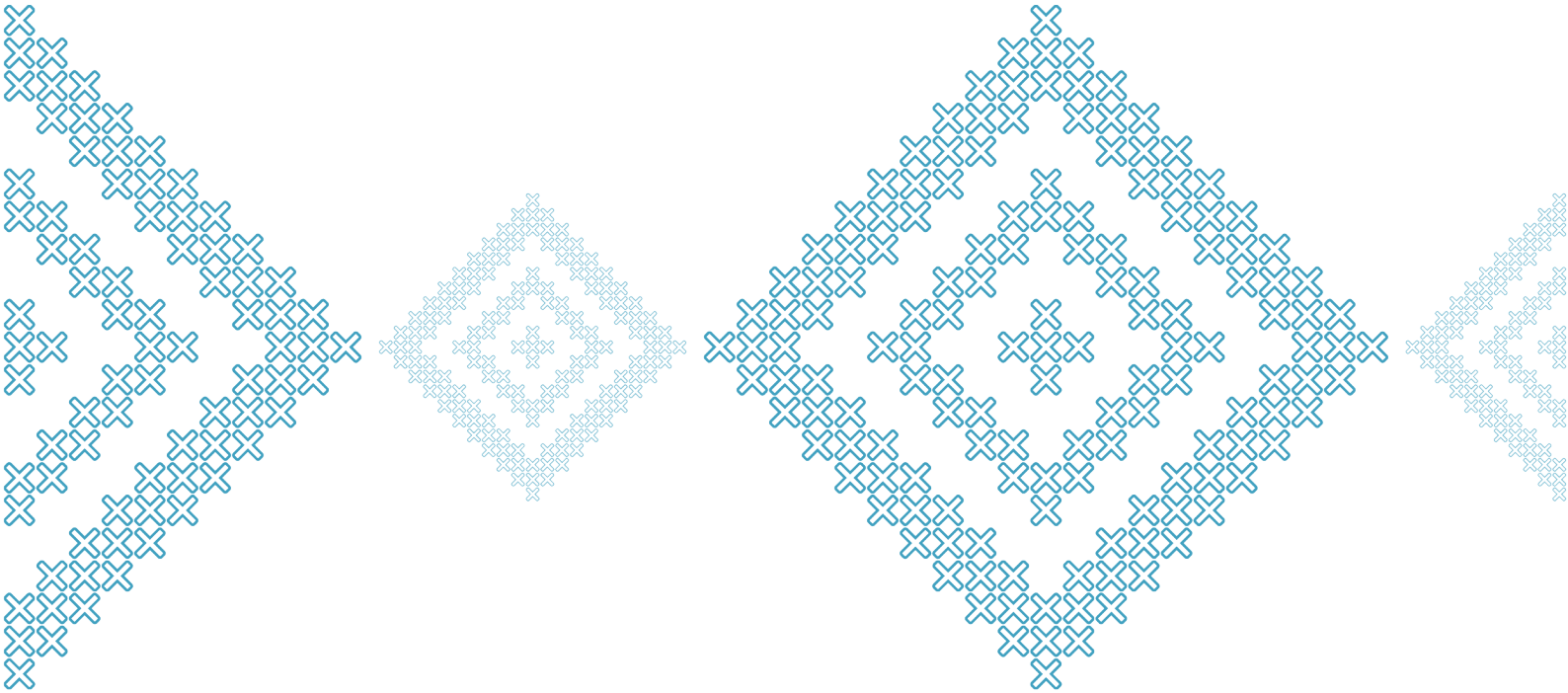
TE TAI ŌHANGA
THE TREASURY

REPORT

Tax and Welfare Analysis (TAWA) Model Methodology Report

August 2024



Te Kāwanatanga o Aotearoa
New Zealand Government

The Treasury URL at August 2024 for this document is https://www.treasury.govt.nz/publications/guide/tax-and-welfare-analysis-tawa-model-methodology-report

Contents

Introduction

Tax and transfer microsimulation models are powerful tools used by governments, research institutions, and policy analysts to assess the distributional impacts of tax and welfare policies. The TAWA (Tax and Welfare Analysis) model is one such model that is widely used in New Zealand. In this report, we provide an overview of the different stages involved in creating and using TAWA.

TAWA is a static or non-behavioural[1] microsimulation model used to analyse the New Zealand tax and transfer system. It is regularly used as part of the Government's budget process to estimate the costs and distributional impacts of potential personal tax and welfare policies (Nolan, Wang, and Stephens, 2022). Analysis is commissioned by other government agencies, the Tax Working Group, and the Welfare Expert Advisory Group and the model has been used in academic collaborations. TAWA is also used to understand the distribution of household incomes, net worth, and expenditure and to estimate future child poverty rates.

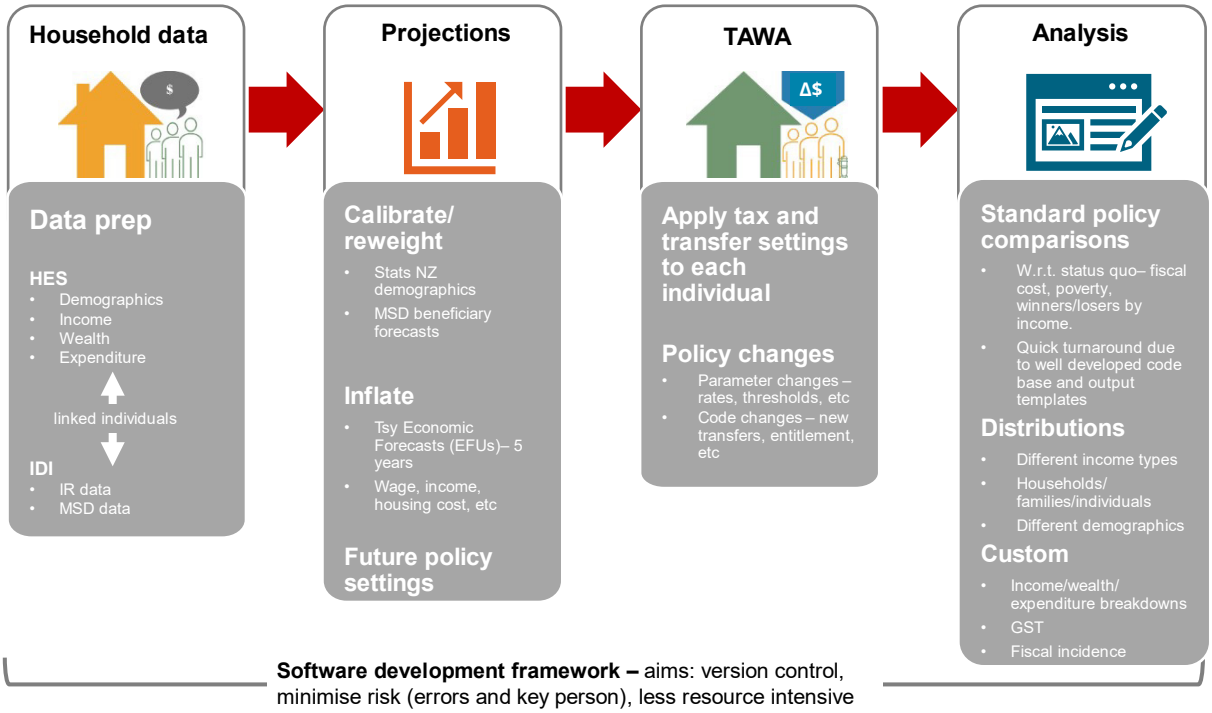
TAWA has made significant contributions to child poverty reports (New Zealand Treasury, 2023), targets, and research (Stephens, 2022), and has provided advice on welfare reform, including changes to Working for Families and child support pass-on. It has also been used to understand the distribution of wealth (Symes, 2022) and expenditure (Wang, 2022), model the distributional impacts of the emissions trading scheme (Davis, Hart, and Stubbing, 2024), model Social Unemployment Insurance, investigate income adequacy of NZ Super for housing costs, and analyse fiscal incidence in New Zealand (Wright and Nguyen, 2024).

[1] No allowance is made for the possible effects of tax and transfer changes on a modelled individual's consumption plan or labour supply (Creedy et al., 2002).

Overview of the TAWA modelling process

Figure 1 shows the four stages of the TAWA modelling process: creating an input dataset, projecting the data into the future, modelling the tax and transfer system, and processing and analysing outputs.

Figure 1 – Overview of the TAWA modelling process



Creating an input dataset

The TAWA input dataset requires data that provides a representative sample of New Zealand households. The base data for TAWA is sourced from the Household Economic Survey (HES), which provides information on family and household structure, individual characteristics, location, housing costs, and non-taxable income[2]. This data is then linked to the Integrated Data Infrastructure (IDI) to incorporate wage/salary income, self-employment income, receipt of core benefit and NZ superannuation, and other non-modelled payments such as youth payment and student allowance.

Key input data sources

Key input data sources for the TAWA model are described in Table 1 below.

Table 1 – Key input data sources

Name	Description	Source
Individual characteristics	Includes age, sex, ethnicity, disability status, labour force status, studying status, highest qualification	HES survey data
Household identifier	Defines members of a household	HES survey data
Dependent flag	Identifies whether an individual is a dependent child. A dependent child means a person who is not in a marriage, civil union, or de facto relationship, and who: • is aged 15 years or less, or • is aged 16 or 17 years and is not financially independent, or • is aged 18 years and is a person for whom a tax credit is allowed[3]	Derived from survey and admin data
Family identifier	Defines members of a family based on WFF rules	Derived from survey and admin data
Housing costs	Housing costs (eg, rent, mortgage payments, and council rates), which are used to define after housing costs poverty measure.	HES expenditure data
Wage / salary income	Total wage and salary income by period.	IDI – EIE tables[4], eg, EI_E_OctDec2021_20220111
Self-employment income	Total self-employment income by period.	IDI – IR3 table[5] ird_rtns_keypoints_ir3_202111 Note survey data used as a proxy if admin data missing

[2] Note taxable income collected from the HES survey may be used as a proxy if income information is missing due to lags in filing return dates or missing data linkages.

[3] For more information see pages 2883, 2901, and 2926 of the Income Tax Act 2007.

[4] For EIE data dictionary see IDI Tax & Income data (IR) - Stats NZ DataInfo+

[5] For IR3 data dictionary see IDI SLA IRD CDE IR3 Keypoints (Published) - Stats NZ DataInfo+

Name	Description	Source
Interest, dividend, and rent income	Total income by period for interest, dividend, and rent	IDI – IR3 table ird_rtns_keypoints_ir3_202111 Note survey data used as a proxy if admin data missing
Core benefit receipt	Flags if individual is in receipt of Job Seeker Support (JSS), Supported Living Payment (SLP), Sole Parent Support (SPS) in a given period	IDI – MSD spells and tier1 expenditure tables
Super receipt	Flags if individual is in receipt of Super in a given period	IDI – MSD tier1 expenditure table
Accommodation supplement (AS) receipt	Flags if individual is in receipt of AS in a given period	IDI – MSD tier2 expenditure table

TAWA target population

As it is based on the HES, the target population for the TAWA input dataset is the usually resident individuals of private dwellings in urban and rural areas in North Island, South Island, and Waiheke Island of New Zealand. The people out of scope include overseas visitors, residents of hotels, motels, hostels, long-term resident of hospitals and/or prisons and persons in home for the aged (including rest homes).[6]

TAWA input dataset structure

The TAWA input data is in CSV format. Each row represents a single person for a given period, with 24 periods in a year. The dataset includes family and household identifiers, individual characteristics (eg, age), income components (eg, income from wages), benefit receipt flags, and housing costs.

[6] The design of the HES is described in Child poverty statistics: Year ended June 2021 – technical appendix | Stats NZ

Projections

In the second stage, the input data is projected into future years to estimate the impacts of different policies. Three mechanisms are used: inflation, calibration, and policy parameter changes.

Inflators

The different components of household income, such as wages, self-employment, and investment income, are inflated into future tax years using forecasts. The same is done for housing costs, which are a significant expense for many households. These inflation rates are obtained from Treasury's economic forecast updates (EFUs) and are based on economic assumptions about future growth rates.

TAWA has seven inflators: Earnings, CPI, Interest, Dividend, HPI, Rent, and HLPI-CPI. The levels for each of these are calculated from historical data and future forecasts. The growth rate of these levels is then used to inflate the values of particular variables from the survey base year into a future year.

Table 2 – TAWA inflators

Inflator	Input forecast	Frequency	Adjustments	Applied to
Earnings	Total weekly earnings	Quarterly	4-quarter average	Income (wage/salary, self-employment income, parental leave, redundancy, other taxable, and other chargeable)
CPI	Consumer Price Index	Quarterly	4-quarter average	Non-modelled benefits, student allowance & scholarships, child support income, other non-taxable income, and regular income
Interest	90-day Bank Bill Rate (BBR)	Quarterly	3-quarter lagged, 5-quarter average of 90-day BBR	Interest, trust, and other non-dividend investment income
Dividend	Net Operating Surplus & Working age population	Annual (June-year)	Divided by 4-quarter average of working age population (ie, net operating surplus per-working-age person)	Dividend income
HPI	CoreLogic house price index	Quarterly	4-quarter average	Mortgage principle, house insurance, rates
Rent	Rent Forecast	Quarterly	4-quarter average	Rent costs; Total housing costs for Accommodation Supplement
HLPI-CPI	Household living-cost price index (Income quantile 1, all groups less housing and household utilities group and credit services subgroup)	Quarterly	4-quarter average, extended by CPI growth rate as projections not available	AHC-50 child poverty fixed line medians

Calibration

To ensure that the model accurately represents the TAWA target population and future projections, the final survey weights are calibrated to agree with demographic benchmarks. These benchmarks are listed in Table 3. They are based on Stats NZ demographic projections and MSD forecasts of the number families receiving core benefits.

Table 3 – TAWA benchmarks

Source	Frequency	Benchmark	Input forecasts
MSD	Monthly – averaged over tax year	Core benefit recipients	JSS, SPS, SLP recipient forecasts
Stats NZ	Annual – growth rates used to project HES benchmarks into the future	5-year age group by sex	Projections of resident population of New Zealand and labour force, by single year of age and gender
		Number of people by region	Subnational population projections
		Māori by broad age group	National ethnic population projections
		Number of households by household type	Subnational household projections
		Number of 2-adult (and non-2-adult households by region	Subnational household projections

Using the GREG calibration technique[7], final weighted sample counts are calibrated to the benchmarks list above. This technique also ensures that families and individuals within a household all have the same weight.

Replicate bootstrap weights are created to allow for the calculation of standard errors and confidence intervals. The procedure used to create the bootstrap weights is based on the Rao and Wu rescaling bootstrap described in Mach, Saïdi, and Pettapiece (2007).

[7] See Deville and Sarndal (1992).

Policy and indexation changes

Forecasted changes to transfer rates, whether through indexation or policies, are incorporated into the model as modelling parameters. The detailed indexation methods and forecasts used for the indexation are listed in Table 4.

Table 4

Transfer	Forecasts	Detailed method
Core Benefits	Average wage	Indexed by the annual growth of net December quarter Average Ordinary Time Weekly Earnings (AOTWE).
Family Tax Credit (FTC), Best Start Tax Credit (BSTC)	CPI	The indexation only happens once cumulative inflation (CPI) breaches 5%. The indexation is applied on the following 1 April using the cumulative inflation from the starting point to the previous September quarter. The cumulative starting point is then reset to this September quarter.
Superannuation	Average wage or CPI	Indexed by the annual growth of December quarter CPI, with a subsequent adjustment to ensure the total amount received by a couple who both qualify is at least 66% and at most 72.5% of the net December quarter Average Ordinary Time Weekly Earnings (AOTWE). A single person then receives 65% of the total couple rate, or 60% if they live with other adults.

Policy changes can also have flow-on impacts on the transfer rates indexation. For example, a low-tier tax threshold change can impact the average wage, affecting the transfers indexed by it.

Underlying assumptions to the projection method

The projection process necessarily relies on several modelling assumptions, which should be considered when using the model, particularly in times of increased uncertainty. These assumptions include the following.

- The input data is assumed to provide an accurate representation of current New Zealand households.
- The model assumes that market incomes and housing costs will increase uniformly for each household in line with the appropriate forecast.
- By using reweighting to account for forecasted changes to the number of benefit recipients, the model assumes that future beneficiaries will have the same characteristics as existing beneficiaries. In addition, this means the entire working population is down-weighted (up-weighted) when the forecast number of benefit recipients increases (decreases).
- The model assumes that there are no newly eligible recipients of main benefits following benefit changes, and that all existing recipients continue to have 100% take-up of their initial benefit. While this assumption simplifies the model and aligns outputs to observed benefit recipients, it may not accurately reflect the impact of future policy changes in some cases.

- Since TAWA is a static model, it is assumed there are no behavioural responses. This means that people who are working will stay working, for the same number of hours, and people who are not working will not start working.

- Accommodation Supplement take-up rates are assumed to follow a logistic regression model which is trained on initial actual take-up rates. TAWA then assumes that the average take-up rate will always hold. If the modelled average take-up rate increases in a reform, then more families will be given AS. If the modelled average take-up rate decreases in a reform, then some families will have AS removed. This symmetric treatment may not be an accurate description of the actual behavioural response, as families may be less likely to give up their AS entitlements than they are to start receiving them.

- Population forecasts use the median projections from Stats NZ. Actual population growth rates may be higher or lower.

Modelling the tax and transfer system

In the third stage, once the inflators have been applied to adjust for the tax year of interest, tax and transfer policy settings are applied to the individuals in the input dataset. Tax and welfare reforms are implemented through parameter changes (rates, thresholds, etc) and code changes to the procedures (new transfers, entitlement, etc). Procedures are stored in the TAWAproc library. The procedures are run by TAWArun, which looks at the input and output attributes associated with each function and determines the appropriate run order. The procedures can be categorised as follows.

- ***Utility procedures*** – These are used within functions and, for example, apply abatement schedules and calculate thresholds.
- ***Income procedures*** – These calculate different income components that are needed when calculating different payment rates and eligibility.
- ***Transfer payments*** – These functions determine eligibility and payment amounts for transfers such as benefits, superannuation and Working for Families (WfF) tax credit.
- ***Tax and ACC*** – These functions apply income tax schedules and the ACC levy.

Policy changes

Policy changes are modelled by either adjusting parameters or changing the TAWA code. Parameter changes include changes to payments (amount paid to eligible recipients), income thresholds (income levels relevant to tax and welfare policies), and rates (eg, tax and abatement rates). Code changes are required for structural changes. For example, to introduce a new tax or welfare payment, to change eligibility settings, or to change abatement structures.

Modelling limitations

The TAWA model simulates a complex tax and welfare system with input data limitations, so some policies cannot be modelled. These include temporary additional support (TAS) and childcare assistance. Other policies such as child support pass-on can be modelled, but data limitations lead to increased uncertainty.

In addition, the HES collection process means that we must assume that there are no changes to family and household structure over the tax year. This limitation becomes important when comparing TAWA outputs with administrative summaries from Inland Revenue and MSD.

	Best Start	Family Tax Credit	Minimum Family Tax Credit	In-work Tax Credit	Independent Earner Tax Credit	Paid Parental Leave **(not modelled in TAWA)**	Winter Energy Payment	Jobseeker support	Sole parent support	Supported living payment	Young parent payment **(not modelled in TAWA)**	Youth payment **(not modelled in TAWA)**	Accommodation supplement	Income related rent subsidy **(not modelled in TAWA)**	Temporary additional support **(not modelled in TAWA)**
Acronym		FTC	MFTC	IWTC	IETC	PPL	WEP	JSS	SPS	SLP	YPP	YP	AS	IRRS	TAS
Description	Payment for a child's early years, up until the child turns 3.	Income-tested payments available to families with children.	Tops up after-tax income of working families with children not receiving a benefit.	Income-tested payment to working families not receiving a benefit, with children.	An entitlement for individuals who earn between $24,000 and $70,000.	Paid leave for working mothers and adoptive parents to care for their new-born or "adopted" child.	Winter period payment to help with household winter heating costs for those on a main benefit, NZ Super or Veteran's pension.	Provides income support while people search for work.	Provides income support for single parents with young children	Provides income support for people who have a permanent and severe health condition, injury or disability, or for people who are caring for someone who requires full-time care.	Provides income support for 16-19 year olds with children. 16/17 year olds only eligible if parents/guardian are eligible for family tax credit, or can't live with parents/guardian	Provides income support for 16/17 year olds who can't live with their parents or guardian and aren't supported by anyone else.	A weekly payment which helps people with their rent, board or cost of owning a home.	Rent which is subsidised to make accommodation more affordable for those on low incomes.	A weekly payment which helps someone who can't meet their essential living costs from what they earn or from other sources.
Policy rationale	Give all families extra support in the first year of a child's life and targeted support for low- and middle-income families with children aged 1 or 2.	Aims to increase family income to help with the costs of raising their children and maintain a standard of living.	Guaranteed minimum level of income for those in work. More than what they can receive on a benefit- better off in work.	To encourage parents to work by boosting the earned incomes of low to middle income families to ensure they are better off in work.	To increase incentives to work. The lower threshold was set at just under the then-minimum wage full time salary. N/A for beneficiaries, NZS, WFF recipients.	Promotes attachment to work while allowing time to bond with new child. Linked to 1 year unpaid parental leave.	To help with household heating costs during the winter period for those currently on benefits.	There are work expectations, which may be full-time, part-time, or deferred (if due to a health condition, injury, or disability).	There are part-time work expectations when the youngest child turns 3. Sole parents shift to JSS when youngest child turns 14.	Recipients are assessed as being severely restricted in their capacity to work for at least 2 years.	Recipients have education obligations. Evidence shows that education attainment is one of the best pathways to lowering the likelihood of being unemployed, earning a higher income and improving social outcomes. This is why payments are abated to allow for part-time work, but discourage full-time work to allow for young people to study.		Both the AS and the IRR subsidise the housing costs of low to middle income earners in rental accommodation. IRRS is generally significantly more generous than the AS.		Provides temporary last resort financial assistance to alleviate financial hardship for people whose essential living costs cannot be met from their income and other resources, while ensuring that applicants take reasonable steps to reduce their costs or increase their income.
Intended recipients	Families with children aged 0-3.	Low- and middle-income families with children, including those receiving benefits.	Working families with children on very low incomes, not receiving a benefit. Targeted to those who at risk of going onto benefits.	Working families with children, not including those receiving a benefit.	Individuals working and earning within a target income band. Note: Individuals partnered/living with someone with a higher income also receive this payment.	Working parents of new-borns, including those receiving benefits	Individuals on a main benefit, NZ Super or Veteran's Pension.	Out of work and seeking employment.	Sole parents.	Individuals with a health condition, injury or disability, or looking after someone who needs full-time care.	Young people with children and without support, or in low-income families	Youth without support.	Individuals with high housing costs relative to income.	Individuals living in state houses.	Individuals with very high housing, disability, or employment-related costs, anyone with income shortfalls (primarily paid to meet housing costs)

Modelling take-up

Not every family who appears eligible for a transfer will receive it. To accurately model transfer income, TAWA needs a method to determine which families receive payments. There are two sources of information to draw on when modelling transfer receipt:

- Microdata that indicates which individuals receive each transfer. This is not available for all transfers, and where it is, relying completely on this information prevents accurate modelling of policy changes that make new individuals eligible for a payment.
- Microdata that describes an individual's economic situation in enough detail to determine whether they are eligible for a transfer. This includes, for example income, hours worked, and the number of children in the family. Relying completely on this information to determine transfer receipt, which we call the full take-up assumption, has some problems, all of which lead to overestimates of the number of recipients:
  - It neglects the role of choice in determining whether an individual applies for and receives the transfer. There are reasons such as social stigma and inconvenience that may stop people applying for a transfer.
  - Some people may not be aware of their eligibility for opt-in payments.
  - The microdata does not completely describe eligibility criteria for the payment. For instance, families with cash assets over a certain threshold are not eligible for accommodation supplement payments. Currently we have no microdata on cash assets so we can't apply this rule in TAWA; if we were using the full take-up assumption we would be incorrectly distributing the payment.

The approach TAWA takes to the take-up of transfers differs between payments, due to the availability of microdata and the extent to which partial take-up is a problem.

For core benefits (JSS, SPS, SLP), only individuals that receive benefits in their administrative data can be given the benefit in TAWA, but eligibility is also determined in the model, so given policy settings may mean they have no modelled benefit income. For example, their IRD income data coupled with the model abatement settings may "abate away" their benefit. The limitation of this approach is that if policy settings change in a way that makes new families eligible for the benefit, the model does not assign benefits to any new families.

At the other extreme, for Working for Families (FTC, IWTC, MFTC, Best Start) we assume full take-up. Receipt is completely determined by TAWA, based on legislated eligibility criteria and the available microdata. If policy settings change to make more families eligible, then the families are modelled to receive the payment, even if it is very small and the realistic likelihood of the family applying for it is relatively low. This would tend to make TAWA overestimate the number of WfF recipients. However, there are other factors, such as the assumption of fixed family/household structure and age, that make comparison with administrative data, and correcting this bias, difficult.

The problem of partial take-up is particularly acute for Accommodation Supplement (AS). Many non-beneficiaries that look eligible in TAWA do not receive the payment in reality. Assuming full-take up in TAWA would significantly overestimate the number of recipients and overstate the impacts of changes to AS policy. Instead, we apply a hybrid probabilistic method to assign payments. First, we fit different logistic regression models to beneficiaries, superannuitants, and other families, to give the probability of an eligible family taking up the payment. The predictors for these models are family chargeable income and the amount of AS they are eligible for. After a probability is assigned to each family, the expected number of AS recipients is obtained by summing these probabilities over all eligible families.
The model approximately satisfies this expected number of recipients by first assigning the payment to families that receive AS in the administrative data. If the expected number of recipients is greater than the number of administrative recipients, we randomly sample the remaining eligible families according to their probability of receipt to make up the difference. Similarly, if the expected number of recipients is less than the administrative recipients, then the probabilities are used to randomly select recipients to stop receiving the payment.

Improving the modelling of take-up for both WfF and core benefits, potentially using a similar method to that used for Accommodation Supplement, is an option for future work. Such work is largely dependent on the availability of TAWA consistent administrative data with which to fit logistic regression models and determine eligibility.

Processing and analysing outputs

In the final stage, TAWA generates a wide range of outputs to produce aggregated statistics. These statistics are then confidentialised by use of suppression and rounding. The standard outputs can include standard policy comparisons (fiscal cost, poverty, winners/losers by income) and distributions by income types. Results can be summarised at household, family, or individual levels. Custom outputs can include income, wealth, or expenditure breakdowns. The model is also used to produce child poverty projections over the forecast horizon.

Example outputs

- ***Fiscal cost*** – Provides the cost for each reform being modelled, relative to the current legislated settings.

- ***Distributional analysis*** – Provides information on how each reform being modelled impacts the population across the income distribution. This is seen in the winners and losers output, which shows how much the reforms impact families and households across the income distribution. When considering households, we use household equivalised disposable income (HEDI), which considers incomes after accounting for taxes and transfers and uses equivalisation to allow for comparisons across different household compositions. For families we consider winners and losers grouped by family taxable income levels.

- ***Poverty impacts*** – This output shows the impacts of the reforms on families, households, adults, and children in poverty. We primarily use the fixed-line AHC50 and moving-line BHC50 poverty measures. More details of the poverty measures can be found in Stephens (2022).

Research outputs

TAWA outputs are flexible and have been used for a range of research including the following.

Fiscal Incidence

Wright and Nguyen (2024) report fiscal incidence results for the 2019 tax year. They augmented results for household disposable incomes produced using TAWA with estimates of the consumption taxes that households pay, and the cost to the government of education and health services that they receive. They present results for the distributions of household market, disposable and final income – and the components of income support, income and consumption taxes, and in-kind benefit spending that contribute to them.

Wealth and Effective Average Tax Rates

By combining income, wealth, and expenditure data using HES18 (wealth, income) and HES19 (expenditure), Ching, Reid, and Symes (2023) estimated effective average tax rates (EATR, tax divided by income) of families across the income and wealth distributions, using a series of increasingly comprehensive EATR definitions. The most comprehensive definition includes capital gains and imputed rent as income in the denominator, as well as GST paid on consumption in the numerator. Capital gains and imputed rent are modelled using the wealth data linked to average rates of return and average rents by area and bedroom size. GST is modelled by using the method of multiple imputation by chained equations (MICE) to impute the GST distribution of households in HES19 onto the households of HES18.

Inflation

Wang (forthcoming) examined the distributional impact of inflation on New Zealand households using the TAWA dataset for income and demographic information and expenditure data from Household Economic Survey (HES). In this analysis, expenditure was inflated by CPI from to 2020-22 and incomes were projected using TAWA. The analysis shows that higher-income deciles experienced large increases in their residual income in 2020, with a positive correlation between the decile and the magnitude of the increase. This trend reversed somewhat in 2021, with lower-income deciles seeing larger increases than the higher deciles. In 2022, all deciles experienced declines except the 8th and 9th deciles.

Child poverty

Stephens (2022) outlined an experimental approach that used TAWA to provide insights into three different indicators of poverty, identifying seven different categories of children in poverty and describing the characteristics of children in each group. This exploratory analysis confirmed that the relationship between material hardship, income, and housing costs is complex. For some of the identified categories there is a direct relationship between low incomes, either before or after housing costs, and material deprivation. However, for several categories low incomes do not correspond to deprivation and vice-versa.

Output benchmarking

The benchmarking of TAWA's model outputs against other data sources is important to the model's development and validation. Although there are currently no routine validation processes, the following comparisons have provided important insights and improvements.

- Child poverty forecasts are compared to Stats NZ child poverty statistics every year. While differences in time periods and data sources can be expected, these comparisons have identified discrepancies. We are working through these differences with Stats NZ.
- The sensitivity of TAWA's outputs to economic forecasts has also been highlighted through comparisons over time. There are unavoidable uncertainties in economic forecasts and changes between EFUs can have a significant impact on modelled results, especially when predicting future poverty levels.
- TAWA's outputs have also been benchmarked against MSD research using administrative data. This benchmarking has demonstrated the importance of take-up, particularly of WfF payments. The challenges of modelling family and income dynamics have also been highlighted, along with the inherent uncertainty in prospective estimates that rely on past year's data and forecasts for economic and employment prospects.

Standardising the benchmarking and validation of TAWA outputs is another option for future work. Regular comparisons could inform model adjustments ensuring that that TAWA remains accurate and relevant over time.

How the code is organised and developed

The tawaverse

With the exception of the routines that encode NZ’s tax and transfer system, which reside in the tawaproc git repository, all of the code is deployed in several R-packages that perform distinct parts of the modelling process: from the creation of an input database to the generation of analytical reports. The organisation of code into separate packages, as opposed to one large monolithic package, is motivated by the principles of separation of concerns and functional cohesion. The packages are responsible for different parts of the modelling process, which run in different environments, and their code is modified at different frequencies. As these packages change relatively infrequently, they can be installed into the users R-library once and only updated when required. The latest source packages and an installation script that installs them all are in the R-packages/tawaverse directory in the IDI project folder. Table 5 below lists the tawaverse R-packages that most TAWA users will need and describes their purpose.

Table 5 – The R-packages that most users will need

Package name	Purpose
TAWArun	Runs the TAWA model by: • inflating the input data to the required tax-year • determining the order to run the TAWAproc procedures from the TAWAproc procedure input/output variables • saving the requested output variables (with full AS take-up assumed)
TAWAadminAS	Once TAWA has run, produces a new set of outputs with stochastically modelled partial AS take-up
TAWApost	Produces aggregates from the AS assigned TAWA outputs for release. Includes plots, tables and Rmarkdown report generation
TAWAcameos	Produces artificial TAWA input with pre-defined family/household types. Used for testing.
TAWAtemplates	Allows the creation of an R project in RStudio that includes template files for the standard modelling process.

The TAWAproc procedures that encode the tax and transfer are not in an R-package because both the method for determining run order, and the frequent need to make code changes for policy reforms, are not compatible with the functions being in an R-library. Instead, the latest version of the procedures is checked-out into a directory, and the location of this directory is supplied to TAWArun.

Development practices

To help maintain the quality of the tawaverse code, we use version control, code review, and automated testing.

To help ensure correct functionality, unit tests and end to end tests accompany each of these packages, with the aim of achieving test coverage that is as high as reasonably possible.

All the code mentioned above is version controlled using git, with the repositories mirrored in Treasury's Azure DevOps environment and the Stats NZ IDI Gitlab server. All development occurs outside the IDI and git bundles are sent into the IDI when required. We use the "GitHub Flow" method for feature branching and code review, where the master branch is always considered production-ready, and feature branches must have all tests passing and generally two code reviews before merging.

Access to TAWA

TAWA is available to trusted Stats NZ researchers via the Integrated Data Infrastructure (IDI). Please contact the Treasury if you are interested in using the model in your work.

References

Ching, B., Reid, C., & Symes, L. (2023). *Tax and Transfer Progressivity in New Zealand: Part 2 Results.* New Zealand Treasury Analytical Note.

Creedy, J., Duncan, A.S., Harris, M. & Scutella, R. (2002). *Microsimulation modelling of taxation and the labour market: the Melbourne Institute tax and transfer simulator.* Cheltenham, UK: Edward Elgar Publishing.

Davis, C., Hart, B., & Stubbing, B. (2024). *Household cost-of-living impacts from the Emissions Trading Scheme and using transfers to mitigate regressive outcomes.* New Zealand Treasury Analytical Note.

Deville, J.C. & Särndal, E.C., (1992). *Calibration Estimators in Survey Sampling.* Journal of the American Statistical Association.

Mach, L., Saïdi, A., & Pettapiece, R. (2007). *Study of the Properties of the Rao-Wu bootstrap variance estimator: What happens when assumptions do not hold.* In Proceedings of the Survey Methods Section, SSC Annual Meeting.

New Zealand Treasury. (2023). *The Wellbeing Budget 2023*. Retrieved from https://www.treasury.govt.nz/publications/wellbeing-budget/wellbeing-budget-2023-support-today-building-tomorrow

Nolan, P., Wang, Y., & Stephens, M. (2022). *Modelling Child Poverty and Wellbeing: the Treasury's TAWA microsimulation model*. Policy Quarterly, 18(3), 58-63.

Symes, L. (2022). *House prices and wealth inequality in New Zealand*. Policy Quarterly, 18 (3).

Stephens, M. (2022). *Insights from New Zealand child poverty data.* New Zealand Treasury Analytical Note.

Wang, Y. (2022). *An expenditure-based approach to measuring child poverty in New Zealand*. Policy Quarterly, 18 (3).

Wang, Y. (forthcoming). The Distributional Impact of Inflation on New Zealand Households: Evidence from Data. New Zealand Treasury Analytical Note.

Wright, T., & Nguyen, H. (2024). *Fiscal incidence in New Zealand: The effects of taxes and benefits on household incomes in tax year 2018/19.* New Zealand Treasury Analytical Note.

Annex

Using the IDI

As the TAWA model makes use of Stats NZ's Integrated Data Infrastructure (IDI), the following IDI disclaimer is always included with any release TAWA results

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