

# Harmonization of Labor Statistics

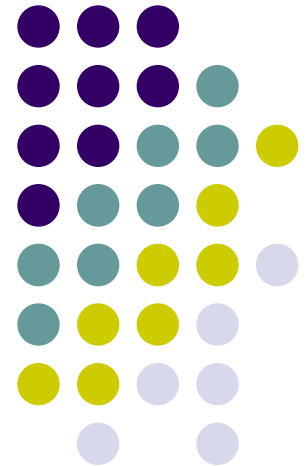
**Nadia Belhaj Hassine**

Economic Research Forum

**Expert Group Meeting**

UNESCWA

June 29-30, 2011



# Presentation Outlines



1. **Overview**
2. **Harmonization process**
3. **Some examples**
4. **Mains challenges**
5. **Recommandations**

# Overview



- Comparable labor statistics across countries, over time and across subpopulations (by gender, age groups, regions) are key for active and effective labor policies.
- Comparability of data on:
  - women, youth, rural participation
  - informal employment, underemployment
  - Precision of estimates of differences poses real challenges

# Overview



- Difference in design, coverage, detail of information collected, methods used to adjust for non-response are serious concerns for labor statistics comparability
- Greater awareness of where non-comparability can creep in...among other things:
  - National specificities, institutional issues, definitions adopted etc..
  - Definition of categories
  - Income from employment

# Overview



- Greater understanding of the sensitivity of key labor indicators to non-comparability
- Objectives, deliverables, and scope
- Requirements

# Harmonization process



- Gathering micro data and documentation
  - codebook, questionnaires etc. should be read carefully to understand the contents of the surveys
  - Basic checks
- Data Cleaning
  - Variable ranges
  - Duplicate records
  - Not applicable cases
  - Inconsistent cases

# Harmonization process



- File restructuring:
  - Availability matrix
  - Select the variables to include in the harmonized file (best practice and international standards..)
  - Define the categories to be used for each variable
  - Carefully document the definition of the variables and changes made
- Data mapping
- Recode or standardize values
- Create output aggregates



# Some examples

- **Roster**
  - **Administration**
  - Location
  - Demographic information
  - Individual records

Unique country identifier
Survey year
Unique person identifier
Weight
Region
Area
Urban/rural



# Some examples



- Roster
  - Administration
  - Location
  - Demographic information
  - Individual person records
    - Demographics
    - Education
    - Health
    - **Labor**
    - Income

Labor force status in the current period (Cat.)

Main activity status in the current period (Cat.)

Status of employment (Cat.)

Classification of occupations for main job (Cat.)

International Standard Industrial Classification of main job (Cat.)

Sector of employment (Cat.)

Tenure in current job (Cont.)

Total weeks worked (Cont.)

# Some examples



- **Labor Force statistics**

*(ILO, OECD ..Classifications)*

- Labor force by age, gender etc..
- Total employment (periodicity)
- Employment by gender
- Employment by industry
- Unemployment (by subgroups, duration??)

# Some examples



- Earnings
  - Hourly earnings
  - Annual wages
  - Average wage (full time employment)
  - ....

# Mains challenges



- Poor quality of data in some surveys
- Lack of adequate documentation for some surveys (questionnaires, coding sheets, methodology, labelled data files), lack of definitions for some country specific expressions, and identification of time periodicity.
- Mismatch between different coding sources for the same data set (questionnaire, coding sheets and the labels provided in the data file).

# Mains challenges



- Inconsistent structure within the same country data over a time period, some variables are added and others are eliminated, change of categories etc.
- Some essential variables are sometimes missing; some main responses in certain variables are aggregated with others.

# Mains challenges



- Differences in survey types (cross-section vs panel, ..)
- Differences in coverage, focus and contents, level of detail
- Different statistical techniques (missing values vs imputed amounts)

# Recommendations



- Use best practices, guidelines and international standard classifications (OECD, ILO-definitions, etc.)
- Extensive documentation, network of contacts in NSA (remaining inconsistencies need to be documented)
- Avoid confusion between concepts within one variable (*usually employed vs employed last week*)
- Confine comparisons to data that have been redefined to be comparable
- Sensitivity tests