Data validation
*Item 9 of the agenda*

Eurostat
ESTAT-F5
Introduction

For information:
  o Review of the validation policy at ESS level

For discussion:
  o ESSPROS validation rules
Background

- **May 2012:** ESSC adopted a number of general principles for a review of the validation policy at ESS level
  - ESSC agreed that implementation of these should be decentralised
  - Each WG has to discuss and adopt a list of validation rules that ensure an agreed "minimum data quality standard"

- **An agreed list of validation rules will facilitate:**
  - Implementation of validation rules in Eurostat/national IT systems
  - Development of shared validation services
  - Automation of the statistical production process
  - Reduction of validation "ping-pong"

- **An agreed list of validation rules will allow a more objective assessment of the quality of the data**
Validation rules

• **First draft of the list of validation rules:**
  - SP-2014-09-Annex 1: rules for the Core system
  - SP-2014-09-Annex 2: rules for the Pension beneficiary module
  - SP-2014-09-Annex 3: rules for the Net Benefit module
  - SP-2014-09-Annex 4: cross-module validation rules

• **The description of each rule should include:**
  - An identification number
  - A description of the rule
  - A sensitivity threshold (optional)
  - An error severity
  - An error message
Validation rules

- **Common language to express validation rules is under development.**
  - Once complete, the rules will be “translated” into this language
  - Should facilitate development of shared validation services

- **Validation rules are classified by “level”:**
  - Level 0: Checks structure/format + dictionary
  - Level 1: Checks the data included in the file
  - Level 2: Checks between 2+ files of the same type from one source
  - Level 3: Checks between 2+ files of the same type from different sources
  - Level 4: Plausibility/consistency checks between ESS statistical domains
  - Level 5: Plausibility checks with any possible correlated data.

Levels to facilitate positioning of checks in the process
For discussion:

Countries are invited to comment on the draft list with the aim of improving it (format and content).

(Box 1)

NO EEV on this item

➡ First preliminary comments today (general and specific)

Improve proposed checks
Other checks?
Presentation to improve? (levels; five columns; ...)

9-10 April 2014

Working Group Social protection
Next steps

• Based on the comments of the WG, an amended list will be submitted to the WG for approval by written procedure.

• Once the list (version 1) adopted, the checks can be implemented within national and Eurostat IT systems (example: in the SDMX framework).

• The list will continue to evolve in the future on the basis of experience (version 2, version 3, etc.).