are within 3 points of DMI score: ●●○ • 95% of scores are within 6 points of DMI score: ●●●	Stability • Scores are more stable than DMI score: ↑ • Scores are less stable than DMI score: ↓	Risk level • Same as DMI score: ●○ • Higher risk than DMI score: ●● Treatment complexity • Simple: ●○ • Complex: ●● Availability of scores • Slightly lower availability than DMI scores: ●○ • Similar availability to DMI scores: ●●	Sensitivity to missingness and imputation • Very insensitive: ●○○ • Insensitive: ●●○ • Sensitive: ●●●	Robustness • Not robust: ○○○ • Less robust than DMI score: ●○○ • Slightly less robust than DMI score: ●●○	Sensitivity to changes in lower end of distribution • Much less sensitive than the DMI score: ●○○ • Less sensitive than the DMI score: ●◐○ • As sensitive as the DMI score: ●●○ • More sensitive than the DMI score: ●●●	Usage • No examples of usage: ○○○ • Used in technical applications: ●◐○ • Widely used: ●●●
First quartile	Direction: ↑ Magnitude: ●●● • Scores tend to increase. • Some increase by a relatively large amount. • Scores increase for the majority of schools in the 96 to 114 DMI score range.	Stability: ↓ • Slightly less stable than the DMI score. • Slightly more schools have some annual change in score. • Average score changes are slightly larger in magnitude than those of DMI score.	Risk level: ●○ Treatment complexity: ●○ Availability of scores: ●○ • Similar risk level and treatments to the DMI score. • Slightly lower availability, due to the larger min. number of contributing family incomes.	Sensitivity to missingness and imputation: ●●● Implementation decisions req'd: No • Sensitive to changes in missingness and imputation strategy. • Methodological decisions that impact on accuracy are not required.	Robustness: ●●○ • Slightly less robust than the median.	Sensitivity to changes in lower end of distribution: ●●● • More sensitive than the median to changes in the lower end of the income distribution. • Expected to be less sensitive than the median to changes at the upper end.	Simplicity: Simple Usage: ●●● Development of quality assurance process and supporting materials req'd: Yes Development of indicators to support interpretation of scores req'd: No • Widely used. • Indicators to support interpretation of scores likely to be similar to current approach.

Statistical summary measure	Relative orientation to the DMI score	Volatility	Confidentiality	Accuracy	Robustness to extreme income values	Sensitivity to distributional differences	Simplicity
Mid-hinge	Direction: ↓ Magnitude: ●●○ • Similar to the mean, but differences are smaller in magnitude.	Stability: ↑ • Slightly more stable than the DMI score.	Risk level: ●○ Treatment complexity: ●○ Availability of scores: ●○ • Same as first quartile.	Sensitivity to missingness and imputation: ●●○ Implementation decisions req'd: No • Like the median, the mid-hinge is not sensitive to changes in missingness and imputation strategy. • Methodological decisions that impact on accuracy are not required.	Robustness: ●●○ • Slightly less robust than the median.	Sensitivity to changes in lower end of distribution: ●●○ • Like the median, sensitive to changes in the lower end of the income distribution.	Simplicity: Simple Usage: ○○○ Development of quality assurance process and supporting materials req'd: Yes Development of indicators to support interpretation of scores req'd: No • No examples of usage. • Indicators to support interpretation of scores likely to be similar to current approach.
Tri-mean	Direction: ↔ Magnitude: ●○○ • Scores are similar to DMI scores, with a large proportion of schools having no difference in score. • Scores decrease for most schools in the 105 to 114 DMI score range, but magnitude of decreases is small.	Stability: ↑ • Slightly more stable than the DMI score.	Risk level: ●○ Treatment complexity: ●○ Availability of scores: ●○ • Same as first quartile.	Sensitivity to missingness and imputation: ●●○ Implementation decisions req'd: No • Same as mid-hinge.	Robustness: ●●○ • Slightly less robust than the median.	Sensitivity to changes in lower end of distribution: ●●○ • Same as mid-hinge.	Simplicity: Simple Usage: ○○○ Development of quality assurance process and supporting materials req'd: Yes Development of indicators to support interpretation of scores req'd: No • Same as mid-hinge.

Statistical summary measure	Relative orientation to the DMI score	Volatility	Confidentiality	Accuracy	Robustness to extreme income values	Sensitivity to distributional differences	Simplicity
Mean	Direction: ↓ Magnitude: ●●● • Scores tend to decrease. • Some decrease by a relatively large amount. • Scores decrease for a large majority of schools in the 96 to 119 DMI score range.	Stability: ↑ • Slightly more stable than, or similar to, the DMI score.	Risk level: ●● Treatment complexity: ●● Availability of scores: ●● • Higher risk level than DMI score. • More complex treatment required to ensure safe data release. • Availability is similar to median. • Further work required to determine safe release strategy. • Potential impact on existing privacy framework.	Sensitivity to missingness and imputation: ●●● Implementation decisions req'd: No • Very stable in the presence of changes in missingness and imputation strategy that typically affect lower income values. • Methodological decisions that impact on accuracy are not required.	Robustness: ●●● • Not robust. • Scores based on the mean change in line with changes in the presence or absence of extreme income values.	Sensitivity to changes in lower end of distribution: ●●● • Less sensitive to changes at the lower end of the income distribution than the median. • More sensitive to changes in the upper end.	Simplicity: Simple Usage: ●●● Development of quality assurance process and supporting materials req'd: Yes Development of indicators to support interpretation of scores req'd: Yes • Widely used. • Indicators to support interpretation of scores require further investigation.

Statistical summary measure	Relative orientation to the DMI score	Volatility	Confidentiality	Accuracy	Robustness to extreme income values	Sensitivity to distributional differences	Simplicity
Trimmed mean	Direction: ↓ Magnitude: ●●○ • Similar to the mean, but differences are smaller in magnitude.	Stability: ↑ • Slightly more stable than the DMI score.	Risk level: ●● Treatment complexity: ●● Availability of scores: ●● • Similar to mean.	Sensitivity to missingness and imputation: ●○○ Implementation decisions req'd: Yes • Sensitivity to changes in missingness and imputation appears similar to the mean but is expected to increase as trimming threshold decreases. • Investigation into trimming thresholds is required, and may need recalibration each year.	Robustness: ●○○ • Less robust than the median.	Sensitivity to changes in lower end of distribution: ●●○ • More sensitive than the mean to changes at the lower end of the income distribution.	Simplicity: Complex Usage: ●●○ Further development of quality assurance process and supporting materials req'd: Yes Further development of indicators to support interpretation of scores req'd: Yes • Used in technical applications. • Indicators to support interpretation of scores require further investigation.
Winsorised mean	Direction: ↓ Magnitude: ●●○ • Similar to the mean, but differences are smaller in magnitude.	Stability: ↑ • Slightly more stable than the DMI score.	Risk level: ●● Treatment complexity: ●● Availability of scores: ●● • Similar to mean.	Sensitivity to missingness and imputation: ●○○ Implementation decisions req'd: Yes • Similar to the trimmed mean.	Robustness: ●○○ • Less robust than the median.	Sensitivity to changes in lower end of distribution: ●●○ • Sensitivity is similar to that of the trimmed mean.	Simplicity: Complex Usage: ●●○ Further development of quality assurance process and supporting materials req'd: Yes Further development of indicators to support interpretation of scores req'd: Yes • Same as trimmed mean.

Introduction

In 2020-21, the Australian Bureau of Statistics (ABS) was engaged by the Department of Education, Skills and Employment (DESE) to evaluate the fitness-for-purpose of alternative statistical summary measures that could be used in the Direct Measure of Income (DMI) methodology for calculating Capacity to Contribute (CTC). As part of this engagement, the ABS is evaluating the following six alternative statistical summary measures:

- the first quartile;
- the mid-hinge;
- the tri-mean;
- the mean;
- the trimmed mean; and
- the winsorised mean.

Definitions of these measures, as well as the median, are provided in Appendix 1. A detailed description of these measures was provided in the January 2021 DMI Refinement Working Group paper, 'Capacity to Contribute: Introduction to alternative statistical measures for use in the DMI methodology', available at the DESE website.

At the February 2021 DMI Refinement Working Group meeting, stakeholders considered the principles which could be used to assess the alternative statistical summary measures. Six principles were proposed by the ABS and stakeholders suggested a seventh principle – simplicity. The resulting conceptual framework is shown in Table 2.

Table 2: Conceptual framework for assessing alternative statistical summary measures for the DMI.

Principle	Definition	Key assessment criteria
1. Relative orientation to the DMI score	The extent to which scores based on alternative measures are the same as, higher than, or lower than the DMI score.	Difference between score based on alternative measure and the DMI score based on the median.
2. Volatility	A measure of the change in a data item over time.	The annual change in score, from 2018-2020.
3. Confidentiality	The requirement to protect the secrecy and privacy of information collected from individuals.	The impact of mitigating confidentiality risks on the availability, accuracy, complexity and interpretability of scores and supplementary data.
4. Accuracy	The degree to which data correctly describes the "real world" object or event.	The impact of technical decisions, such as the trimming parameter and winsorisation threshold. The sensitivity of scores to missingness and refinements to the CTC income imputation strategy.

Principle	Definition	Key assessment criteria
5. Robustness to extreme income values	The extent to which the measure is stable in the presence of extreme income values.	The extent to which scores change when outliers are introduced. Differences between scores based on alternative measures and the DMI score based on the median, for schools with a large proportion of outliers.
6. Sensitivity to distributional differences	The degree to which a statistical measure changes value in response to distributional changes.	How scores based on the alternative measures reflect differences in school income distributions.
7. Simplicity	The state or quality of being simple – i.e. easy to understand, implement, quality assure and describe.	The extent to which the measures are commonly used. Whether an alternative summary measure requires further methodological decision-making and ongoing recalibration. The impact of using an alternative measure on quality assurance and supporting information.

Preliminary analysis of the first three principles – (1) relative orientation to the DMI score, (2) volatility and (3) confidentiality – was presented to the DMI Refinement Working Group in February 2021. The ABS' paper for the February DMI Refinement Working group, 'Capacity to Contribute: Alternative statistical measures for use in the DMI methodology – preliminary results', also contained background information about the legislative and policy context in which DMI scores are used. The results presented in this paper should also be considered in that context.

This paper contains preliminary analysis of the remaining four principles – (4) accuracy, (5) robustness to extreme income values, (6) sensitivity to distributional differences and (7) simplicity.

Methodology

A detailed explanation of the methodology for calculating DMI scores and DMI-based CTC scores is provided in Appendix 2.

For the purposes of this analysis, scores based on the median and alternative statistical summary measures were calculated by:

i. taking the summary measure of the school's family income distribution;
ii. standardising the summary measure to obtain a score;
iii. rounding the score to the nearest integer; and
iv. bottom and top coding to 93 and 125, respectively.

All other aspects of the DMI methodology, such as the income imputation strategy and standardisation process used to convert summary income values into scores, were held constant, unless otherwise noted.[1] This approach ensures that the differences in school scores described in this paper result solely from the use of the alternative summary measures.

Analysis of the potential impact of using alternative statistical summary measures on school funding is out-of-scope of the ABS' engagement. However, to support the analysis of score changes which have an impact on funding, schools with scores below 93 were assigned a score of 93, and schools with scores above 125 were assigned a score of 125 in this analysis.

To be included in a comparison of DMI scores with scores based on an alternative measure, schools must meet the confidentiality requirements of both the median and the alternative measure.

[1] Standardisation is a common statistical process which converts a set of numbers, which may have any average and spread, into a pre-determined average and spread. It does not change the order of school communities in the distribution. For CTC, income imputation refers to the methods used to determine a value of Adjusted Taxable Income (ATI) for those parents whose ATI is missing in the linked administrative data available via the Multi-Agency Data Integration Project (MADIP).

Principle 4: Accuracy

Accuracy is defined as the degree to which data or a statistical measure correctly describes the "real world" object or event it is intended to measure. In relation to the DMI methodology, the accuracy of the summary measure refers to how well it reflects the 'true' income distribution of a school community. Missing income values, as well as the application of methods such as trimming and winsorisation, reduce the extent to which a statistical summary measure can make use of information from the 'true' income distribution of a school community. Therefore, sensitivity to missing income values and imputation strategies, as well as the intrinsic statistical properties of the summary measure, can affect the level of accuracy of a statistical summary measure. This has clear implications for how useful and meaningful the measure is for its intended purpose.

The key criteria for assessing the accuracy of the alternative statistical summary measures are:

- the sensitivity of the alternative summary measures to missing or imputed income, compared with the sensitivity of the median to missing or imputed income; and
- the impact of methodological decisions such as the trimming threshold and winsorisation parameters, on the ability of the alternative summary measures to accurately summarise the school community's income distribution.

Summary

Table 3: Summary of preliminary assessment of accuracy.

Statistical summary measure	Sensitivity to missing and imputed income	Impact of methodological decisions
Median	Stable in the presence of changes to missing or imputed values. The majority of scores (70%) did not change as a result of a refined imputation strategy and for those which did change, the average score change was small (0.14 points).	There are no further methodological decisions associated with using the median that impact on accuracy.
First quartile	Sensitive to changes to missing and imputed values. Scores changed for half of schools as a result of using a refined imputation strategy and for those which did change, the average score change was 0.46 points.	There are no further methodological decisions associated with using the first quartile that impact on accuracy.
Mid-hinge	Similar in sensitivity to the median. The majority of scores (76%) did not change as a result of using the refined imputation strategy, and for those which did change, the average score change was 0.24 points.	There are no further methodological decisions associated with using the mid-hinge that impact on accuracy.

Statistical summary measure	Sensitivity to missing and imputed income	Impact of methodological decisions
Tri-mean	Similar in sensitivity to the median. The majority of scores (77%) did not change as a result of using the refined imputation strategy, and for those which did change, the average score change was 0.18 points.	There are no further methodological decisions associated with using the tri-mean that impact on accuracy.
Mean	Less sensitive than the median. The majority of scores (87%) did not change as a result of using the refined imputation strategy, and for those which did change, the average score change was less than 0.01 points.	There are no further methodological decisions associated with using the mean that impact on accuracy.
Trimmed mean	Similar to the mean. The majority of scores (83%) did not change as a result of using the refined imputation strategy, and for those which did change, the average score change was close to zero (-0.005 points).	Investigation regarding trimming threshold is required, considering trade-off between accuracy and robustness to extreme values / volatility of data over time.
Winsorised mean	Similar to the mean. The majority of scores (85%) did not change as a result of using the refined imputation strategy, and for those which did change, the average score change was 0.09 points.	Investigation regarding winsorisation method and parameters is required, considering trade-off between accuracy and robustness to extreme values / volatility of data over time.

Sensitivity to missing and imputed income

Overall, the extent of missing income values in the CTC population is low. In the 2020 Address Collection, 6.5% of all parent records were excluded from the DMI score calculation due to a lack of income information and 1.8% were imputed a zero income (consisting of 1.2% for whom a low income concession card was available and 0.6% which had a lack of income information), under the existing income imputation strategy[2].

The extent to which scores based on a given summary measure are stable in the presence of missing or imputed income values depends firstly on the summary measure itself, secondly on the perceived structure of the missingness in the data (which cannot be known with certainty) and finally on the imputation strategies used to address that missingness. For example, if incomes at the lower end of the distribution are more likely to be missing, it is reasonable to expect that scores based on the first quartile would be relatively sensitive to that missingness and the strategies used to address it. It is also reasonable to expect that in this situation the mean would be relatively less sensitive to such

[2] For further information about income coverage and imputation in CTC data, see the January 2021 DMI Refinement Working Group paper, 'Capacity to Contribute: Income Imputation – discussion paper', available at: www.dese.gov.au/direct-measure-income-refinement-working-group.

missingness, given that (all else equal), a low income value has less of an impact on the mean than a high income value. Therefore, incorporating (or not) a low income value into the statistical summary measure might be have a greater impact on the 'truthfulness' of an estimate of the first quartile than it would for an estimate of the mean. Thus, both the perceived nature of the missingness and the choice of statistical measure may affect the ability with which an estimate determined using a given statistical measure accurately represents the corresponding true income distribution of a school community.

To assess the sensitivity of the alternative statistical summary measures to missing or imputed income values, two sets of scores were calculated for all schools which met the confidentiality requirements. The first set of scores, based on the median and each alternative statistical summary measure, used incomes assigned according to the existing imputation strategy for the DMI methodology. The second set of scores used incomes assigned according to a refined income imputation strategy, which incorporated the additional information sources:

- government payments data available in the DOMINO Centrelink Administrative Dataset via MADIP; and
- modelled ATI estimates, for parent records which linked to MADIP.

For further information about the existing imputation strategy for the DMI methodology and the refined strategy used in this analysis, see Appendix 3.

Inclusion of DOMINO and modelled incomes

The inclusion of DOMINO data in the refined imputation strategy led to a change in income for approximately 2% of students in 2020. Almost every school (96%) included at least one student with a parental income from DOMINO.

Applying a modelled income estimate to parents who linked to MADIP but for whom no income data was available in the administrative sources resulted in a change in income for 0.55% of students in 2020. Approximately 70% of schools had at least one student with a modelled parental income.

Impact of refined income imputation strategy on scores

For each statistical summary measure, scores based on the existing imputation strategy were subtracted from those based on the refined income imputation strategy. The difference represents how much a school's score would increase (for positive differences) or decrease (for negative differences), if the refined income imputation strategy were used.

For scores based on the median and the six alternative statistical summary measures, the proportion of schools which had no change in score, as a result of applying the refined income imputation strategy to 2020 CTC data, is shown in Figure 1, below. For DMI scores based on the median, the use of the refined income imputation strategy resulted in no change in score for 71% of schools. For scores based on the first quartile, half of schools (50%) had no change in score when the refined income imputation strategy was used.

Figure 1: Proportion of schools with no difference between the score based on the original and the refined income imputation strategies, in 2020.

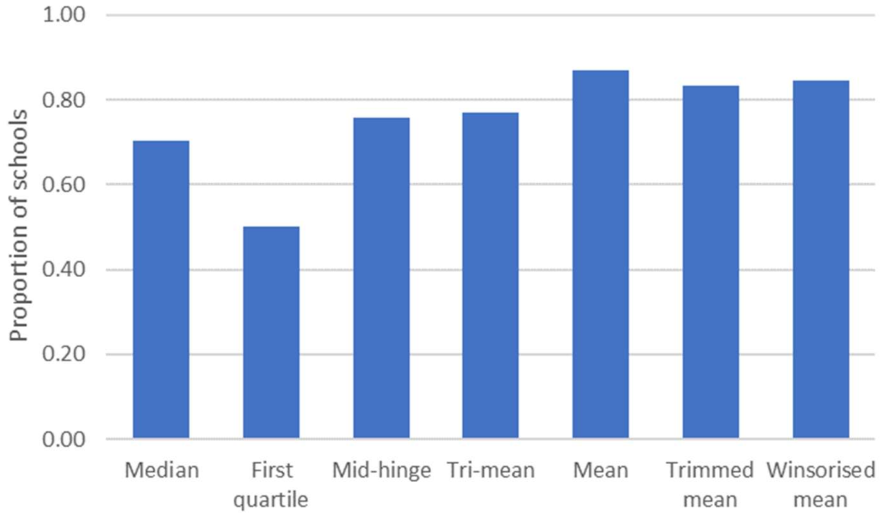
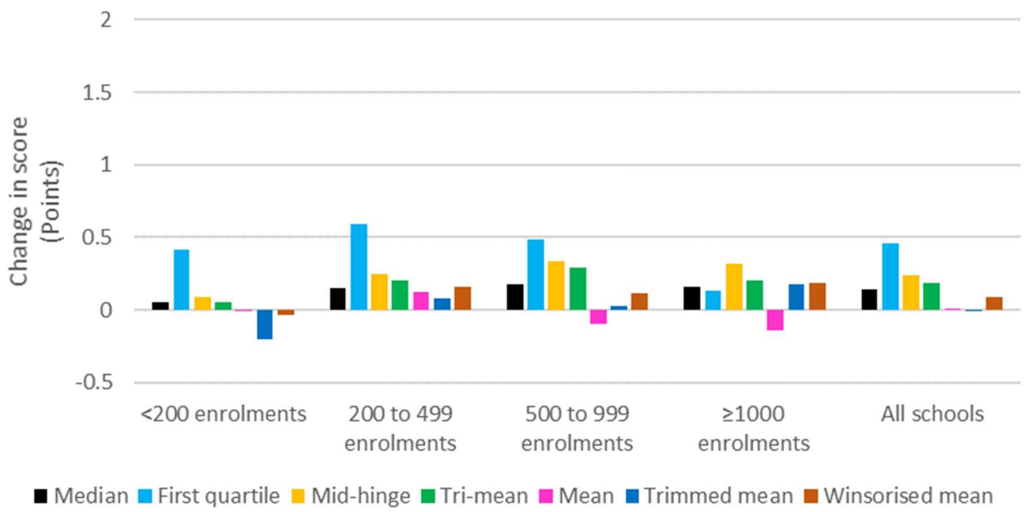


Figure 1 shows that for the majority of schools, the change in income imputation strategy did not have an impact on the school's score, regardless of the measure used. This may be because:

- the quality of the income data for the school community was high. In these cases, few incomes were missing and needed to be imputed, and / or the incomes which were imputed did not have a large enough impact on the statistical summary measure to result in a change in score; or
- the school's score changed, but remained below 93 or above 125, despite the change in income imputation strategy.

For schools whose scores changed due to the use of the refined income imputation strategy, the average change in scores is shown in Figure 2.

Figure 2: Average difference between scores calculated using the original and the refined income imputation strategies, by statistical summary measure and school size category, in 2020.



The use of the refined income imputation strategy had the greatest impact on scores based on the first quartile, which increased by almost half a point (0.46 points) on average. Overall, scores based on the median increased by an average of 0.14 points when the refined income imputation strategy was used. The average change in scores based on the mean was almost zero (0.01 points).

Impact of methodological decisions on accuracy

One aspect of the accuracy of the statistical summary measure to be used in the DMI methodology is the extent to which it summarises the 'true' income distribution of a school community. As the DMI score is a relative measure, it is important that the summary measure accurately reflects the 'true' income distributions of school communities to support the comparison of schools. For example, the mean – as the sum of all the incomes divided by the total number of incomes – makes use of all available data to summarise the school community income distribution. However, in the case of the trimmed mean, some incomes in the distribution are not used, and for the winsorised mean, some incomes in the distribution are reduced. These processes affect the extent to which the trimmed mean and the winsorised mean can represent the 'true' income distribution of a school community.

This raises the question – why consider the trimmed or winsorised mean at all and not simply use the mean?

As noted above, the mean makes use of all the available data about the population of interest. However, this does not necessarily imply that it is the "best" measure. When the mean is used to summarise a skewed and/or volatile phenomenon, such as income and especially for small populations, it becomes apparent that the mean is not robust to extreme values. This is because any value that is magnitudes larger than the rest will have a similarly large effect on the mean.

There are several ways of dealing with extreme values:

- changing the weights of the extreme values (this has not been investigated for use in the DMI methodology);
- changing their values (e.g. winsorisation); and
- removing the extreme values from the calculation of the mean (e.g. trimming).

Therefore, considering the fitness-for-purpose of the mean, trimmed mean and winsorised mean for use in the DMI methodology involves considering the trade-off between accuracy (as defined above) on one hand, and robustness to extreme values and volatile data over time, on the other.

For the trimmed mean and the winsorised mean, the trade-off between accuracy and robustness to extreme values and volatile data over time is affected by where the trimming threshold is set and the degree to which extreme values are winsorised (or dampened). For example, a trimmed mean that is more robust would exclude a larger number of extreme values, but in doing so, it would be calculated using less of the income distribution of the population of interest. In other words, a lower threshold for the trimmed mean (i.e. discarding more observations), or a greater dampening parameter for the winsorised mean (i.e. bringing extreme values down further) implies that some accuracy is being sacrificed for improved robustness.

From a statistical perspective, setting an optimal threshold aims to achieve a high level of accuracy and robustness, by effectively identifying and treating only those extreme values which have the greatest influence on the summary statistic. Such an approach would take into account various factors, including:

- the distribution of income in a school in a given year;
- whether changes in the distribution of income over time are representative of a trend, or whether they are just due to chance; and
- what constitutes an 'extreme' value, in either absolute or relative terms.

It is important to note that the ‘statistically optimal’ threshold for trimming or winsorisation may not be the same for all schools. For example, some schools may have quite symmetrical income distributions, with no extreme values, in which case the ‘statistically optimal’ approach would be not to trim or winsorise at all.

Simple implementation choices were used by the ABS to enable the preliminary assessment of the trimmed mean and winsorised mean as alternative statistical summary measures in the DMI methodology. The implementation of either of these summary measures would require further investigation to determine the level of trimming or the method and degree of winsorisation to be applied to school community income distributions. Further development of indicators to quantify accuracy, robustness and/or volatility to inform implementation decisions would also be required.

Principle 5: Robustness to extreme income values

This principle is defined as the extent to which a statistic is stable in the presence of extreme values. For use in the DMI methodology, it is preferable that a statistical summary measure is not heavily impacted by the presence of extreme family income values and represents the majority of family income values in the school community.

The key criteria for assessing the robustness to extreme income values are:

- The extent to which scores change when extreme outlier income values are introduced into school community income distributions.
- The extent to which scores change in line with changes in the proportion of extreme outlier values in school community income distributions.

Summary

Table 4: Summary of preliminary assessment of robustness to extreme values.

Statistical summary measure	Robustness to single extreme income value	Robustness to observed changes in extreme outliers from 2019 to 2020
Median	Very robust. On average, scores based on the median do not change with the introduction of an extremely high income value.	Relatively robust. For schools with a decrease in the proportion of extreme outliers, scores based on the median decreased by less than half a point, on average. For schools with an increase in the proportion of extreme outliers, scores increased by approximately 0.1 points on average.
First quartile	Slightly less robust than the median.	Less robust than the median.
Mid-hinge	Slightly less robust than the median.	Very robust.
Tri-mean	Slightly less robust than the median.	Very robust.
Mean	Not robust. Scores based on the mean increase on average by 0.5 points with the introduction of an extremely high income value. For small schools, the average increase in score is greater than 1 point.	Not robust. Scores based on the mean increase by 1 point on average when there is an increase in the proportion of outliers, and decrease by over 1 point, on average, when there is a decrease in the proportion of outliers at a school.
Trimmed mean	Less robust than the median.	Less robust than the median.
Winsorised mean	Less robust than the median.	Less robust than the median.

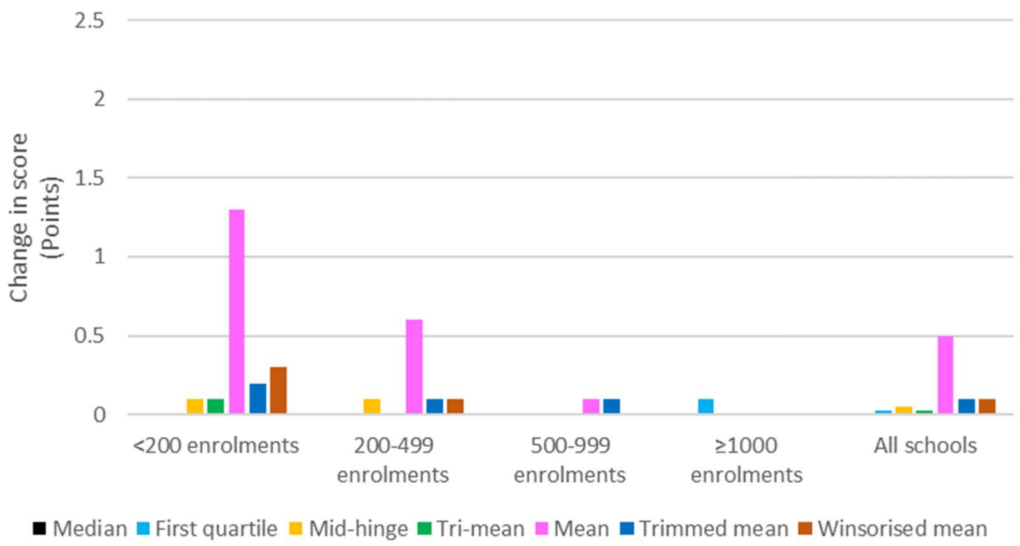
Scenario analysis: Introduction of an extreme income value

This analysis considers a scenario in which one very high-income family joins a school.

For a school, a family income of $2,000,000 is added to the 2020 income distribution, and new scores, based on each statistical summary measure, are calculated. The difference in scores is calculated by subtracting the scores based on the school's original income distribution from the new scores, for each statistical summary measure. Therefore, the difference represents the increase in a school's score that would result from a high-income family joining the school.

As the arrival of a high-income family is likely to affect small schools more than larger schools, this analysis was repeated for 10 randomly selected schools in each of four size categories: (1) less than 200 students, (2) 200 to 499 students, (3) 500 to 999 students, and (4) 1000 students or more. Figure 3 shows the average increase in score, by statistical summary measure, for the 40 schools overall and for the 10 schools in each size category, using 2020 CTC data.

Figure 3: Average increase in score, in points, when a single high-income family is added to a school's 2020 income distribution, by statistical summary measure and school size category.



Overall, the mean is the least robust of the statistical summary measures to an additional extreme income value, with scores based on the mean increasing by 0.5 points on average, for all schools. The median is the most robust summary measure overall, with no average change in score as a result of the extreme income value.

As expected, the impact on scores is greater for smaller schools. For schools with less than 200 students, scores based on the mean increase by an average of 1.3 points with the addition of a single high-income family to the schools' 2020 income distribution.

The ABS notes that the results presented above should be interpreted with caution, as they are based on a small sample of schools. Schools in a given size category are likely to exhibit a variety of differently shaped income distributions. Therefore, although the 10 schools in each size category were randomly selected, they may not be representative, in terms of their income distribution, of other schools in the same size category. A more comprehensive analysis, which repeats the process described above for all schools, is possible but would require further analysis.

Impact of a change in extreme income values over time on school scores

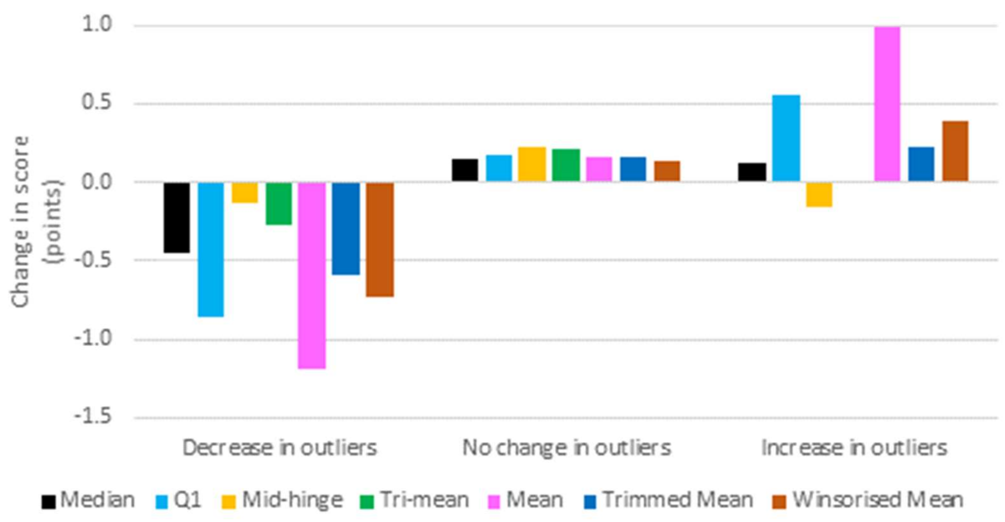
This section aims to assess how school scores are affected when the proportion of extreme income values (outliers) at a school changes from one year to the next. An extreme outlier was defined as an income value greater than three times the Inter-Quartile Range (IQR) above the third quartile value of a schools' income distribution[3]. That is: Q3 + 3 x IQR, where the IQR is defined as Q3 - Q1.

The analysis presented in this section is based on observed changes in school income distributions from 2019 to 2020, rather than on deliberately constructed scenarios. This analysis consisted of:

- calculating the annual change in each school's score, using each statistical summary measure. The difference was calculated as the 2020 score minus the 2019 score.
- calculating the proportion of income values classified as 'extreme outliers' at each school in 2019 and 2020.
- calculating the change in the proportion of extreme outliers at each school from 2019 to 2020, and classifying schools according to this change. Schools were classified as having:
  - a decrease in extreme outliers from 2019 to 2020, for changes in the proportion of extreme outliers at a school of less than -0.5%;
  - no change in extreme outliers from 2019 to 2020, for changes in the proportion of extreme outliers between -0.5% and 0.5% (inclusive); or
  - an increase in extreme outliers from 2019 to 2020, for changes in the proportion of extreme outliers of more than 0.5%.

For schools which had a change in score from 2019 to 2020[4], Figure 4 shows the average change in score, based on the median and the alternative summary measures, with schools grouped by their level of change in the proportion of extreme outlier income values from 2019 to 2020.

Figure 4: Average size of score change, in points, by statistical summary measure, for schools with a decrease, no change, or an increase in the proportion of extreme outlier incomes, from 2019 to 2020.



[3] There are various methods for defining and detecting outliers in statistical analyses, and there is no universal approach. This simple, non-parametric approach is known as Tukey's fences (Tukey, J. (1977) Exploratory data analysis, Addison Wesley Publishing, Reading, Mass.).
[4] Preliminary analysis of the proportion of schools with a change in score from 2019 to 2020 is provided in the volatility section of 'Capacity to Contribute: Alternative statistical measures for use in the DMI methodology – Preliminary results.'

DMI scores, based on the median, were relatively robust to changes in the proportion of extreme outliers from 2019 to 2020. For schools which experienced a decrease in the proportion of extreme outliers during that period, DMI scores decreased by less than half a point, on average. For schools which experienced an increase in the proportion of extreme outliers, DMI scores increased by approximately 0.1 points on average. The largest increases in DMI scores typically occurred when the distribution of incomes in the school was generally higher and/or broader in 2020 than in 2019 (and vice versa, for the largest decreases in score).

In contrast, the mean was the least robust of the statistical summary measures to changes in the proportion of extreme outliers from 2019 to 2020. For schools which experienced a decrease in the proportion of extreme outliers, scores based on the mean decreased by almost 1.2 points, on average, from 2019 to 2020. For schools which experienced an increase in the proportion of extreme outliers, scores based on the mean increased by approximately 1 point, on average. This is expected, given that large income values have a proportionately large impact on the mean. The largest positive (or negative) changes in scores based on the mean occurred under similar circumstances to those of scores based on the median.

By looking at the change in a school's score over time, holding the statistical summary measure constant, this analysis aims to focus on changes in score associated with changes in a school's income distribution, and in particular, with the number of extreme outliers in the income distribution. However, it should be noted that scores may also change for reasons other than the presence or absence of extreme outliers. For example, the results for the first quartile, above, may also reflect findings regarding accuracy and volatility. The decrease in the mid-hinge for schools with an increase in the proportion of extreme outliers may reflect the robustness of the third quartile to outliers, as well as other changes in the income distributions of schools. For example, if distributional changes result in a decrease in a school's third quartile income, its IQR will also decrease, which means that more observations might meet the definition of an extreme outlier, and somewhat counter-intuitively, the proportion of outliers may increase.

Principle 6: Sensitivity to distributional differences

This principle is defined as the degree to which a statistical measure changes value in response to distributional changes. Responsiveness to distributional changes is a desirable property of a statistical summary measure. For example, if the arrival of new students and the departure of existing students results in a change to the incomes of a significant proportion of a school community, this should be reflected in the summary measure and DMI score for that school.

The key criterion for assessing the sensitivity of a score to distributional changes is:

- The extent to which scores change when there is a change in a school's income distribution. For example:
  - when a new cohort of low-income families is introduced into a school's income distribution; or
  - when the proportion of low-income families at a school has changed over time.

Summary

Table 5: Summary of preliminary assessment of sensitivity to distributional differences.

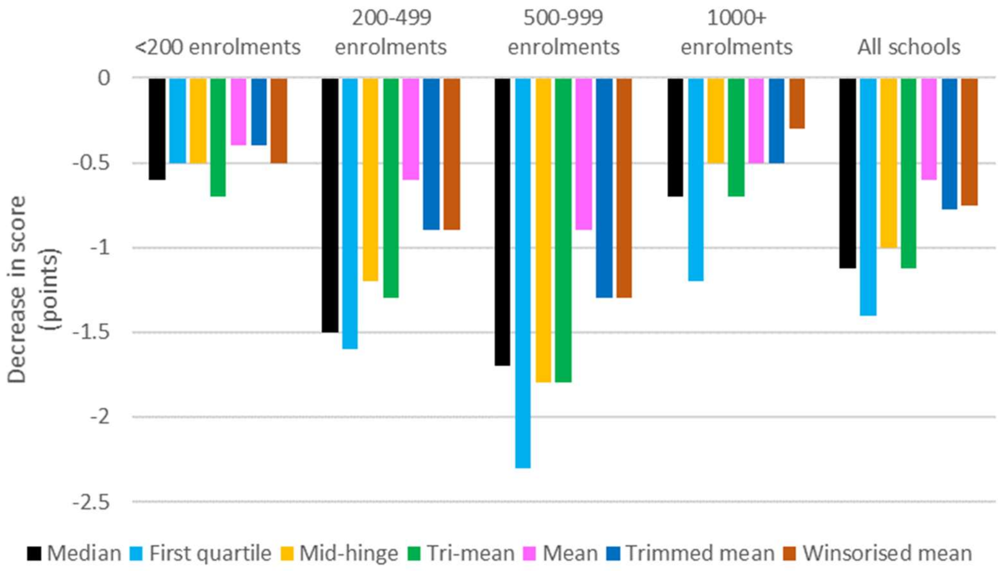
Statistical summary measure	Sensitivity to new cohort of low-income families	Sensitivity to observed changes in proportion of low incomes from 2019 to 2020
Median	Sensitive to introduction of a new cohort of low-income families into a school's income distribution.	Somewhat sensitive to changes in proportion of low incomes.
First quartile	More sensitive than the median to a new cohort of low incomes.	More sensitive to changes in the lower end of the income distribution than the median, but less sensitive to changes at the upper end.
Mid-hinge	Sensitivity is similar to that of the median.	Sensitivity is similar to that of the median.
Tri-mean	Sensitivity is similar to that of the median.	Sensitivity is similar to that of the median.
Mean	Less sensitive than all other measures to a new cohort of low incomes.	More sensitive to changes in the upper end of the income distribution than the median, but less sensitive to changes at the lower end. This is because the former can be expected to represent a larger dollar amount.
Trimmed mean	More sensitive than the mean to a new cohort of low incomes, but less sensitive than the median.	More sensitive than the mean to changes in the proportion of low incomes. This is because, due to the trimming, the distributional changes are averaged across a smaller number of incomes.
Winsorised mean	More sensitive than the mean to a new cohort of low incomes, but less sensitive than the median.	Sensitivity is similar to that of the trimmed mean.

Scenario analysis: A new cohort of low-income families

This analysis considers a scenario in which a new cohort of low-income families joins a school. For a single school, a cohort of students from low-income families, representing approximately 10% of the students at the school, is added to the income distribution. For this analysis, 'low-income' has been defined as an income between the minimum wage (approximately $40,000) and the median income of the school. New scores, based on each statistical summary measure, are calculated. The difference in scores is calculated by subtracting the scores based on the school's original income distribution from the new scores, for each statistical summary measure. Therefore, the difference in score represents the decrease in a school's score that would result from the new cohort of low-income families joining the school.

This analysis was repeated for 10 randomly selected schools in each of the four size categories. Figure 5 shows the average decrease in score, by statistical summary measure, for the 40 schools overall and for the 10 schools in each size category, using 2020 CTC data.

Figure 5: Average decrease in score, in points, when a group of low-income families is added to a school's 2020 income distribution, by statistical summary measure and school size category.



Overall, the median is sensitive to the increase in the proportion of low-income families, decreasing by 1.1 points, on average, across all 40 schools to which the scenario was applied. The mid-hinge and tri-mean are similar in sensitivity to the median, decreasing by an average of 1 point and 1.1 points, respectively. As expected, the first quartile is the most sensitive summary measure to a change in the proportion of low-income families at a school, decreasing by 1.4 points on average for all schools. The mean, trimmed mean and winsorised mean are least sensitive to an increase in the proportion of low-income families at a school, decreasing by an average of 0.6 points, 0.8 points and 0.8 points, respectively.

As with the scenario analysis conducted to assess the robustness of scores to extreme values, these results should be interpreted with caution, as they are based on a small sample of schools. Schools in a given size category are likely to exhibit a variety of differently shaped income distributions. Therefore, although the 10 schools in each size category were randomly selected, they may not be representative, in terms of their income distribution, of other schools in the same size category. A more comprehensive analysis, which repeats the process described above for all schools, is possible but would require further analysis.

It is also important to note that the results described above reflect the choice of scenario – namely, the impact on scores of a number of low-income families joining a school. Various other scenarios may be of interest to stakeholders – such as a scenario where a number of high-income families join a school – and similar analyses could be undertaken of those scenarios.

Impact of a change in the proportion of low-income families over time

This section aims to assess how school scores are affected when the proportion of students with low family incomes changes over time. The proportion of students with low family incomes may increase for various reasons, such as:

- a group of students with low family incomes joins the school;
- the family incomes of existing students may decrease; or
- a group of students with higher family incomes leaves the school.

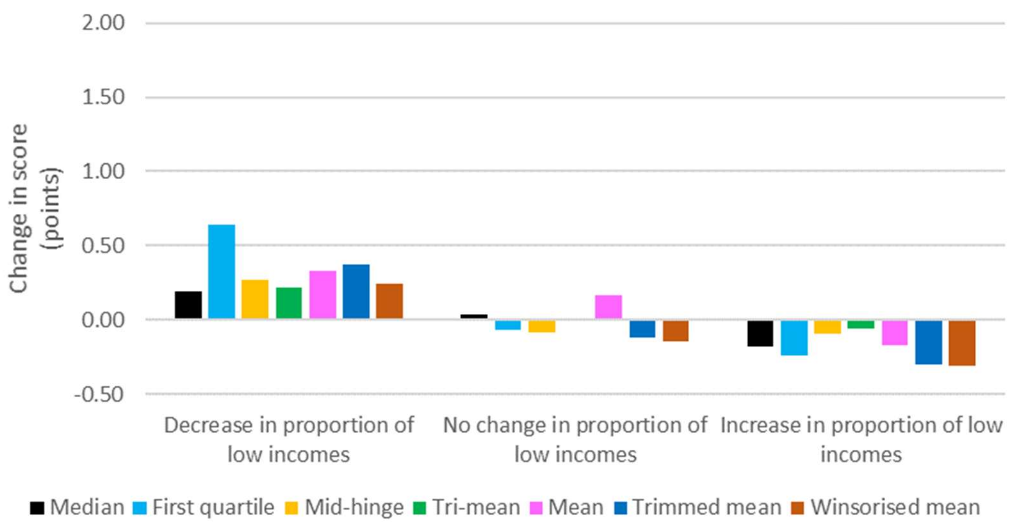
The analysis in this section is based on actual changes in school income distributions from 2019 to 2020, rather than on a hypothetical scenario.

The analysis described in this section consisted of:

- calculating the change in a school's score over time, for all schools which met the confidentiality requirements in both years, using the median and each of the alternative statistical summary measures.
    - The difference was calculated as the 2020 score minus the 2019 score.
- calculating the proportion of incomes classified as 'low' at each school in 2019 and 2020.
    - A low income was defined as an income value less than the first quartile of all incomes in the CTC data for the relevant year.
- calculating the change in the proportion of low incomes in each school from 2019 to 2020, and classifying schools according to this change.
    - Schools were classified as having:
        - a decrease in the proportion of low incomes from 2019 to 2020, for changes in the proportion of low incomes in a school of less than -1%;
        - no change in the proportion of low incomes from 2019 to 2020, for changes in the proportion of low incomes between -1% and 1% (inclusive); or
        - an increase in the proportion of low incomes from 2019 to 2020, for changes in the proportion of low incomes of more than 1%.

For schools which had a change in score from 2019 to 2020[5], Figure 6 shows the average change in score, by statistical summary measure, for schools according to the change in the proportion of low income values in their income distribution, over that period.

Figure 6: Average size of score change, in points, by statistical summary measure, for schools with a decrease, no change, or an increase in the proportion of low incomes in the school income distribution, from 2019 to 2020.



When the proportion of low incomes in a school's income distribution decreases, it is expected that (in the absence of other distributional changes), a school's score would typically increase, and this was observed for scores based on the median and the six alternative summary measures from 2019 to 2020. Scores based on the median were least sensitive to a decrease in the proportion of low incomes, increasing by 0.2 points on average. As expected, scores based on the first quartile were most sensitive to a decrease in the proportion of low incomes, increasing by over 0.6 points on average.

Conversely, when the proportion of low incomes in a school's income distribution increases, it is expected that (in the absence of other distributional changes), a school's score would decrease, and this was observed for scores based on all summary measures from 2019 to 2020. Scores based on the trimmed mean and winsorised mean were the most sensitive to an increase in the proportion of low incomes at a school, both decreasing by approximately 0.3 points on average.

By looking at the change in a school's score over time, holding the statistical summary measure constant, this analysis aims to focus on changes in score associated with changes in a school's income distribution, and in particular, with the proportion of low incomes in the income distribution. However, it should be noted that scores based on the different summary measures may also change for other reasons, such as improvements in data quality over time, the presence or absence of outliers, and changes in incomes at the higher end of the income distribution.

[5] Preliminary analysis of the proportion of schools which had a change in score from 2019 to 2020 is provided in the volatility section of 'Capacity to Contribute: Alternative statistical measures for use in the DMI methodology – Preliminary results.'

Principle 7: Simplicity

Definition

The 'state or quality of being simple', with simple defined as 'easy to understand, deal with, use, etc' and 'not complex or complicated' (Macquarie Dictionary, 2020). For a statistical summary measure for use in the DMI methodology, simplicity encompasses the ease with which the measure is understood by data users, as well as how easy or difficult it is to implement, quality assure and describe.

Related concepts, which aim to support the ability of users to understand and correctly interpret statistics, exist in statistical data quality frameworks.

The ABS Data Quality Framework includes the dimension of interpretability, which refers to the availability of information to help provide insight into statistics and their quality. This information may include glossaries, data item definitions, information about the methodology, data quality indicators and data quality statements.

The European Statistics Code of Practice includes the principle of 'accessibility and clarity'. This principle refers to the presentation of statistics 'in a clear and understandable form, released in a suitable and convenient manner, available and accessible on an impartial basis with supporting metadata and guidance' (p.18).

Rationale

This principle was suggested by members during the February DMI Refinement Working Group meeting. Working group members noted the desirability of a broad understanding of the statistical summary measure and how it is calculated. It was also noted that some summary measures may require further methodological assessment and regular recalibration to support their implementation. This would need to be clearly explained and made available as part of the suite of supporting information.

Key assessment criteria

The simplicity of a statistical summary measure is difficult to assess objectively. The ABS proposes the following questions to support the consideration of stakeholders' needs in assessing simplicity:

- is the measure commonly used?
- does implementation of the measure require further methodological assessment and / or ongoing recalibration?
- how does implementation of the measure affect the:
    - data quality assurance process?
    - information available to support an understanding of the methodology?
    - information available to support stakeholders' understanding of scores and their quality?

The ABS notes that the principle of simplicity may not necessarily be equally weighted when compared with the six principles described in the January DMI Refinement Working Group meeting, and the importance of each principle may vary.

Summary

Table 6: Summary of preliminary assessment of the simplicity of the median and alternative statistical summary measures.

Statistical summary measure	Usage in income / economic analysis	Further methodological assessment & recalibration	Quality assurance process, supporting information and indicators
Median	Widely used	No	Exists as part of current methodology
First quartile	Widely used	No	Further development of quality assurance process and supporting materials required. Indicators to support interpretation of scores likely to be similar to current approach.
Mid-hinge	No examples of usage	No	Further development of quality assurance process and supporting materials required. Indicators to support interpretation of scores likely to be similar to current approach.
Tri-mean	No examples of usage	No	Further development of quality assurance process and supporting materials required. Indicators to support interpretation of scores likely to be similar to current approach.
Mean	Widely used	No	Further development of quality assurance process and supporting materials required. Indicators to support interpretation of scores require further investigation.
Trimmed mean	Used in technical applications	Yes	Further development of quality assurance process and supporting materials required. Indicators to support interpretation of scores require further investigation.
Winsorised mean	Used in technical applications	Yes	Further development of quality assurance process and supporting materials required. Indicators to support interpretation of scores require further investigation.

Preliminary assessment

Usage of the statistical summary measures

A wide range of statistical summary measures are commonly used in economic analysis. It may be reasonable to assume that more commonly used statistical summary measures will be more widely understood. The ABS reviewed the usage of the median and the six alternative statistical summary measures in national and international income and economic statistics publications as part of the introduction to alternative statistical summary measures for use in the DMI methodology, presented to the [DMI Refinement Working Group](#) in January 2021. This review found that:

- the median is used extensively as a statistical summary measure in national and international income and economic statistics publications;
- the mean is also widely used as a statistical summary measure, even though it is not robust to extreme values. To partly counteract this, the mean is often provided alongside other summary statistics such as the median;
- like the mean, the first quartile is typically used alongside other percentiles (such as deciles or quartiles) as a statistical summary measure;
- the trimmed mean is used extensively to calculate Consumer Price Index (CPI) movements;
- the winsorised mean is used by the ABS for selected earnings, expenditure and employment statistics; and
- no examples of the use of the mid-hinge or tri-mean in national and international income and economic statistics publications were able to be found.

Further methodological assessment and recalibration

Further methodological assessment would be required if the trimmed mean or winsorised mean were to be used in the DMI methodology. This would require significant further investigations and decisions to determine the method and amount of trimming or winsorisation to be applied to school community income distributions.

In general, and as described in relation to Principle 4: Accuracy earlier in this paper, the advantage of using a trimmed or winsorised mean is to improve the robustness of a measure of central tendency compared with the mean. The decision regarding how to set the trimming threshold or winsorisation parameters aims to identify outliers in a distribution and remove or, in the case of winsorisation, reduce the impact of outliers on the summary measure. This is complex in the context of CTC as the point at which outlier values are identified in school income distributions is likely to differ, both for different schools and over time. An approach to trimming or winsorisation that is based on identifying and removing or dampening outliers would need to be recalibrated on a regular basis. This introduces challenges for implementation as well as for supporting stakeholders' understanding of the method.

Identifying and documenting an agreed process for setting and recalibrating the trimming parameter or winsorisation threshold each year would assist in supporting the ability of stakeholders to understand the methodology and interpret scores.

Data quality assurance process, supporting information and information to understand scores

For DMI scores based on the median, there is an existing data quality assurance process. As part of this process, each school's score is assessed using indicators of missingness, accuracy and volatility, and further review of certain scores is carried out based on these quality indicators.

The data quality assurance process, along with the methodology for calculating school scores, is described in '[A Data Quality Framework for the Australian Government's Direct Measure of Income for Capacity to Contribute]'. There is a range of additional information on the DESE website, including fact sheets about aspects of the school funding policy framework and the DMI methodology.

If an alternative statistical summary measure were implemented, the data quality assurance process and supporting information would need to be reviewed and updated. This includes not only the data quality indicators used, but also the thresholds at which the quality of scores are flagged for further review. For example, a score based on the first quartile may be more sensitive to missingness than a score based on the median or mean, and therefore if the first quartile were used in the DMI methodology, the threshold of indicators of missingness may need to be lowered to identify scores for further review.

Under the current methodology, additional information, such as coverage rates, annual DMI scores and income quartile information, is provided where it is safe to release, to assist stakeholders to interpret their school scores. The impact of adopting an alternative statistical summary measure on confidentiality requirements was described in ABS' paper for the DMI Refinement Working Group in February 2021. In addition to the confidentiality implications, further consideration would also need to be given to what additional information would help stakeholders understand their scores, should an alternative measure be adopted. In particular, the existing suite of indicators may be mostly satisfactory for alternative measures based on quartiles. In contrast, the standard deviation is often used to describe spread around a mean (or trimmed mean or winsorised mean), but using it can be misleading if the underlying distribution is skewed. As income distributions are often skewed, further investigation of meaningful indicators to support an understanding of mean-based statistical summary measures would be required if such a measure were adopted for use in the DMI methodology.

Appendix 1: Definitions of selected statistical summary measures

Statistical Summary Measure	Definition
Median (or second quartile, Q2)	After sorting all family incomes in ascending order, the middle value or the average of the middle two values (depending on whether the dataset has an odd or even number of incomes respectively).
First quartile (Q1)	Similar to the median, but is instead the income value that lies 25% of the way through the ordered set of income values.[6]
Mid-hinge	The average of first quartile and the third quartile (Q3)[7] of the school community's income distribution; that is: $$\text{Mid-hinge} = \frac{Q1+Q3}{2}.$$
Tri-mean	The weighted average of the first quartile (Q1), the median (Q2) and the third quartile (Q3) of the school community's income distribution; that is: $$\text{Tri-mean} = \frac{Q1+2*Q2+Q3}{4}.$$
Mean[8]	The sum of all income values in the school community's income distribution divided by the number of income values in the distribution. That is, if there are n income values in the distribution, then: $$\text{Mean} = \frac{\sum_{i=1}^{n} x_i}{n}.$$
Trimmed mean[8,9]	Calculated by discarding extreme income values (either a certain number or proportion) from one or both sides of the school community income distribution and only using the remaining values in the calculation of the mean. If the number of discarded values is m, then the trimmed mean is the sum of the remaining (n-m) values divided by (n-m), not n.
Winsorised mean[8,9]	Similar to the trimmed mean but the extreme income values are dampened or brought closer to a pre-determined, less extreme value and are still used in the calculation of the mean, instead of being discarded.

[6] Note that if the 25% mark falls between two income values in the distribution, then Q1 is an appropriately weighted average of those two income values.
[7] Note that Q3 is calculated in a similar manner to Q1, but is instead the income value that lies 75% of the way through the ordered set of income values.
[8] If the mean, trimmed mean or winsorised mean is chosen, a decision will need to be made regarding the treatment of negative incomes. For the purposes of the analysis described in this report, any negative incomes identified in the data will be treated as zero in the calculation of means.
[9] Note that if the trimmed mean or winsorised mean were to be used in the DMI methodology, significant further investigations and decisions would be required to determine the method and amount of trimming or winsorisation to be applied to school community income distributions.

Appendix 2: Methodology for calculating DMI-based CTC scores

The Direct Measure of Income (DMI) score

The DMI score is based on the median Adjusted Taxable Income (ATI) of each school community. It is created by:

- calculating the total income for each student by summing the incomes of up to two parents or guardians;
- identifying the median family income for each school; and
- converting the median incomes for all schools into DMI scores via standardisation[10].

The resulting DMI score represents the anticipated capacity to contribute of a school community, relative to other school communities.

The DMI score uses data from the Student Residential Address and Other Information Collection (the Address Collection) to identify the school community population. Income data is obtained via the Multi-Agency Data Integration Project (MADIP) and includes Personal Income Tax (PIT) data, payment summary data and low income concession card information from the DOMINO Centrelink Administrative dataset (formerly provided in the Social Security and Related Information) data. These data sources enable the DMI to use the most accurate and timely income data available for school communities. The PIT and payment summary income data are from the financial year ended 18 months earlier (table 2.1). The DOMINO data aligns with this reference period.

For a detailed description of the DMI methodology, see [www.education.gov.au/quality-schools-fact-sheets.](www.education.gov.au/quality-schools-fact-sheets)

The Capacity to Contribute (CTC) score

In 2020, a DMI-based CTC score is the average of DMI scores for 2018 and 2019. This is because the first Address Collection to which administrative data in MADIP were linked took place in 2018. From 2021, a DMI-based CTC score will be the average of the previous three years' DMI scores (table 2.1).

Table 2.1 Reference periods of income data used in DMI-based CTC scores.

		Address Collection and DMI score reference year			
		2018	**2019**	**2020**	**2021**
CTC Score	**2020**	2015-16 income	2016-17 income		
	2021	2015-16 income	2016-17 income	2017-18 income	
	2022		2016-17 income	2017-18 income	2018-19 income

[10] Standardisation is a common statistical process which converts a set of numbers, which may have any average and spread, into a pre-determined average and spread. It does not change the order of school communities in the distribution.

Appendix 3: Income imputation under the DMI methodology and a provisional refined strategy

Income imputation under the DMI methodology

The DMI methodology uses linked administrative data available via the Multi-Agency Data Integration Project (MADIP) to determine a value of Adjusted Taxable Income (ATI) for parents and guardians of students at non-government schools. Income imputation refers to the methods used to determine a value of ATI for those parents whose ATI is missing. An ATI value for a parent may be missing because:

- the parent did not link to MADIP; or
- the parent linked to MADIP but did not have an income value in the available administrative data sources[11].

A multi-stage approach is used to assign income values from available data sources to parent records.

- First, an income value is sought from Personal Income Tax (PIT), spouse-reported PIT, and payment summary data, in that order.
- Second, if the above data sources are unavailable and the parent has a low income concession card flag, then zero income is imputed for that parent.
- Third, if no income has been assigned, an income value is sought from the previous year's PIT, previous year's spouse-reported PIT, or previous year's payment summary data, in that order.
- Fourth, if no information is available for a parent, they are:
    - imputed zero income if the student has two parents in the Address Collection and the other parent has an income value; or
    - excluded from the calculation.

A more detailed description of this income imputation strategy is provided in '[A Data Quality Framework for the Australian Government's Direct Measure of Income for Capacity to Contribute'](#) (see page 10).

It should be noted that tax information is not expected to be available for all parents in the Address Collection in a given year. Some parents are not required to submit a tax return, for example if they earn no income or earn less than the tax-free threshold. ABS Survey of Income and Housing (SIH) data indicates that, among households with a student attending a non-government school, approximately 12% of parents and guardians earned less than the tax free threshold in 2017-18. Also, some parents may lodge their tax return too late for it to be included in the linked data. It

[11] The ABS is investigating potential refinements to the DMI methodology to reduce both of these sources of missingness as part of DESE's DMI refinement work program. Initiatives to improve linkage outcomes are described in 'Progress on CTC data linkage improvements' and initiatives to improve the income imputation strategy are described in 'Capacity to Contribute: Income imputation – discussion paper'. These papers were presented to the January DMI Refinement Working Group meeting and are available at: www.dese.gov.au/direct-measure-income-refinement-working-group.

should also be noted that parents who are missing in one year of the DMI score may nevertheless be represented in the CTC score, as a result of having income information in one of the three DMI scores included in the rolling average CTC score.

A refined income imputation strategy

As part of DESE's DMI refinement work program, the ABS is investigating two initiatives for potential refinement of the income imputation strategy:

1. the use of government payments data available via MADIP in the DOMINO Centrelink Administrative Dataset; and
2. the incorporation of modelled ATI estimates.

For a more detailed description of this work, see 'Capacity to Contribute: Income imputation – discussion paper', available on the DESE website.

Government payments data

Data relating to a range of government payments is sourced from the Department of Social Services (DSS) DOMINO Centrelink Administrative Dataset and linked to MADIP. This data is of interest because of its potential to complement PIT and payment summary data as sources of income values for parents across a range of income and labour force participation categories.

The ABS is working with DESE and DSS to define an estimate of ATI for parents using the DOMINO data available via MADIP, and incorporate this estimate into the income assignment strategy for the DMI methodology. As this work is ongoing, a provisional approach has been applied for the purposes of the analysis described in this paper; however, ABS notes that further improvements may be made to this approach.

Estimating ATI using statistical modelling

The ABS developed a statistical linear regression model, using data from the 2017-18 Survey of Income and Housing Basic Confidentialised Unit Record File (SIH CURF), to predict ATI for parents whose ATI is missing. Linear regression models use available information, or predictor variables, to estimate an outcome variable. For CTC, predictors of income – such as sex, occupation, whether a person has a tertiary education and whether a person received government benefits – available for a parent via MADIP, can be used to produce an estimate of the parent's income that takes that extra information into account.

Further refinement of the model described in the January 2021 DMI Refinement Working Group paper has since occurred.

Provisional refined income imputation strategy

Due to the ongoing work noted above, a provisional refined income imputation strategy has been used to calculate scores based on the median and alternative statistical summary measures, for the purposes of assessing the sensitivity of scores to changes in missingness and income imputation. This strategy consists of the following steps:

- First, an income value is sought from Personal Income Tax (PIT), spouse-reported PIT, and payment summary data, in that order.
- Second, if no income has been assigned, an income value is sought from DOMINO data.
- Third, if the above data sources are unavailable and the parent has a low income concession card flag, then zero income is imputed for that parent.
- Fourth, if no income has been assigned, an income value is sought from the previous year's PIT, previous year's spouse-reported PIT, previous year's payment summary data or previous year's DOMINO data, in that order.
- Fifth, if no income information is available for a parent and the parent record links to MADIP, a modelled income estimate is used.
- Sixth, remaining unlinked parent records are:
    - imputed zero income if the student has two parents in the Address Collection and the other parent has an income value; or
    - excluded from the calculation.