PIAAC Data Analysis in Stata:
A practical guide

Video 1: Introduction to PIAAC Data
GESIS – Leibniz Institute for the Social Sciences
In a Nutshell

**Goal:**
- Practical guide for three existing tools to perform PIAAC data analysis in Stata

**Target Group:**
- Researchers with some experience in Stata, but little or no experience on analysis of PIAAC data in Stata

**Structure:**
- General overview of how the three tools work, as well as several exemplary analyses with the PIAAC data
Video 1: 
Introduction to PIAAC Data
Short Introduction to PIAAC

- International large-scale assessment of adult skills
- Initiated by the Organization for Economic Cooperation and Development (OECD)
- Three rounds of data collection between 2011 and 2017 in 38 (OECD and partner) countries
- In each country, nationally representative sample of adults between 16 and 65 years (N ≈ 5,000)
- Interview structure:
  1. Background questionnaire (personal interview)
  2. Skill assessment (self-administered)
Interview Structure

Background Questionnaire → Computer Experience?

- no
- yes

Paper-Based Assessment

Computer-Based Assessment
Skill Assessment in PIAAC

- Skill assessment in three key domains
  - Literacy
  - Numeracy
  - Problem-solving in technology-rich environments (PS-TRE)

- Results reported
  - On proficiency scales from 0-500 points or
  - As proficiency levels

- Further information:

https://www.oecd.org/skills/piaac/
PIAAC Research Topics I

- Skill differences across sociodemographic groups, e.g., gender, migration background, ...
- Skill differences across countries, educational systems, ...
- Returns to skills and skill mismatch in the labor market
- Skill use in the workplace and during leisure time
- Skills and social outcomes, e.g., risk, trust, ...
- Methodological aspects, e.g., interviewer effects, non-response bias analysis, ...
- And many more...
PIAAC Research Topics II

Reference:
Important facts for analyzing PIAAC data

Plausible Values (PV):

▪ Each respondent only worked on a subset of items
▪ Unbiased proficiency estimates: 10 PVs are drawn from distribution of latent proficiency, conditional on observed response scores and background information
▪ PIAAC provides 10 PVs for each of the three skill domains

Complex Sampling:

▪ Sampling variance needs to be taken into account
▪ PIAAC applies replication method which produces up to 80 replicate weights to be considered in addition to a full sample weight

→ All analyses in PIAAC that include proficiency have to use the ten PVs values
→ Variance estimation in PIAAC needs to reflect uncertainties due to measurement of skills and due to sampling
Further Reading on PIAAC
