

Proceedings of Statistics Canada Symposium 2024: The Future of Official Statistics

Survey-admin Hybrid Measure of Persistent Child Poverty in New Zealand

by Keith McLeod, Adam O'Neill, Ryan Sutcliffe, and Robert Templeton

Release date: September 8, 2025



Statistics
Canada

Statistique
Canada

Canada

Survey-admin Hybrid Measure of Persistent Child Poverty in New Zealand

Keith McLeod¹, Adam O'Neill^{1†}, Ryan Sutcliffe¹, and Robert Templeton¹

Abstract

The Child Poverty Reduction Act (2018) outlines a need for the New Zealand Government to set three- and ten-yearly persistent child poverty reduction targets come end of 2024. In the absence of longitudinal survey data, a survey-administrative data hybrid method that will facilitate the production of these reduction targets and official estimates of persistent child poverty once reporting is required for the 2025/2026 financial year onwards is outlined. This hybrid approach leverages off the cross-sectional Household Economic Survey (HES), administrative-based beneficiary's family data, and recent advances developed for the construction of households within the Administrative Population Census (APC) at Statistics New Zealand. With increasing data collection challenges due to rising non-response and costs, this survey-admin hybrid method represents an alternative to longitudinal survey data collection, ensuring ongoing sustainable and quality statistics to produce persistent child poverty estimates.

Key Words: Longitudinal surveys; Non-response; Administrative data; Child poverty.

1. Introduction

The Child Poverty Reduction Act 2018 (the Act) was introduced to direct the measurement and reporting of the number of children in poverty in New Zealand within a given year. It outlines a requirement for the New Zealand Government to set three- and ten-year targets on four primary measures of child poverty, three of which are cross-sectional (single-year) estimates of poverty, with one measure of poverty persistence (Stats NZ, 2019).

To annually measure the progress towards meeting these targets the Household Economic Survey (HES), a key survey traditionally used to measure the economic well-being of New Zealand households, was optimised in 2018 for child poverty reporting (Stats NZ, 2019). HES is, however, a cross-sectional survey, selecting and interviewing a unique set of addresses every year. Thus, while HES could be used to inform on the cross-sectional measures of child poverty, a new method was required to measure both the cross-sectional and persistent child poverty rates for the 2025-2026 financial year onwards. A method would also be required in 2024 to outline the baseline persistent child poverty estimates required under the Act to facilitate government in setting three- and ten-years targets.

Attempts to model poverty persistence using a combination of cross-sectional survey data and modelled household dynamic changes (informed from the Household Labour Force Survey (HLFS)) were explored but deemed unreliable for official estimates (Bentley, A., Brooking, C., & Northern T., 2020).

¹Statistics New Zealand, 8 Willis Street, Wellington, New Zealand.

† adam.oneill@stats.govt.nz

Disclaimer: These results are not official statistics. They have been created from the Integrated Data Infrastructure (IDI) which is carefully managed by Stats NZ. For more information about the IDI 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.

A longitudinal survey called ‘Living in Aotearoa’ was established in 2022 that aimed to interview the same individuals over a six-year-period should they be living in New Zealand with the intent to produce both the cross-sectional estimates and persistent poverty measure. However, this was discontinued after two years due to ongoing challenges in data collection and participant retention, as well as increasing costs. These challenges meant alternative approaches to ensure ongoing sustainable and quality statistics of the persistent child poverty measure were required.

The use of administrative data for statistical purposes has been well recognised internationally as well as in Statistics New Zealand. The main benefits of using administrative data for official statistics include: reducing the collection costs, reducing respondent burden, and improved quality of certain variables. Income data and key demographic information has been derived from administrative data and combined with HES since 2018/2019 (Stats NZ, 2024a). The 2023 New Zealand Census also adopted a combined model which included using administrative data as alternative data sources (Stats NZ, 2024b).

The Integrated Data Infrastructure (IDI) managed by Statistics New Zealand offers opportunities for integrating administrative and survey information. The IDI represents a large research database containing de-identified microdata about people and households present in New Zealand, that has been sourced from government agencies (including Statistics New Zealand), and non-government organisations. This data includes routinely collected information as well as that collected from surveys. Data from these different sources are linked together through a central ‘spine’, facilitating societal and economic research with high coverage of the New Zealand population and accurate tracking of individuals (Black, A., 2016). Privacy and other risks associated with the IDI are robustly managed by Statistics New Zealand, with all personal identifiers being either removed or encrypted (Black, A., 2016; Stats NZ, 2017).

Despite the increased use of administrative data in supporting surveys, unfortunately no administrative sources currently directly inform on all individuals present within a household at a given date, making it difficult to measure child poverty which relies on knowing household income and composition (Stats NZ, 2024a). Several administrative sources do, however, provide information on where people live, as well as relationship information independent of address (Stats NZ, 2023a).

Predicting usual residence address from the multitude of available administrative data sources is thus a critical step in the formation of admin-based households. Several recent improvements have occurred in recent years, one of which uses a machine learning model to create admin-based households as part of the Administrative Population Census (APC): an experimental output created as part of Statistics New Zealand’s census transformation programme (Stats NZ, 2023a; Stats NZ, 2024c).

This machine learning model aims to select the best address out of all the addresses present within the administrative environment for a given individual. Trained off people responding to the 2023 Census, the model predicts 89.9% of people in the right household and has high coverage (99.8%) of the admin resident population (Stats NZ, 2024c).

This paper explores the use of administrative data to aid in the creation of a persistent child poverty measure in New Zealand, leveraging off recent admin-based household formation advances.

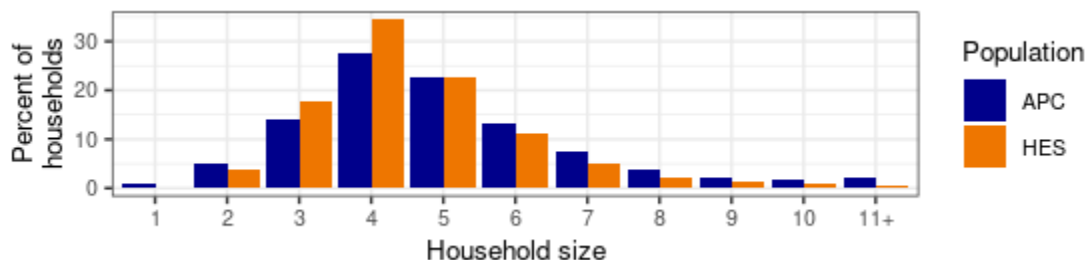
Note that the results presented in this paper were used to aid the government to set targets for child poverty before the end of 2024 as required by the Act. While Statistics New Zealand is confident in the methodology described below, the poverty rates presented are not to be regarded as official statistics. Statistics New Zealand will continue to investigate whether further improvements can be made that would be incorporated once regular persistent child poverty reporting begins in 2027. A more detailed version of this paper is available on the Statistics New Zealand website (Stats NZ, 2024d).

2. Admin-only household characteristics

Given the recent advances in household formation using administrative data (Stats NZ, 2024c), we first explored the potential for an ‘admin-only’ measure of persistent child poverty. To this end, we compared the household size and household equivalized disposable income distributions between HES and those formed from the APC.

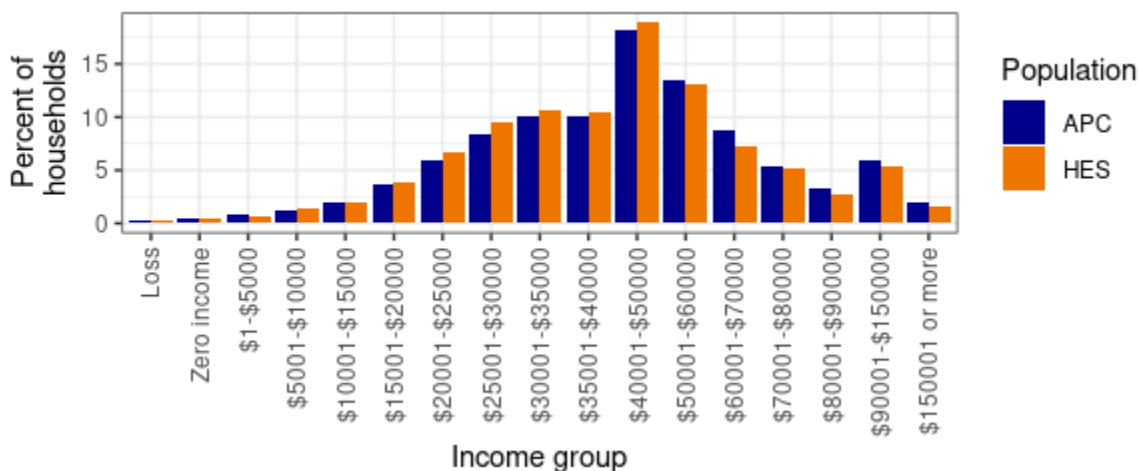
Admin-based households with children in the APC were more likely to contain six or more individuals compared with HES, which tended to have a greater proportion of three- and four-person households (Figure 2-1). This is consistent with what has been described in earlier reports when the APC households are compared to those in the 2018 Census (Stats NZ, 2023b). A proportion of admin-only households (0.8%) also only contained children (i.e., no adults).

Figure 2-1
Size distributions of households with children in admin-only (APC) and HES for the year 2021.



Despite the differences in household size distributions, the household equivalized disposable income (HEDI) distributions showed better agreement (Figure 2-2). HEDI is a crucial component in the calculation of a child’s poverty status within a given year. Children in households whose HEDI is lower than 60% of the national median are in poverty according to one of the measures outlined in the Act (Stats NZ, 2024a). HEDI is calculated using the disposable income remaining after the deduction of taxes and equivalizing it to adjust for household composition, incorporating the age of the household members. This equivalization allows for more valid comparisons between households, however, in doing so, also minimises the difference originally observed for changes in household size (Figure 2-2).

Figure 2-2
Admin-only (APC) and HES household equivalized disposable income distributions for the year 2021.



While promising, these results nevertheless suggest that further improvements are needed for an admin-only approach to be considered for official persistent child poverty reporting.

3. Survey-admin hybrid measure of persistent child poverty

Given the need to develop baseline estimates of persistent child poverty, we therefore sought to investigate whether improvements could be made to forming households with administrative data, by using it in combination with survey data. The sections below outline the proposed survey-admin hybrid method and analysis done to validate the method.

A child is defined to be in persistent poverty in this paper if their HEDI (before deducting housing costs) is below 60% of the national median in the current cross-sectional year of interest (Y , also termed base-year) and at least two of the previous three years.

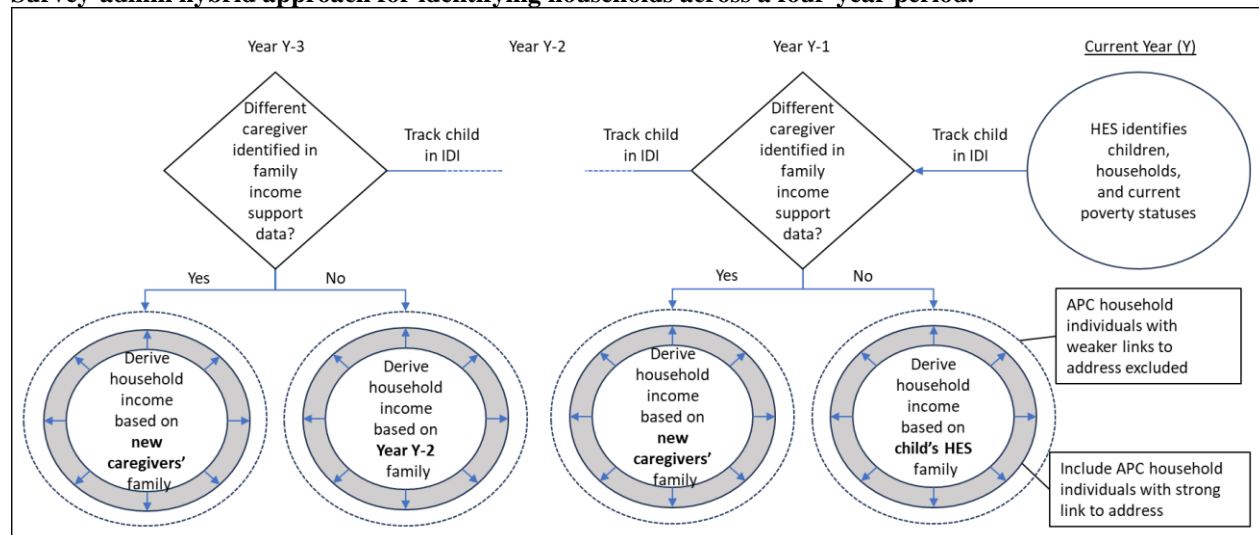
3.1 Survey-admin hybrid household formation

To measure persistent poverty, the poverty status of children needs to be known for each year in the four-year period. The specific survey dataset used in this hybrid approach is HES given its optimisation for reporting and capturing year-on-year changes in cross-sectional child poverty estimates since the Act was established in 2018. Within this hybrid approach, HES provides the sample of children, their household composition, HEDI, and thus their poverty status in the base year (Y). These children are then tracked back through administrative data for the three previous years to assess their presence or absence within the population (i.e., $Y-1$ to $Y-3$).

To facilitate household formation in the three previous years, we leveraged off family income support data within the IDI, household information present in the APC, and the HES survey itself (Figure 3.1-1). The IDI has high coverage (95%) of families who have experienced periods of low income within its family income support data (a combination of inland revenue and social welfare data sources associated with the “Working for Families” income support programme). Thus, for children in such families, we are likely to know their parents or caregivers in a given year. For children not present in family income support data we used their parents or caregivers as documented in HES. Additional household members are included into either of these families in a given back year if the machine learning model predicting admin addresses outlines the individual to reside at the same address as the caregiver with a probability of 0.8 or higher (a stronger criteria than that used to form the admin-only households in section 2) (Stats NZ, 2024c). The HEDI was then calculated using admin income data, before the child’s poverty status is determined for each year (Figure 3.1-1).

For children aged two years or younger who can not be observed across the four-year persistence window, we used their birth mother’s household information to inform on the child poverty status for the pre-birth years, allowing four-years of data for each child (Stats NZ, 2024d).

Figure 3.1-1
Survey-admin hybrid approach for identifying households across a four-year period.



3.2 Household size and composition

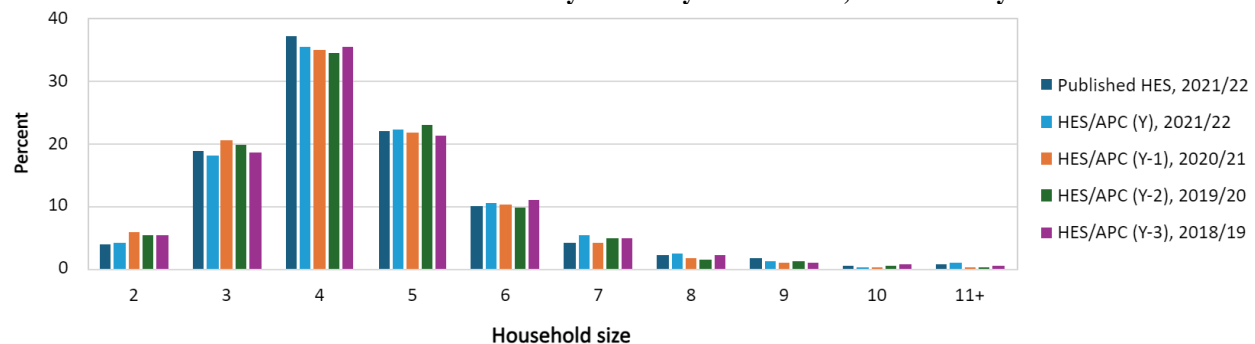
The survey-admin hybrid approach (referred to hereafter as ‘HES/APC’ method) resulted in consistent household size distributions to those observed in HES (Figure 3.2-1). To assess the quality of household composition, we assessed HES/APC households against households for the same children from the 2018 Census, excluding those census households that were administratively enumerated. This was done using children in HES samples, at intervals approximately 0-, 1-, 2-, and 3-years post-census (i.e., with the 2018/2019, 2019/2020, and 2020/2021 HES surveys respectively), giving quality estimates of the households at 1, 2, and 3 years prior to HES respectively. The HES/APC method was compared against two other household formation methods: the first of these was to simply use the HES household across each of the four years, essentially assuming that household composition is static across time (apart from the loss of recently born children), while the second involved using the households constructed for the APC. Household membership agreement was measured using the method described elsewhere (Gath and Bycroft, 2018).

The HES/APC longitudinal sample showed an increased percentage of ‘exact’ household membership agreement to the census for at least years *Y-2* and *Y-3* and the lowest ‘poor’ household membership agreement compared to if HES or the APC datasets are used in isolation. For *Y-1*, the survey-admin hybrid approach was comparable to that of HES with respect to the number of ‘exact’ household membership agreements. For the current year, *Y*, the HES households tend to match more closely with those of the census with a 67% ‘exact’ agreement, noting that differences in reference time-points and eligibility between the census and HES are expected to cause differences in household agreements within the same year. The HES/APC households were also superior when assessing consistency and bias in household size relative to the census (see Supplementary table 8 of Stats NZ, 2024d).

Household composition is critical for identifying the correct equivalized income for the household, and therefore each child’s poverty status. The results from these investigations showed that the HES/APC method provided more reliable households than the admin-only approach described in Section 2.

Figure 3.2-1

Household size distribution in HES and the survey-admin hybrid datasets, for the base year ended June 2022.

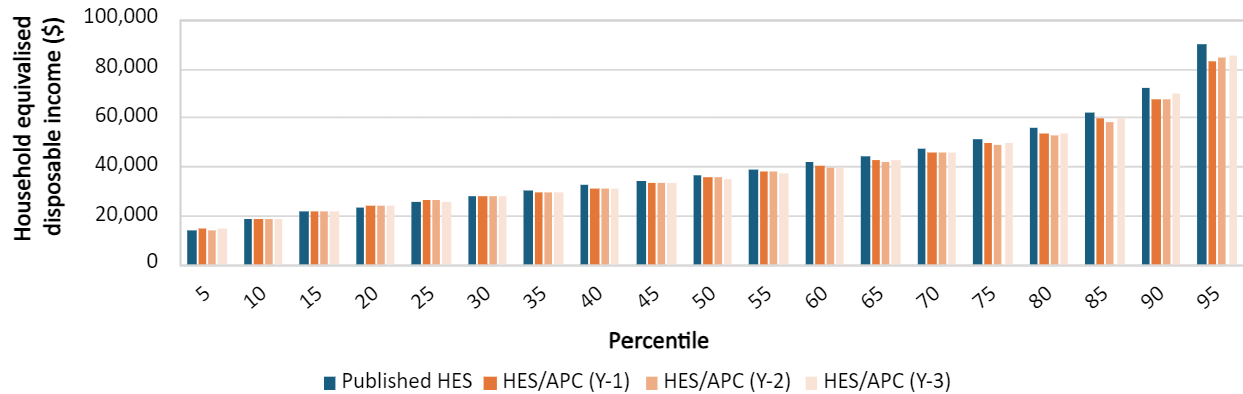


3.3 Household equivalized disposable income distributions

The HEDI distributions produced for years *Y-1* to *Y-3* under the HES/APC method (where income data comes from admin source alone) were strongly consistent with those produced in HES for those same reference years. Figure 3.3-1 shows the income value at each quantile produced using the HES year ended June 2019 data, and the *Y-1*, *Y-2*, and *Y-3* datasets of the subsequent longitudinal cohorts, such that incomes for 2019 are being compared. There is some discrepancy at the top end of the distribution, with HES quantile values tending to be higher than those produced by the *Y-1*, *Y-2*, and *Y-3* years. The impact of this difference is somewhat mitigated by using poverty thresholds set by the published HES data for each year, rather than deriving a new threshold value from the longitudinal dataset directly using admin-based incomes.

These results indicate that the HES/APC method can deliver HEDI and household structures within the *Y-1* to *Y-3* back-years like those found in the survey, supporting the use of this hybrid approach.

Figure 3.3-1
Annual income at selected percentiles in HES and each back year of the longitudinal cohort, year ended June 2019.

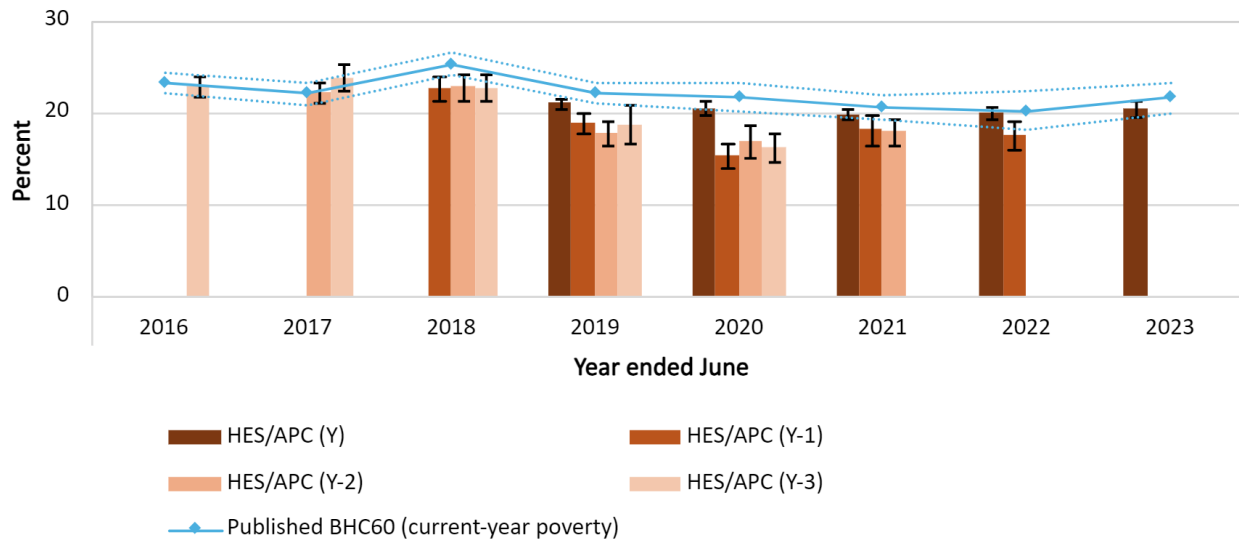


3.4 Cross-sectional poverty

We next sought to check if the cross-sectional poverty rates produced for the base-year (*Y*) by the HES/APC method were similar to those produced in HES. Only marginal differences in these cross-sectional poverty rates were expected given, (1) the longitudinal HES/APC and cross-sectional HES samples only differ with respect to those children who are usually resident in the country for the full four-year persistence window; (2) the proportion of children in each income band is nearly identical between the HES and the longitudinal samples (Stats NZ, 2024d). Figure 3.4-1 demonstrates the single-year poverty rates for each base-year (*Y*) within a longitudinal HES/APC cohort (in dark orange for 2019 to 2023) were broadly consistent with the rates published in HES (blue points), albeit marginally lower.

In addition, the cross-sectional poverty rates produced for the *Y-1*, *Y-2*, and *Y-3* back-years were also broadly consistent, though also tended to be lower, than the published HES rates for a given reporting period. This indicates good consistency in the income distributions produced by the HES/APC method across the four-year persistence window for each of the longitudinal datasets created (with base years from 2019 to 2023) (Figure 3.4-1).

Figure 3.4-1
Percentage of children in current-year poverty, for each back year of a longitudinal cohort, year ended June 2016–2023



3.5 Persistent poverty in other data sources

Although we do not have survey estimates of persistent child poverty over a four-year period for the years of interest (i.e., 2019 to 2023), we can leverage off surveys that repeatedly interview the same individuals across a two-year period to assess how well the HES/APC method compares. The Household Labour Force Survey (HLFS) is a quarterly survey measuring employment in New Zealand. Although it surveys the same *address* (rather than individuals) every quarter across a two-year period, we can calculate a two-year persistent poverty rate for those children who remained at the same address (about 70% of those initially surveyed). HLFS respondents were assigned income as per the admin-income method used for the HES/APC. We were also able to measure two-year persistent poverty using the two completed waves of the Living in Aotearoa Survey.

Figure 3.5-1 indicates similar two-year persistent poverty rates between the HES/APC, HLFS, and Living in Aotearoa survey, again suggesting that HES/APC methodology can be used to produce reliable estimates of persistent child poverty.

Figure 3.5-1
Percentage of children in two-year persistent poverty, by dataset, 2020-2023.

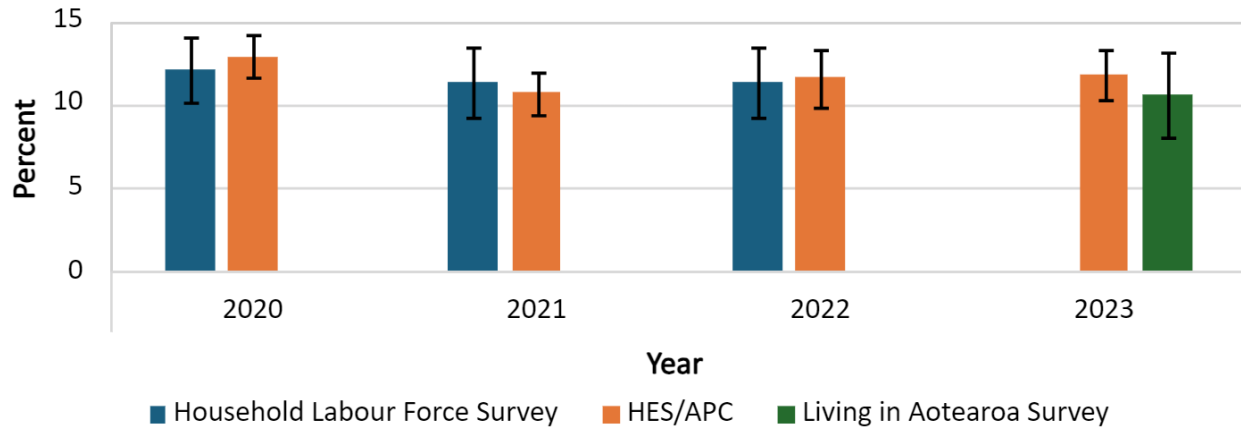
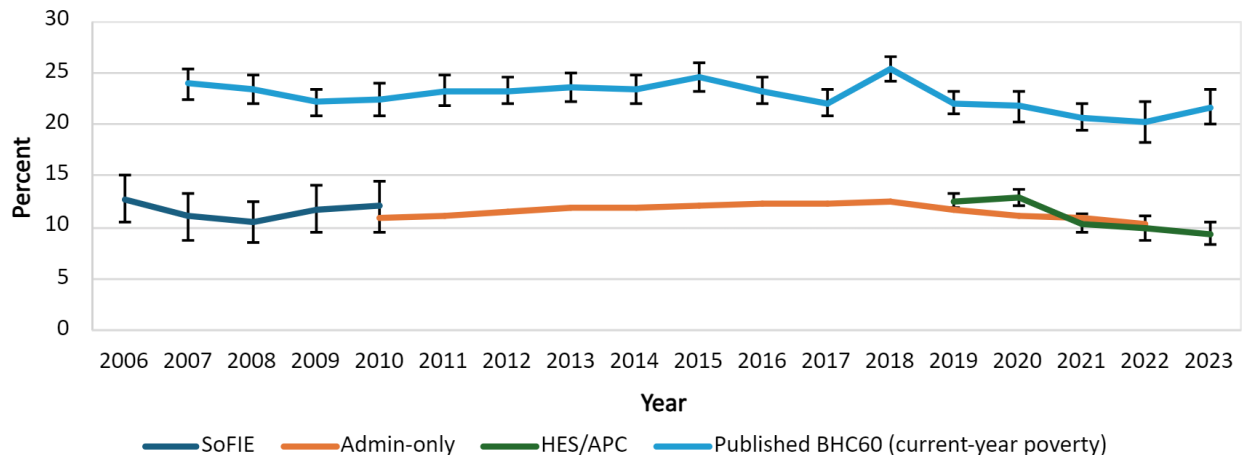


Figure 3.5-2 shows that the HES/APC persistent child poverty rates (over four years) are at a level that is consistent with those produced from an earlier Statistics New Zealand longitudinal survey, called the Survey of Families, Income and Employment (SoFIE), carried out between 2002 and 2010. The HES/APC persistent child poverty rates are also about half the magnitude of the equivalent cross-sectional rates, an expected feature given previous investigations (Jenkins & Van Kerm, 2014).

Figure 3.5-2
Percentage of children in poverty, by dataset, 2006-2023.



4. Discussion

With the discontinuation of the longitudinal Living in Aotearoa survey Stats NZ needed to develop an alternative means of measuring persistent child poverty in New Zealand. Recent advances in the use of admin data from the APC, supplemented with survey information, and family income support data to assess family relationships across time has demonstrated to be a viable option; enabling the production of baseline persistent poverty rates to facilitate the production of poverty reduction targets and provides confidence for official persistent poverty reporting.

In addition to the investigations discussed in this report, a more detailed version of the investigations carried out can be found elsewhere (Stats NZ, 2024d). Further technical documentation for persistent poverty will also be made available once regular reporting begins. This will describe any improvements, refinements, or changes that have been made for the final methodology. There is potential for these methodological refinements to impact the baseline estimates, and in that case the baseline estimates would likely be revised. For the purposes of setting initial persistent poverty reduction targets, however, Statistics New Zealand is confident in the quality of what has been produced and that it is fit for the core purpose.

References

- Bentley, A., Brooking, C., and Northern, T. (2020), “Panels of Poverty Persistence Year (PoPPY): a hybrid survey-administrative estimator of longitudinal income poverty”, unpublished report, New Zealand: Statistics New Zealand.
- Black, A. (2016), “[The IDI prototype spine’s creation and coverage. \(Statistics New Zealand Working Paper No 16–03\)](#)”. Retrieved from www.stats.govt.nz.
- Gath, M., and Bycroft, C. (2018), “[The potential for linked administrative data to provide household and family information](#)”. Retrieved from www.stats.govt.nz.
- Jenkins, S.P., and Van Kerm, P. (2014), “[The Relationship Between EU Indicators of Persistent and Current Poverty](#)”, *Social Indicators Research*, 116(2), pp. 611–638. Retrieved from www.jstor.org.
- Stats NZ (2017), “[Integrated Data Infrastructure: Overarching privacy impact assessment](#)”. Retrieved from www.stats.govt.nz.
- Stats NZ (2019), “[Expanding the Household Economic Survey to obtain good measures of child poverty](#)”. Retrieved from www.stats.govt.nz.
- Stats NZ (2023a), “[Experimental administrative population census \(third iteration\): Data sources, methods, and quality for household information](#)”. Retrieved from www.stats.govt.nz.
- Stats NZ (2023b), [Experimental administrative population census \(third iteration\): Information by variable](#). Retrieved from www.stats.govt.nz.
- Stats NZ (2024a), “[Child poverty statistics: Year ended June 2023 – technical appendix](#)”. Retrieved from www.stats.govt.nz.
- Stats NZ (2024b), “[Methodology for using admin data to count people in the 2023 Census](#)”. Retrieved from www.stats.govt.nz.
- Stats NZ (2024c), “[Predicting usual residence address from admin data in the 2023 Census](#)”. Retrieved from www.stats.govt.nz.
- Stats NZ (2024d), “[Developing a methodology to measure persistent child poverty using survey and admin data. Version 1](#)”. Retrieved from www.stats.govt.nz.