EUROMOD
Modelling Conventions

adjusted for UNU-WIDER-ISER-SASPRI (and local partners) Project

Sanja Vujackov

1st UNU-WIDER / ISER / SASPRI Webinar
EUROMOD, University of Essex, United Kingdom, December 2015

EUROMOD

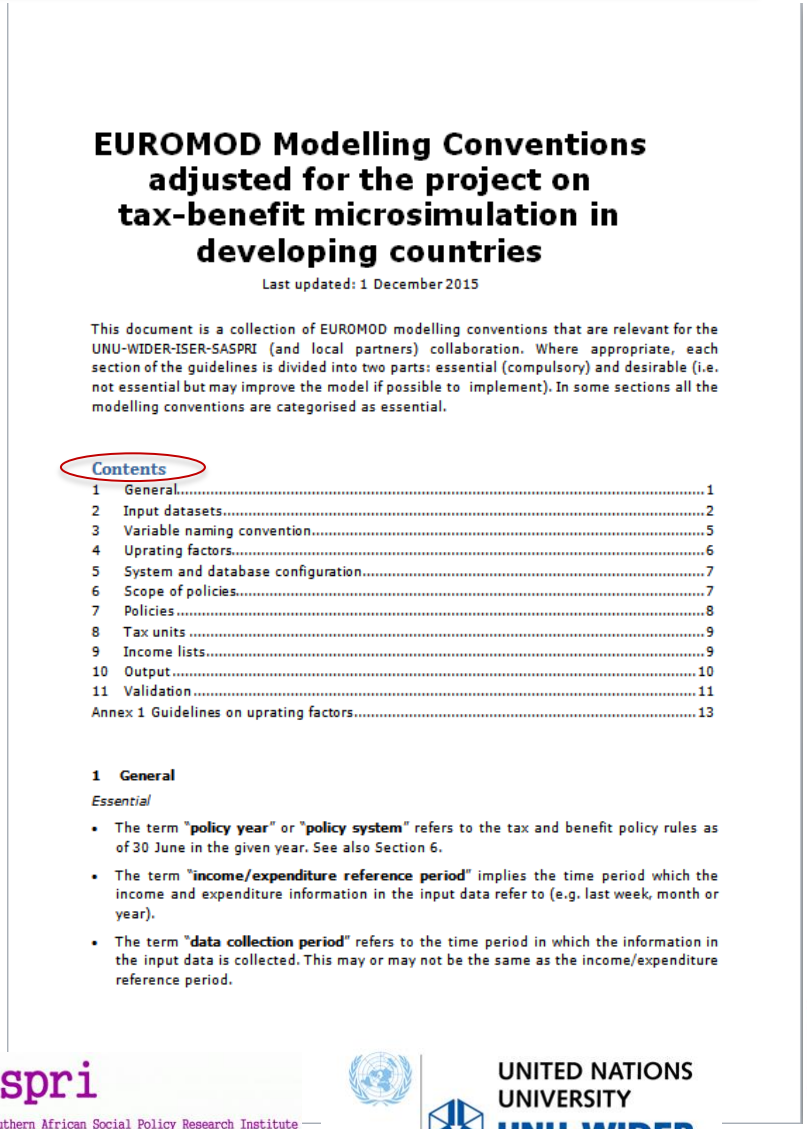
saspri
Southern African Social Policy Research Institute
Southern African Social Policy Research Insights

UNITED NATIONS UNIVERSITY
UNU-WIDER

ISER
INSTITUTE FOR SOCIAL & ECONOMIC RESEARCH

Introduction

- Importance of *consistency* and *harmonization*
- This document is a collection of best practises after years of experience with EUROMOD, SAMOD and NAMOD
- Each section is divided into:
  - *Essential* (compulsory)
  - *Desirable* (to improve the model if it is possible to be implemented)
- It contains information about input and output data, income lists and variable`s names, policies, uprating factors, validation, etc.



EUROMOD Modelling Conventions adjusted for the project on tax-benefit microsimulation in developing countries

Last updated: 1 December 2015

This document is a collection of EUROMOD modelling conventions that are relevant for the UNU-WIDER-ISER-SASPRI (and local partners) collaboration. Where appropriate, each section of the guidelines is divided into two parts: essential (compulsory) and desirable (i.e. not essential but may improve the model if possible to implement). In some sections all the modelling conventions are categorised as essential.

Contents

1 General

Essential

- The term "**policy year**" or "**policy system**" refers to the tax and benefit policy rules as of 30 June in the given year. See also Section 6.
- The term "**income/expenditure reference period**" implies the time period which the income and expenditure information in the input data refer to (e.g. last week, month or year).
- The term "**data collection period**" refers to the time period in which the information in the input data is collected. This may or may not be the same as the income/expenditure reference period.

EUROMOD

saspri
Southern African Social Policy Research Institute
Southern African Social Policy Research Insights

UNITED NATIONS UNIVERSITY
UNU-WIDER

ISER
INSTITUTE FOR SOCIAL & ECONOMIC RESEARCH

General

- Some basic definitions:
  - "*policy year*" ("*policy system*")
  - "*income/expenditure reference period*"
  - "*data collection period*"
  - "*base-year simulation*"
  - "*target-year simulation*"
  - "*baseline simulations*"

- An agreed two-letter country acronyms

Input datasets

1/2

Goal: to organize/compile the input micro data in a consistent and meaningful manner and to make it usable for the purpose of tax-benefit micro-simulation

- **Input dataset** must be:
  - in (tabulated) *text format*
  - named as CC_year_a# (country acronym_data collection year_version number) – e.g. SA_2012_a1.txt
  - must have *compulsory variables* (identification, demographic and labour status variables)
  - all variables must be documented in a Data Requirement Document (DRD)
  - income data (and desirable expenditure) should be *as detailed as possible*

Input datasets 2/2

- **Data** must be at *individual level* (with people grouped into households) in way that:
  - all *monetary variables* need to be at individual level (if at household level, then to be assigned to one person in household)
  - *non-monetary variables* should be assigned to all persons in the household
- income and expenditure data expressed in *monthly terms*
- presented in *national currency*
- all income variables must be provided with *gross values* and expenditure variables *inclusive of VAT*
- *Missing values* are not allowed

Variable naming convention

Goal: to have comparable across countries and meaningful variable names

- All variable names *must begin with* a **Class 1** acronym *followed by* at least one **Class 2** acronym:
  - *Class 1: one-character that identifies type of variable (asset, benefit, demographic, etc.)*
  - *Class 2: two-character specific for each variable type*
- For example: income from employment is named:
  "*yem*" – "*y*" for *market income* + "*em*" for *employment*
- The *id-variables* are exception as they have two-letter Class 1 acronym (e.g. idmother)
- Check whether the *variable name* is *already defined* before creating a new one

Uprating factors

Goal: to project income and expenditure levels in the input micro data up to the policy year

- Where income reference period and policy year do not match, *monetary variables* need to be *uprated* to the policy year
- Using relevant *uprating factor(s)*
- If not possible, then by using *CPI*
- *Factor = 1* should be used when no change was observed
- **Annex 1.** with further information about *possible sources for uprating* different income components - such as market incomes, pensions, non-simulated benefits, etc.

Scope of policies

Goal: to simulate a set of (comparable) policies in a consistent manner across countries

- Policies are simulated for the corresponding policy year
- *Most important* ones – social insurance contributions, income, property, wealth and other personal direct taxes, cash benefits
- Benefit *non-take-up* and *tax evasion*
- *Regional differences* to be simulated as far as possible

Output

Goal: to make clear what information is included in the output micro data

- Results, by default, must be outputted at the **individual level**
- **Output data** must be in *national currency*
- **Needs to include** all: income variables, simulated income variables, socio-demographic variables, income lists and identifiers for all tax units
- We should not have any *temporary* or *intermediate variable* or any *variable* included twice

Validation

1/2

- **Micro validation:**
  - Check *eligibilities and the amounts* of taxes and benefits simulated by the model (case-by-case validation for a selection of particular households)
  - Compare *simulated values* against data recorded values in the same survey on case-by-case basis
  - Check *descriptive statistics* for outcome variables (min, max, mean)
  - Check that *results* for some basic indicators (e.g., average tax rate) make sense for all observations in the sample

Validation

- **Macro validation:**
  - Compare the *sum of each income component and the number of recipients* (both original incomes and tax-benefit instruments) with external statistics;
  - Focus should be on *relative differences*, e.g. how one instrument compares to another, whether the bias has an expected sign and whether trends over years are in the expected direction.
  - Calculate and compare *inequality measures* (Gini, S80/S20 ratio) and *poverty measures* with external statistics

Thank you!

http://www.euromod.ac.uk

euromod@essex.ac.uk

EUROMOD

saspri
Southern African Social Policy Research Institute
Southern African Social Policy Research Insights

UNITED NATIONS UNIVERSITY
UNU-WIDER

ISER
INSTITUTE FOR SOCIAL & ECONOMIC RESEARCH