

The Treasury

Miscalculation of Disposable Income Release

February 2014

Release Document

www.treasury.govt.nz/publications/informationreleases/disposableincome

Key to sections of the Official Information Act 1982 under which information has been withheld.

Certain information in this information release has been withheld under one or more of the following sections of the Official Information Act, as applicable:

- [1] 9(2)(a) - to protect the privacy of natural persons, including deceased people
- [2] 9(2)(g)(i) - to maintain the effective conduct of public affairs through the free and frank expression of opinions

Where information has been withheld, a numbered reference to the applicable section of the Official Information Act has been made, as listed above. For example, a [2] appearing where information has been withheld in a release document refers to section 9(2)(g)(i).

In preparing this Information Release, the Treasury has considered the public interest considerations in section 9(1) of the Official Information Act.

Review of accuracy of Taxwell dataset and associated documentation provided to Statistics NZ

Executive Summary

This report details the work undertaken to check the Taxwell/HES (Household Economic Survey) data and metadata sent to Statistics New Zealand (Statistics NZ) for the six years 2006/07 – 2011/12. This review was triggered by the discovery of an error in the metadata that resulted in the Accommodation Supplement being double-counted in some data supplied to Ministry of Social Development (MSD) and others. A separate review covers co-production processes used in the preparation of this data with Statistics NZ and Treasury.

Several methods have been used to check the data and associated documentation.

- The data has been checked against the documentation provided, which found only insignificant errors.
- Averages of all the numerical variables have been compared year-on-year, which found two further significant problems causing double counting of Working for Families income and underestimation of modelled Accommodation Supplement.
- The Taxwell procedure code and the Taxwell files used to prepare the dataset for Statistics NZ have also been compared year-on-year. This check found the original problem again, insignificant errors, expected bug fixes and changes made to Taxwell code for legislation changes. This check would not find any changes in legislation that have not been included in Taxwell, but Taxwell is frequently used to analyse significant policy changes and policy teams within Treasury will be aware and have advised of those changes.
- A similar check on the dataprep code was attempted. The significant time required to complete this check was considered to be unproductive, as there are compensating checks done elsewhere in this review and plans to completely revise the dataprep process.
- Documentation illustrating the links between the income variables sent to Statistics NZ has been prepared, and sent to Statistics NZ for comment and comparison with their use of the variables.
- Statistics NZ have also sent Treasury documentation on how they combine variables to confirm they are using them correctly.

The review focuses largely on the year-on-year changes that were made since the production of the initial 2006/07 Taxwell dataset. While a detailed review of Taxwell code is out of scope for this report, testing undertaken both in the initial development of the Taxwell model, the file conversion program HesToTaxml, and the 2006/07 dataset and yearly tests in the annual update of HES satisfy us that there are unlikely to be significant problems with the parts of the Taxwell code in regular use. Statistics NZ also have their own testing procedures before releasing data to Treasury and other users.

As well as checking this data, a brainstorming session was held to come up with ideas for effective and efficient ways to check the data and metadata sent to Statistics NZ each year. These have not yet been evaluated, but those that can be quickly and easily implemented and will be included in the checks done on the revised 2007/08 dataset are included in this report. Further evaluation will inform an improved QA process to be implemented in time for the 2012/13 dataset and beyond.

This review has identified 3 significant problems with the Taxwell/HES data and metadata, affecting 2007/08 and 2009/10 – 2011/12. The seriousness of these problems meant the implications for policy advice also needed to be checked. While some distributional analyses for the Tax Working Group and Budget 2010 were affected, recalculations on a corrected dataset show any effects were small and most unlikely to have changed policy advice.

Recommendations:

Immediate:

1. Prepare and provide to Statistics NZ a corrected dataset and documentation for 2007/08 by late February 2014
2. Prepare and provide to Statistics NZ corrected data dictionaries for all six years by late February 2014

Medium term:

3. Evaluate the brainstorm ideas prepared for this review by early March 2014 and implement as appropriate
4. Develop a new, fully documented dataprep process for preparing the 2012/13 and future datasets by late March 2014 and invite Statistics NZ to provide QA
5. Use the new dataprep code and revised data from Statistics NZ to prepare revised datasets and documentation to Statistics NZ for all six historical years, incorporating the fixes described in this review by April 2014
6. Prepare and provide to Statistics NZ documents that describe the year-on-year changes for each of the five years from (and including) 2007/08 by April 2014
7. Develop an annual process for checking published legislation changes that affect the modelled variables in Taxwell by June 2014.

Review of accuracy of Taxwell dataset and associated documentation provided to Statistics NZ

Introduction

On 10 December 2013 it was discovered that data on household incomes double-counted accommodation supplement for low-income individuals and families. The data originates from the Household Economic Survey HES (which is unaffected), modelled through Taxwell by the Treasury, and then provided to Statistics New Zealand (Statistics NZ). This materially affects the data used for some measures of poverty and housing affordability for low income families.

A separate review relates to the annual process for preparing this data (refer *Review of Communication and handover processes relating to the 2009-10 and subsequent transfer of HES data from Treasury to Statistics New Zealand*). By comparison, this review considers the reliability of the Taxwell datasets and associated documentation provided to Statistics NZ.

Objectives of the review

The objectives of this review are to:

1. Assess the accuracy of the Taxwell dataset and associated documentation provided to Statistics New Zealand
2. Identify opportunities for improvements to verification procedures.

The full Terms of Reference for this review can be found at [\(ToR for taxwell output review:2809940\)](#)

Context

Given the nature of the issue that has arisen the review focuses largely on the year-on-year changes that were made since the production of the initial 2006/07 Taxwell dataset. While a detailed review of Taxwell code is out of scope for this report, we have considered the sources of assurance related to implementation of the Taxwell Model. The purpose of this was to satisfy us that there are unlikely to be significant problems with the parts of the Taxwell code in regular use that hasn't been changed since implementation.

This includes internal sources of assurance such as testing undertaken both in the initial development of the Taxwell model, the 2006/07 dataset and yearly tests in the annual update of HES. We also considered other external sources of assurance such as the various checks and reviews by other agencies on the datasets provided by Taxwell.

Some examples of these checks include:

Initial testing:

- Distributional comparison of Taxwell output with Taxmod output (the predecessor to Taxwell) on weights, disposable income, gross income and tax¹.
- Comparison of Taxwell weights and weights used by Statistics New Zealand²
- QA of SAS procedures and code³
- Comparison of poverty analysis on equivalised household disposable income.⁴

Yearly checks

- Taxable income totals compared to Budget forecast for personal tax
- Taxable income distributions compared to administrative data from Inland Revenue
- Aggregate totals for benefits, Working for Families, and other key variables are regularly compared to administrative data and Budget forecasts
- SAS weighting programme - the weighted totals add to the benchmarks, and this is confirmed in the program, and from 2011/12 has also been checked in Excel.

Where discrepancies were identified, records show these were discussed with a group of Taxwell “wise heads” to determine the cause, and either accept, remedy, or investigate it further.

Identifying what testing of Taxwell and the surrounding processes has occurred in the past has been challenging, and highlights the need to improve documentation and record-keeping.

Checks

Seven checks were made for this review, as follows:

1. Compare the historical data with the data dictionaries
2. Year-on-year comparison of averages
3. Compare Taxwell project files year-on-year
4. Compare Taxwell procedure files year-on-year
5. Compare Taxwell dataprep files year-on-year
6. Map the relationship between income variables provided to Statistics NZ
7. Confirm Statistics NZ using variables correctly

These checks are intended to provide overlapping coverage of places where problems could occur, and significant problems should be detected by several of these checks. The impact of each problem found is based on a combination of what the impacted variable is used for, how much it is used, and the size of the effect on the variable relative to sampling error within a reasonable margin.

¹ See for example <http://treasury-intranet/SiteDirectory/taxwell/Wiki%20Pages/Taxwell%20vs%20Taxmod.aspx>

² See for example ([HES07: analysis of income by family type:1091662](#))

³ See for example log file ([QA of Dataprep programs:985795](#))

⁴ See for example ([Initial Taxwell analysis of equivalised household disposable income:1025406](#))

This review has taken a risk-based approach, where checks have been focussed on areas that are considered higher risk. In addition, lower risk areas that were easier to check have also been checked. As problems due to changes between years had already been detected, this was considered the riskiest area and hence the area of biggest focus. Other checks focussed on methods that would find material problems with the data and metadata, irrespective of the cause, as these are the final products of the process.

Check 1: Compare the historical data with the data dictionaries

In this check, [1] made a careful comparison between the data dictionary and the data sent to Statistics NZ for all six years (2006/07 – 2011/12). Explicitly, he checked the following:

- Compared the variables names in the data dictionary to the dataset sent to Statistics NZ.
- Accuracy of the cover sheet in the data dictionary.
- Line-by-line check of the data dictionary to check the variable descriptions are accurate.
- Checked the variable names from Taxwell match the alias mappings given to Statistics NZ.
- Checked the modelling methodology still matches the metadata, in regard to dependent children, principle earner definitions and so on⁵.

A number of minor discrepancies were found, mostly typos, as described in Table 1 in the Appendix. This check also found again the incorrect statement in the data dictionary about disposable income and accommodation supplement which triggered this review. None of these discrepancies require new datasets to be sent to Statistics NZ, but new data dictionaries will need to be provided.

Check 2: Year-on-year comparison of averages

For all variables where a meaningful average could be calculated, [1] calculated the average for each survey and compared them with previous and subsequent years. Any large variations or unexpected results such as unexpected zeros are recorded in Table 2. [1] then investigated these variations for their impact and their cause⁶.

This check found a number of variables that were unexpectedly zero, or significantly smaller or larger than expected. Several had a common cause (those relating to benefits and Family Assistance), and this has a significant impact on 2007/08 data, with approximately \$900million in total disposable income double counted. This was due to a miscoding of income in the original HES data provided by Statistics NZ, who have since revised their data. Records are insufficient to determine if/how Statistics NZ raised this error with Treasury, and whether consideration was given to revising the

⁵ The details of this check are recorded in [\(ASSESSMENT OF ACCURACY OF STATEMENTS MADE IN THE DATA DICTIONARY SENT TO STATISTICS WITH TAXWELL OUTPUT:2829467\)](#)

⁶ These investigations are recorded in [Q:\Taxwell drive prototype - Do not edit\Taxwell\Output\Data requests\2013-12 - Double Counting AS\3 - Analysis\Summary.xlsb](#) and [Q:\Taxwell drive prototype - Do not edit\Taxwell\Output\Data requests\2013-12 - Double Counting AS\3 - Analysis\WeightedTotals.xlsx](#); however access to these files is restricted by law as they contain unit record data.

2007/08 Taxwell/HES data. To fix the data problem, the revised data from Statistics NZ will need to be used to prepare a new version of the 2007/08 data to Statistics NZ and their clients. This further highlights the need for improved record-keeping and documentation.

There was also a problem with the HHAnnualRent variable where the Treasury dataprep process did not pick up most rental expenditure in 2007/08. This variable is used in the modelled Accommodation Supplement variable (which had previously been considered unreliable) and resulted in an underestimate of the amount of Accommodation Supplement by about \$650million. This problem can be treated in the same way and at the same time as the Family Assistance problem described above.

The significance of these problems meant the implications for policy advice also needed to be checked. This was done by tracing all saved instances of Taxwell that used the 2007/08 database, and also by the tax and welfare teams checking their advice from that time period for anything that might contain Taxwell calculations using the affected database. Many of those calculations were unaffected as they either did not use the affected variables (using, for example, taxable income instead) or were differences where both problems cancelled out (e.g. in difference calculations). However, some distributional analyses for the Tax Working Group and Budget 2010 were affected. These were recalculated, and the corrections were small and most unlikely to have changed policy advice.

This check also found that the proportion of investment income attributed to PIE sources changed every year, from 0% to 100% and back. This only has a very small impact (0.1%) on average disposable income, but could become a problem if someone was using this data to study this part of the tax system. In the short term, this can be corrected by adding a note to the data dictionaries, but in the medium term a consistent value should be chosen for this parameter and revised datasets prepared.

The other issues found in this check do not have a significant impact, but can be found in Table 2 in the Appendix. However, they should be corrected if new datasets for the affected years are provided to Statistics NZ.

Check 3: Compare Taxwell project files year-on-year

[1] compared each Taxwell project file (.tpf) used to provide data to Statistics NZ to the tpf for the previous year. Some differences are expected, such as labels for the relevant year but only unexpected differences are included here. Most of these differences were also picked up in Checks 1 or 2, including the original problem of including the Accommodation Supplement in the calculation of disposable income. None of these differences require new datasets to be sent to Statistics NZ in the short term.

This check has been logged in [\(Log of tpf comparisons 2006/07 - 2011/12:2830634\)](#) and the results listed in Table 3 in the Appendix.

Check 4: Compare Taxwell procedure files year-on-year

This involved checking for differences between each year in the procedure files (.polml) used by the Taxwell project files, done by [1] Some differences are

expected and are due to code fixes, legislation changes, and model enhancements. This report only notes those that materially affect the data sent to Statistics NZ, excluding those due to legislation changes. None of these differences require new datasets to be sent to Statistics NZ, but if new datasets are sent they should be corrected.

This check has been logged in [\(Log of Taxwell procedure comparisons 2007 - 2011:2830637\)](#) and the results listed in Table 4 in the Appendix.

This check would not identify any instances where legislative changes have *not* been incorporated. However, analysis from Taxwell output for policy purposes is checked with the relevant policy teams in Treasury who have knowledge of the significant legislative changes – where significant changes have been omitted these would be expected to be identified through this process. Taxwell is also frequently used for policy analysis of significant changes before those changes are passed into law. Minor legislative changes may have been missed, for example the removal of the child tax payer rebate totalling \$11million, and these are unlikely to have a large effect on the calculation results. Should they have a large effect, this will be noticed when preparing data for Budget publications each year, where the results are compared against the official Budget forecasts that are produced independently within Treasury.

To prevent such issues going forward, an annual process to check published legislation changes should be put in place.

Check 5: Compare Taxwell dataprep files year-on-year

This involved checking for differences between each year in the data prep files, done by [1]. However, this proved to be a large piece of work and beyond the scope of this review. The way in which the dataprep code is written makes it difficult to determine in retrospect which code changes are significant and which are not. Furthermore, any documentation to explain changes was not found. However, removing this check from the suite of checks documented in this review is unlikely to be a problem, due to the overlapping nature of the checks. Any significant problems that this check would have detected such as the problem with HHAnnualRent will also have been detected by other checks, particularly Check 2 which looked at the overall averages of each variable. Furthermore, Statistics NZ compare the counts and sums of income variables of its original HES data file with Treasury's file used in the Taxwell model. This would also pick up any significant issues if the wrong input file was used in the Taxwell model.

However, to prevent further problems in the longer term, it is recommended that the dataprep code be rewritten and thoroughly documented. This work was underway before the problems that triggered this review were identified, but should now be given a higher priority to ensure it is ready for use with preparing the 2012/13 HES data. This will take a similar length of time as attempting to understand the existing data prep code, and will give better long term results.

The results of this check have been logged in [\(Log of dataprep comparisons 2008/09 - 2011/12:2830649\)](#)

Check 6: Map the relationship between income variables provided to Statistics NZ

This check involved mapping the relationships between income variables as calculated in Taxwell, and providing to Statistics NZ for them to check their usage of those variables. This was done by [1], and no further issues were found.

Documentation of links between income variables:

[\(Income Variables for Stats from Taxwell:2812829\)](#)

Email reply from Statistics NZ:

[\(RE: disposable income \[1\]\)](#)

[:2823924\)](#)

Check 7: Confirm Statistics NZ using variables correctly

[1] at Statistics NZ sent a description of how Statistics NZ were combining Taxwell/HES variables to create the statistics requested by the OECD, which [1] checked was a correct use of the data. No issues were found.

Variables created by Statistics NZ:

[\(Re: Fw: disposable income \[1\]\)](#) [:2823933\)](#) [\(RE: Fw: disposable income \[1\]\)](#) [:2823931\)](#)

Other points to note

Source data

Each year Statistics NZ quality assures the Household Economic Survey (HES) data as part of standard output procedure, before the HES microdata set is provided to users such as the Treasury. This includes checking:

- that systems have worked as intended
- the validity of extreme values in the data
- that unusual movements in aggregate figures can be adequately explained⁷.

also checked to ensure respondent confidentiality is maintained.

File Conversion

Evidence of testing of the HesToTaxml program which converts the data into the appropriate format for Taxwell gives us comfort on the file conversion process. In addition, problems such as changes in numbers of records, variable sums, and character truncations would be picked up at several stages of the process, including by Statistics NZ when they check the data we send to them against the original HES data.

Improvements to Verification Process

[1] organised a brainstorming session (using the method Brainwriting 6-3-5) on Friday 17th January 2014. [1] from Statistics NZ and [1]

[1] also participated. [1] from MSD,

[1] from Statistics NZ, and [1]

were invited but could not attend and instead contributed ideas via email. The results from this brainstorming session also included many ideas relevant to the overall process of preparing data for Statistics NZ.

⁷ However, as the HES is a sample survey, it is subject to both sampling and non-sampling error.

The question used for the brainstorming session was:

How do we ensure that the Taxwell / HES data and metadata that Treasury send to Stats is correct and is used correctly?

At the time of writing this review, the ideas resulting from the brainstorming session had not been evaluated. However, those that can be quickly and easily implemented and will be included in the checks done on the revised 2007/08 dataset are listed here.

- Basic checks that total type variables based on raw HES data add to the sum of their components.
- Sample-check variables that have missing values or are zero to ensure they should be missing or zero.
- Sample-checks on types of households or families e.g. average income of all families with two adults and two children.
- Compare totals or averages for all variables where this can be meaningfully calculated and compare with other years.
- Ensure more than one person looks at the checks, especially where judgement required.
- Ensure changes are documented in the metadata.

Further evaluation of the brainstorm session will inform an improved QA process to be implemented in time for the 2012/13 dataset.

Conclusions

This review has found several problems with the Taxwell/HES data and metadata for the six years investigated that require immediate correction. It will be straightforward to create and provide to Statistics NZ a corrected data dictionary for each year, which will decrease the risk of incorrect use of the data in the future. The 2006/07 and 2008/09 – 2011/12 datasets do not have any significant problems and do not need immediate revision. However, the two significant problems with the 2007/08 dataset means this dataset should be revised immediately and the revised version provided to Statistics NZ. In addition, a review of policy advice that used Taxwell and other calculations from the original, incorrect 2007/08 dataset has been performed and found policy advice would not have changed.

However, this review has also shown that there are substantial risks to the accuracy of future datasets should the existing dataprep process continue to be used. As such, it recommends increasing the priority of work already underway to develop new dataprep code so that the new code can be used to prepare Taxwell/HES 2012/13 data.

This would also be a suitable opportunity to recreate the datasets for 2006/07 – 2011/12 using the latest revised datasets from Statistics NZ and correcting all the minor issues found in this review. These would then be provided to Statistics NZ to replace the existing datasets for those years. These datasets would need new data dictionaries, and in addition a document describing what has changed between each year will be a useful addition to the metadata.

Finally, a brainstorming session on ways to improve the quality of the dataprep processes was held. The results of this session have yet to be evaluated, but appear

to include a number of useful suggestions for improving the process, and particularly the checks performed on the data before it is finalised. This review recommends evaluating these ideas before redeveloping the dataprep process, and implementing the suitable ideas as appropriate.

This review has identified 3 significant problems with the Taxwell/HES data and metadata, affecting 2007/08 and 2009/10 – 2011/12. Implementing the recommendations from this report will address these problems and, it is unlikely there are any remaining significant problems. There are a number of smaller problems that have been identified that should be corrected in the longer term, and a number of process problems that should be improved, but it is unlikely that these will have a significant effect on the existing data.

Recommendations:

This review makes seven recommendations. The first two recommendations are remedial work required immediately to fix known existing problems. The rest of the recommendations are intended to improve the historical data and help prevent future problems of this type.

Immediate:

1. Prepare and provide to Statistics NZ a corrected dataset and documentation for 2007/08 by late February 2014
2. Prepare and provide to Statistics NZ corrected data dictionaries for all six years by late February 2014

Medium term:

3. Evaluate the brainstorm ideas prepared for this review by early March 2014 and implement as appropriate
4. Develop a new, fully documented dataprep process for preparing the 2012/13 and future datasets by late March and invite Statistics NZ to provide QA
5. Use the new dataprep code and revised data from Statistics NZ to prepare revised datasets and documentation to Statistics NZ for all six historical years, incorporating the fixes described in this review by April 2014
6. Prepare and provide to Statistics NZ documents that describe the year-on-year changes for each of the five years from (and including) 2007/08 by April 2014
7. Develop an annual process for checking published legislation changes that affect the modelled variables in Taxwell by June 2014.

Review prepared by [1]
Treasury on 20 February 2014.

Analyst, Economic Research and Analysis, The

Appendix

Table 1: Comparison of dataset with data dictionaries

Discrepancy	Years Affected	Impact	Notes	Action
Data dictionary explicitly and incorrectly states Accommodation Supplement is excluded from Disposable Income calculation	2009/10 2010/11 2011/12	Significant	This is the discrepancy that triggered this review and other related work. Actions described here are limited to those needed to correct the data and metadata	Send corrected data dictionaries to Statistics NZ for affected years Statistics NZ will need to provide corrected data to their users
Principal earner definition inaccurate in data dictionary	2006/07 2007/08 2008/09 2009/10 2010/11 2010/12	Very minor	Dataprep adds wages to self-employed income before taking absolute value, but should probably be other way around. Also not quite correctly defined in data dictionary	Correct in data dictionary before revised version sent to Statistics NZ If revised data set prepared, correct in data set and metadata
Definition of dependent person inaccurate in data dictionary	2006/07 2007/08 2008/09 2009/10 2010/11 2010/12	Very minor	Some of the conditions are not listed in the metadata and the metadata is not worded correctly as there is potential ambiguity on the application of and/or in the first sentence.	Correct in data dictionary before revised version sent to Statistics NZ
In the variable descriptions, the code associated with line 45 should be attached to line 44 in the variable list	2006/07 2007/08 2008/09 2009/10 2010/11 2011/12	Very Minor	No investigation required	Correct in data dictionary before revised version sent to Statistics NZ
Variable labelled PTCAmuont in data and PTCAmount in data dictionary	2006/07 2007/08 2008/09 2009/10 2010/11 2011/12	Very Minor	Statistics NZ will need to be explicitly informed when this variable name is corrected so they can change any code they use with this variable.	Correct in data dictionary before revised version sent to STATISTICS NZ If revised data set prepared, correct in data set and metadata
Line 23 of Taxwell Notes in the data dictionary should	2006/07 2007/08 2008/09	None	No investigation required	Correct in data dictionary before revised version

be Assistance not Assistance	2009/10 2010/11 2011/12			sent to Statistics NZ
The variables in the data dictionary are sorted by household/family/person level except for FamCountPeople. This order does not match the order in the dataset	2006/07 2007/08 2008/09 2009/10 2010/11 2011/12	None	No investigation required	Ensure data and data dictionary variables are listed in the same order when data sets are sent to Statistics NZ in the future
Some variables have shifted order (fullTimeWeeks and TawellWeight for instance) as the data set progresses	2006/07 2007/08 2008/09 2009/10 2010/11 2010/12	None	No investigation required	Determine a suitable variable order, and apply consistently to all data sets sent to Statistics NZ in the future
Sheet name not updated for new year	2008/09 2011/12	None	No investigation required	Correct in data dictionary before revised version sent to Statistics NZ
Dates in cell A7 of "Taxwell Notes" not updated	2010/11 2011/12	None	No investigation required	Correct in data dictionary before revised version sent to Statistics NZ

Table 2: Year-on-year comparison of averages

Discrepancy	Years Affected	Impact	Notes	Action
Variable BenOtherBen too high	2007/08	Significant	BenMiscAmount too high. This is where missing FamAssist amounts (see below) have been assigned. This results in FamAssist amounts being double-counted in Disposable Income.	Get corrected 2007/08 data from Statistics NZ Provide corrected 2007/08 database to Statistics NZ Check and recalculate results for policy advice
Variables FamAssistCTCRawAmount FamAssistIRDRawAmount	2007/08	Significant	In HES, no income recorded under codes	See above

FamAssistMSDRawAmount FamAssistPTCRawAmount zero for everyone			3.2.0.01, 3.2.0.02, 3.2.0.03, 3.2.0.04, instead included in more general categories that feed into BenMiscAmount. Due to a coding error in the data supplied to Treasury by Statistics NZ.	
Variable BenAllNonTaxable too low	2008/09	Significant	Looks like actually BenMiscAmount too high in 2007/08 – see above	See above
Variable HHRentAnnual too low	2007/08	Significant	Looks like expenditure 04.1.01.1.0.02 Rent paid for primary property has not been included in this variable in this year. Used to calculated modelled Accommodation Supplement in Taxwell, not used directly by Statistics NZ	Provide corrected 2007/08 database to Statistics NZ Check and recalculate results for policy advice
Proportion of investment income attributed to PIE investments different every year	2006/07 2007/08 2008/09 2009/10 2010/11 2011/12	Moderate	Modelling suggests that this changes average disposable income by about 0.1% (well within modelling error) and measures of inequality are unaffected.	Document and use a consistent proportion for PIE for each year Correct if revised data is sent to Statistics NZ for the affected years Add proportion to corrected

				data dictionaries
Variables TotalIrregular, TotalNZInvestments, TotalNZOther, TotalOverseas zero for everyone	2008/09	Minor	Checked with Statistics NZ to see if they use these variables – they don't. Zero also for 2006/07 and 2007/08.	Correct if revised data is sent to Statistics NZ for affected years

Table 3: Compare Taxwell project files year-on-year

Change	Years Affected	Impact	Action
Parameter UseAccomSup changed from 'No AS' to 'Use Taxwell AS'	2008/09 2009/10	Significant	As described in Check 1
Parameters PIE/Percent/Divident and PIE/Percent/Investment changed from 0 to 1	2007/08 2008/09	Moderate	As described in Check 2
Parameters PIE/Percent/Divident and PIE/Percent/Investment changed from 1 to 0	2009/10 2010/11	Moderate	As described in Check 2
Parameters PIE/Percent/Divident and PIE/Percent/Investment changed from 0 to 0.5	2010/11 2011/12	Moderate	As described in Check 2
Variable alias WBStart incorrectly changed to WBEnd	2006/07	Minor	Correct before sending any revised datasets

Table 4: Compare Taxwell procedure files year-on-year

Change	Years Affected	Impact	Action
Variable F/Attributes/FamilyType definition incorrect	2007 2009	Minor	Correct before sending any revised datasets
Core benefits incorrectly treated as integers, corrected to doubles	2009 2010	Minor	Correct before sending any revised datasets
Working for Families income definition corrected	2010 2011	Minor	Correct before sending any revised datasets
Tax now specified as non-negative – previously could take any value including negative values	2010 2011	Minor	Correct before sending any revised datasets