

# Income measures in the IDI

Income is an important driver of health inequalities in New Zealand (for example see the [New Zealand Census-Mortality Study](#)) and it is important to take into account when studying health outcomes in the Integrated Data Infrastructure (IDI). There are many ways of examining income. We discuss the two main sources of income information available nationally. The first is summary income information from the Inland Revenue Department (IRD) and the second is income reported in the 2013 census.

## Inland Revenue income

The strength of this IRD personal income data is that it provides an objective and longitudinal measure of a person's income over time, and it can be cut in many different ways, whilst also accounting for time outside the country, e.g., income in past year, average income over past five years, changes in income etc. Only government researchers can apply for access to business data which includes self-employed income (due to the IR Tax Act), the full IDI Inland Revenue dataset and the Longitudinal Business Database. However in October 2017, StatsNZ released clarification about who has this access. Inland Revenue's definition of government researcher was broadened to incorporate not just NZ government agency researchers but also researchers from a NZ government funded research institution (including NZ universities). This is good news for NZ university researchers wanting to apply for access to more detailed IRD data (including income in self-employed people).

Government and NZ university researchers can apply to access the `income_tax_yr_summary` table in the 'data' schema, which is an aggregated `income_tax_yr` table. This table draws on data from the Employer Monthly Schedule (EMS), IR3 (currently this excludes the `ird_rtns_keypoints_ir3` data on overseas interest and income from trusts), IR4 and IR20 tables. See the Clarke (2014) data dictionary below for more information on how variables are derived. Income data are available for the following categories;

- 'W&S' Wages and salaries
- 'BEN' Benefit payments from the Ministry of Social Development (MSD)
- 'CLM' Accident Compensation Corporation (ACC) payments
- 'PEN' Pension payments from MSD
- 'PPL' Paid parental leave payments from MSD
- 'STU' Student allowance payments from MSD
- 'C00' Total shareholder Salary amount (ir4s variable `ir_ir4_tot_sholder_sal_809_amt`)
- 'C01' Company director/shareholder receiving PAYE deducted income
- 'C02' Company director/shareholder receiving WHT deducted income
- 'P00' Sole trader income (IR20 variable `ir_ir20_tot_share_of_inc_865_amt`)
- 'P01' Partner receiving PAYE deducted income
- 'P02' Partner receiving withholding tax deducted income
- 'S00' Sole trader income (IR3 variable `ir_ir3_net_profit_amt`)

- 'S01' Sole Trader receiving PAYE deducted income
- 'S02' Sole Trader receiving withholding tax deducted income
- 'S03' Net rent income (IR3 variable ir\_ir3\_net\_rents\_826\_amt)

Individual income data is available from 1 April 1999 onwards up to 2015 (currently). Longitudinal income data such as these may be used to measure trends in household income over time particularly as information on household composition improves, including a monthly record of income. Data on some individuals may be repressed including sensitive occupations such as judges. The table does not contain interest, dividends tax credits, overseas interest, trust incomes, income earned and taxed overseas and income not taxed at source like the accommodation supplement. There is information on benefits paid for MSD's main benefits. These are all coded under the source code 'BEN' and currently include unemployment benefits, domestic purposes benefits, sickness benefits, invalid's benefits, widow's benefits and emergency benefits.

There are other Inland Revenue tables with additional income data also available. These include new variables in the ird\_rtns\_keypoints\_ir3 table for sources of income such as tax credit entitlements, overseas and estate trust income, covering the years 1999 to 2015. A new IR Personal Tax Summary table 'ird\_pts' has been made available (since the June 2017 refresh) that contains information on sources of income such as interest, dividends and tax credits and covers the years 1999 to 2015.

For 'non-government' researchers, a more restricted aggregated set of personal income data can be used. The Employer Monthly Schedule (EMS) table 'ems\_cal\_yr\_summary' was produced in 2015 and is located in the IDI sandpit for the years 1999 to 30 Sep 2015 (currently). The main caveats are that this table excludes information on self-employed income (including salaried self-employed income), investment income, income not taxed at source (eg the accommodation supplement, a non-taxable government transfer) and income earned and taxed overseas. In the Healthier Lives Earthquake Cardiovascular Disease project, we found that in 2010 income records were available from the IRD for approximately 90% of the population 25+ years old.

### **Census income**

The second main source of income information in the IDI is from the 2013 census. There are many ways of creating summary income variables depending on the research question. The Jensen Equivalised Income is one we have found useful for our particular research to account for the composition of the household (Ministry of Health, 2005). Jensen Equivalised income tends to separate lower incomes and takes account separately the number of adults and children in a household, whereas the Luxembourg Equivalised Income method creates a different income distribution and only uses the total number of people in the household. Jensen equivalised income is used routinely by Statistics New Zealand. This method is also used to create a variable used in the New Zealand Deprivation Index, and in the analyses in the New Zealand Census-Mortality and CancerTrends studies for equivalised household income quintiles (Ministry of Health, 2005). An example of VHIN code is available – see [code sharing](#). Finally, census income can also be analysed at a personal level. An example of code that extracts personal income is available from the Virtual Health Information Network catalyst projects ([code here](#)).

For more information on the IRD information in the IDI please see;

- [Suei, S. Comparing income information from the census and administrative sources.](#)
- [Clarke, A. \(2014\) Data dictionary for IR derived data in the IDI.](#)

- Statistics New Zealand (2015). *IDI Data Dictionary: IR tax data (September 2015 edition)*. Available from [http://www.stats.govt.nz/browse\\_for\\_stats/snapshots-of-nz/integrated-data-infrastructure/idi-data/irtax-data.aspx](http://www.stats.govt.nz/browse_for_stats/snapshots-of-nz/integrated-data-infrastructure/idi-data/irtax-data.aspx)

## References

MINISTRY OF HEALTH 2005. Decades of Disparity II Socioeconomic mortality trends in New Zealand, 1981–1999. *Public Health Intelligence Occasional Bulletin Number 25*. Wellington: Ministry of Health. Accessible; <http://www.otago.ac.nz/wellington/otago024500.pdf>

*By Andrea Teng, June Atkinson*

Version: *Original 3 April 2017 and updated 7 November 2017.*



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).