

## 4.1 Data treatments: top and bottom coding, treatment of negative incomes

### Context

- Once recorded on the field, income values are submitted to a series of treatments during the survey processing stage
- Most of these modifications aim at minimizing measurement errors that occurred during the field work: treatments for negative income and bottom coding are one of those
- Treatments can also be applied to ensure a sufficient level of privacy and confidentiality to respondents: top coding is one of these
- But alterations of the tails of the income distribution' will affect inequality and poverty figures, e.g. :

 Burkhauser-Feng-Jenkins (2009), "Using the P90/P10 index to measure U.S. inequality trends with current population survey data: a view from inside the Census Bureau vaults", *Review of Income and Wealth*

## In practice: top and bottom coding

- Some countries apply top coding at different stages of their process (Ireland, Japan, Netherlands, United States), other countries provide some information above the top coding threshold (i.e. mean cell values mainly)
- Changes in country practices on top coding (e.g. In PUF) can lead to spurious changes in inequality (e.g. US experience)
- OECD Terms of Reference asks countries to provide estimates based on non-top coded values (e.g. US) but this is not always possible (e.g. EU)
- Very little information on country practices on bottom coding



## In practice: treatment of negative values

- Most OECD countries reprocess negative incomes for detailed income components, but practices differ:
  - Retained
  - Set to zero
  - Retained but excluded from computations
- For the OECD income distribution database:
  - All individual components of market income showing negative values are set to zero after equivalization
  - Alternative treatments of these negative values can lead to different results for income inequality and poverty measures (e.g. Australia)



## For discussion

- Get better understanding of country practices on top/bottom coding and treatment of negative values
- How do these practices impact upon the cross-country comparability of key measures
- Is there a need for different recommendations on top/bottom coding and negative values in OECD Terms of References?

