

Data

We use information from the person-level integrated data asset (PLIDA) provided by the ABS DataLab environment. Using this data asset, we can link de-identified individual tax records with university enrolments data (ABS DataLab, 2023).

- To obtain employment income, we add up gross income from all jobs an individual worked in a given year. This data stems from payment summaries that are provided by employers for income tax purposes. In theory, this data will encompass all individuals in formal employment for the financial year. We use the wage price index to convert all figures into 2022 dollars.
- We supplement the income tax data with 2016 Census data on people's age and country of birth. The primary motivation behind this is so we can filter the sample for people who hypothetically completed high school in Australia between 2004-2018, and would have enrolled in university before 2019 (we have university enrolments data between 2005-2019).
- To obtain ATAR data, we use higher education enrolments data collected as part of the administration of Commonwealth supported education loans. If an individual enrolled in a university course using their ATAR at some point, we record their ATAR in the data.
- The taxation data, the 2016 Census data, and the higher education data are joined at the individual level via a hashed individual identifier.

Disclaimer

The results of these studies are based, in part, on data supplied to the ABS under the Taxation Administration Act 1953, A New Tax System (Australian Business Number) Act 1999, Australian Border Force Act 2015, Social Security (Administration) Act 1999, A New Tax System (Family Assistance) (Administration) Act 1999, Paid Parental Leave Act 2010 and/or the Student Assistance Act 1973. Such data may only be used for the purpose of administering the Census and Statistics Act 1905 or performance of functions of the ABS as set out in section 6 of the Australian Bureau of Statistics Act 1975. No individual information collected under the Census and Statistics Act 1905 is provided back to custodians for administrative or regulatory purposes. Any discussion of data limitations or weaknesses is in the context of using the data for statistical purposes and is not related to the ability of the data to support the Australian Taxation Office, Australian Business Register, Department of Social Services and/or Department of Home Affairs' core operational requirements.

Legislative requirements to ensure the privacy and secrecy of these data have been followed. For access to MADIP and/or BLADE data under Section 16A of the ABS Act 1975 or enabled by Section 15 of the Census and Statistics (Information Release and Access) Determination 2018, source data are de-identified and so data about specific individuals has not been viewed in conducting this analysis. In accordance with the Census and Statistics Act 1905, results have been treated where necessary to ensure that they are not likely to enable the identification of a particular person or organisation.