

# Monitoring Indications of Quality Variation in Sampling Activities in PIAAC

Tom Krenzke, Westat

March 28, 2014

CSDI Conference, Bethesda, Maryland

# PIAAC Facts

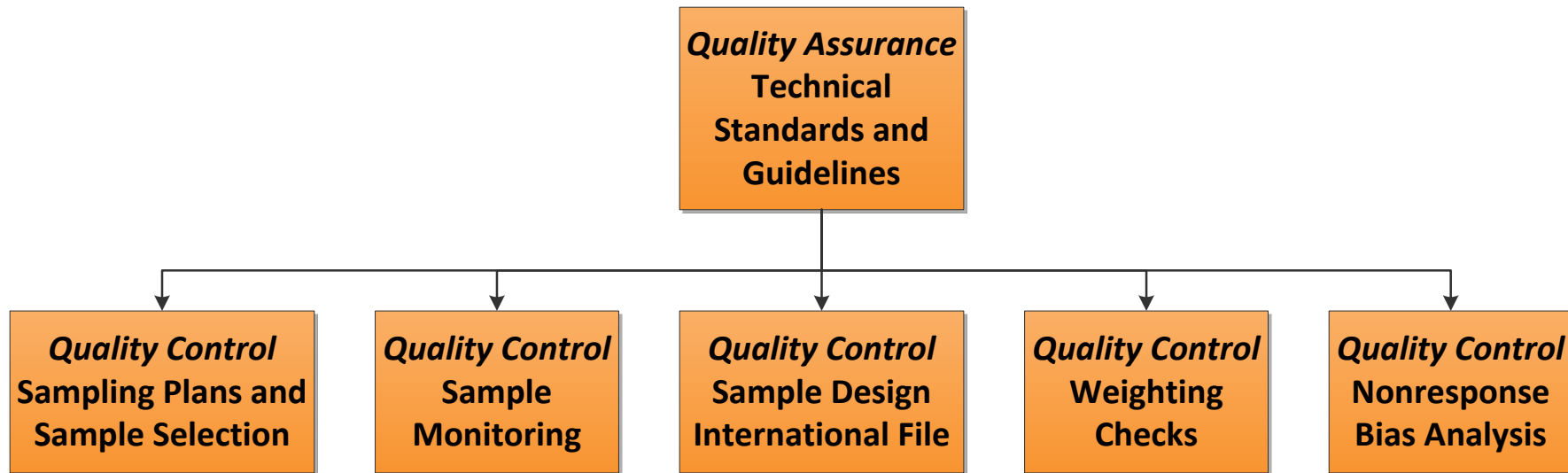
- Study objective
  - From Organisation of Economic and Co-operation Development (OECD) website:
    - OECD collects and analyses data that assist governments in assessing, monitoring and analyzing the level and distribution of skills among their adult populations as well as the utilization of skills in different contexts
- Target population
  - 16-65 years old
  - Noninstitutionalized population
- 24 Round 1 countries (collected data in 2011-2012)
- 9 Round 2 countries (data to be collected in 2014)
- 5,000 Background questionnaires and assessments

## PIAAC Facts (2)

- Obvious indications of quality variation
  - Response rates
    - Round 1 ranged from 45% to 75%
  - Design effects (DEFFs)
    - Increase to variance due to complex design
      - Round 1 ranged from 0.80 to 15.77
        - » Highest 3 → 15.77, 6.62, 3.53
  - Appears to be wide variation in quality among countries
    - However...
      - Implementing standards helped to reduce bias
      - Many high DEFFs are due to oversampling by design
  - We explain how quality variation among countries was reduced in terms of bias and variance

# Quality Control Objectives

- To ensure that the Technical Standards and Guidelines for sampling are followed so that survey results are comparable in quality across countries

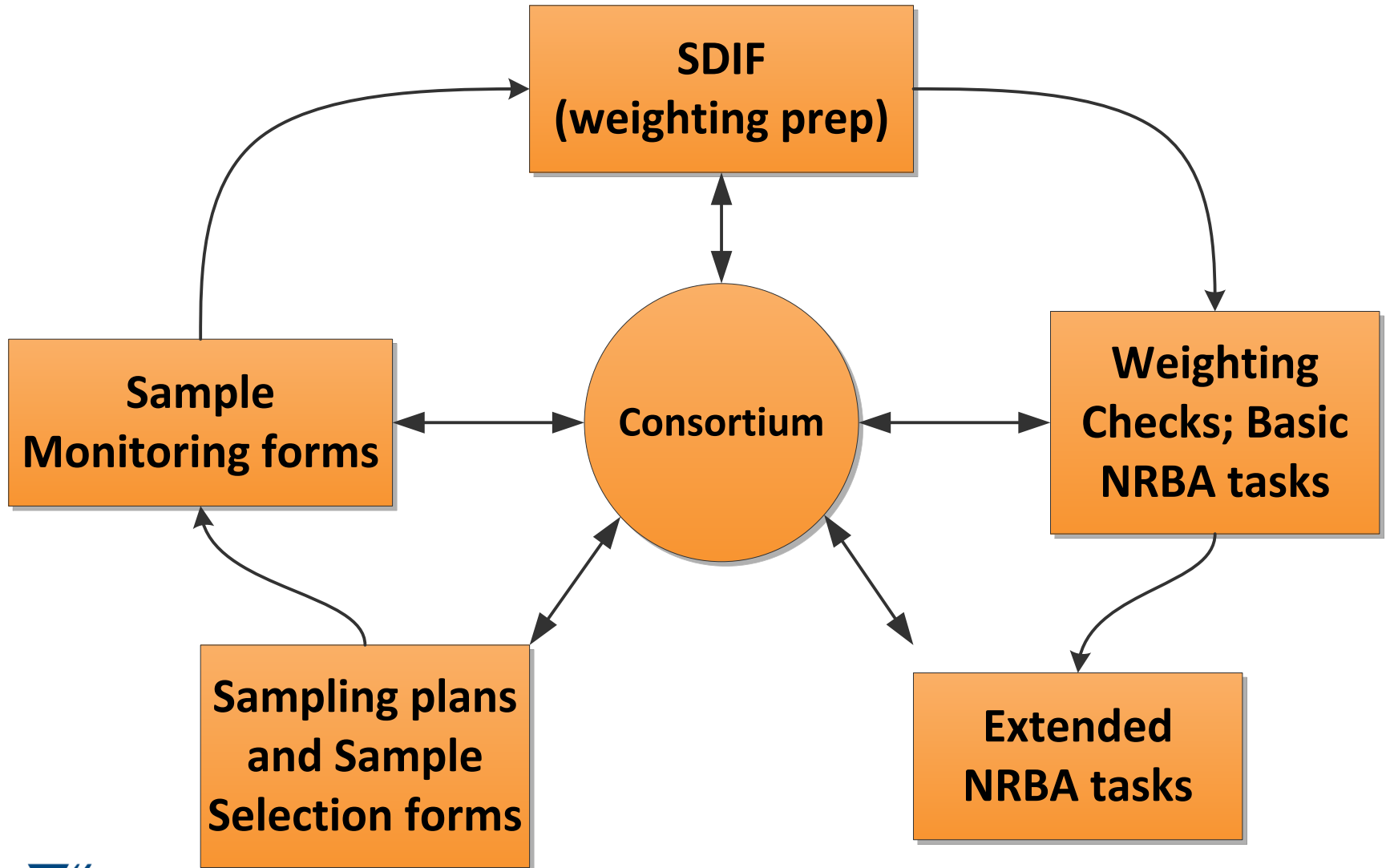


- To retain quality at a high level
- To reduce the amount of quality variation among countries

# Sources of Quality Variation in Sampling Activities

- Sampling plans and sample selection
  - Sample bias and variance
- Sample monitoring
  - Response rates → Nonresponse bias
  - Sample yields → Variance
- Sample Design International File (SDIF)
  - Quality of probabilities of selection, sampling and weighting variables, disposition codes
- Weighting
  - Weighting variables → Nonresponse bias
  - Design effects → Variance
- Nonresponse bias analysis (NRBA)

# Detecting Quality Variation in Sampling Activities



# QC Sampling Plans and Sample Selection

- Sampling plans template (Excel)
- Sample selection forms (Excel)
- Examples of initial detection of quality variation
  - Bias
    - Disposition codes initially not aligned with standard codes
    - Initially only listed eligible units
    - Reserve sample planned to come from 1<sup>st</sup> half of frame
    - Initial exclusion of non-native speakers
  - Variance
    - Initially too few Primary Sampling Units
    - Unequal probabilities of selection
    - Outdated frame information
    - Improved design by using implicit stratification rather than simple random sample

# QC Sample Monitoring

- Objectives of forms
  - To identify potential shortfalls in the sample
  - To identify problems in achieving the desired response rate
  - To identify potential for nonresponse bias
  - To employ procedures while it is still possible to meet goals associated with sampling and data quality
- Types of forms
  - Basic (monthly)
    - Sample yields and response rates
    - Actual, projected and goals
  - Extensive (once or twice during data collection)
    - Classification tree analysis
    - Identify subgroups with a low response rate



# QC Sample Monitoring (2)

- Examples of initial detection of quality variation
  - Bias
    - Misclassification of some cases to disposition codes
    - Field test showed large number of literacy-related nonrespondents. Did not translate to all official languages
    - Distribution across age unexpected – sampling issue
    - Low RRs in some subgroups led to special mid-data collection efforts
  - Variance
    - Low overall sample yield
  - Other
    - Field test
      - Showed no screener dispositions available
      - Could not compute RRs for subgroups, code provided

# QC SDIF

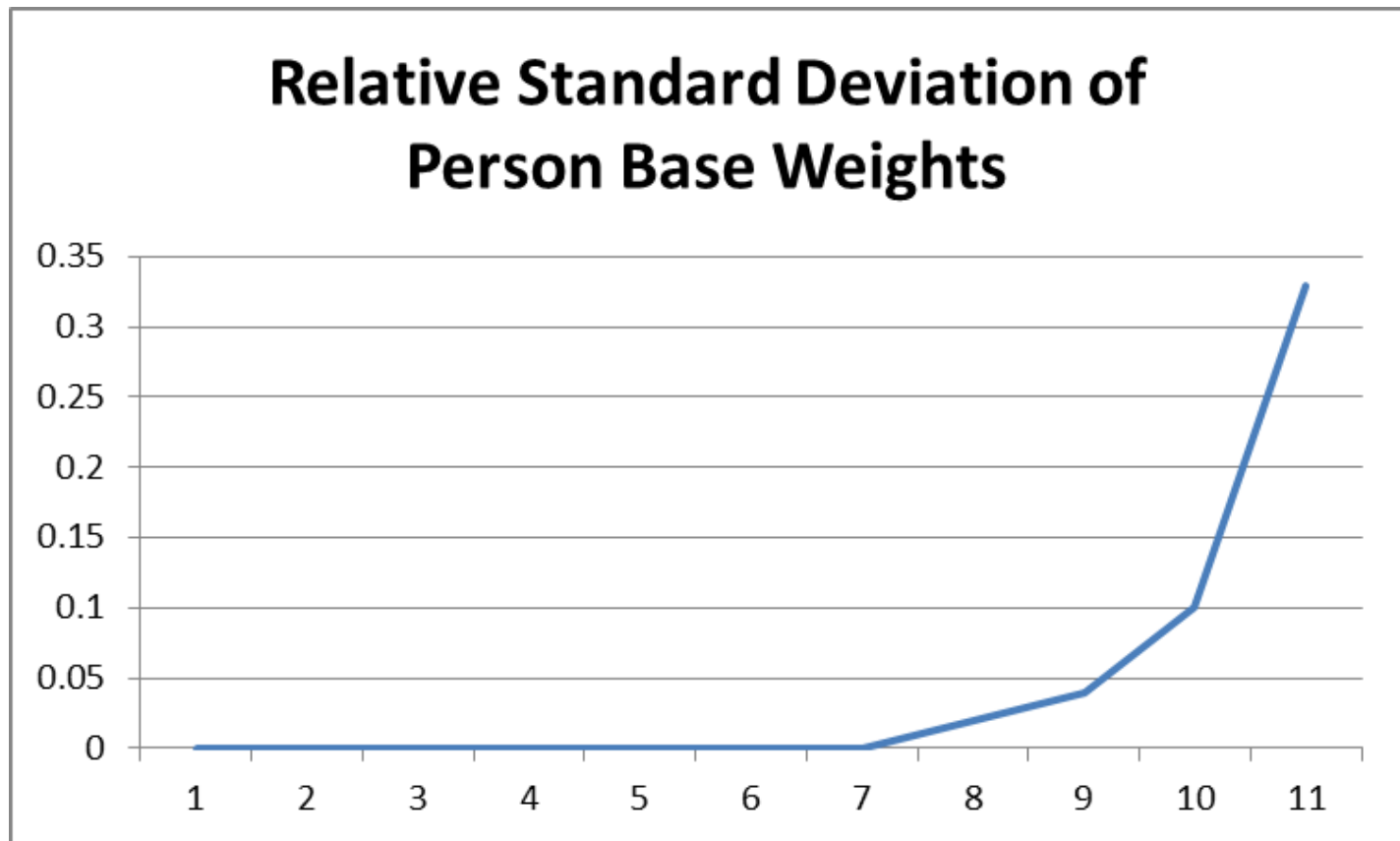
- Objective: To prepare a base file for weighting
- Examples of initial detection of quality variation
  - PSU, SSU, and DU counts different than reported
  - Probabilities not between 0 and 1
  - Sum of base weights not aligned with population estimates
  - PSUs were not aligned with strata
  - Sorting variables were not one-to-one with sampling unit
  - Group quarters flag was not used
  - Calibration variables not same for sample and controls
  - Assessment disposition codes not assigned correctly
  - Variance strata and variance units not assigned according to PSU strata and units

# QC Weighting

- Objective: Double check plans and results through Excel spreadsheets and data files
- Select checks of initial detection of quality variation
  - Descriptive statistics on full sample and replicate weights
  - Identify extreme weights
  - Compute the weighted frequencies of key survey variables
- Examples of initial detection of quality variation
  - Bias
    - Missing or zero weights
    - Frame totals not representing target population
    - Weak weighting variables

## QC Weighting (2)

- Variance. Relative standard deviation of base weights for registry countries without oversampling in Round 1



# QC NRBA

- Objectives of NRBA
  - Basic NRBA (all countries)
    - To evaluate the relationship of response status to available auxiliary variables
    - To provide an indication of nonresponse bias prior to weighting adjustments
    - To use results to select weighting variables to reduce potential nonresponse bias

## QC NRBA (2)

- Objectives of NRBA (continued)
  - Extended NRBA (countries < 70% RR; or < 80% in a stage)
    - To evaluate the effect of weighting
    - To study nonresponse and undercoverage bias
    - To investigate the relationship between auxiliary variables and PIAAC main study proficiency estimates
    - To estimate the potential bias if nonrespondents are very different from respondents within a weighting cell

## QC NRBA (3)

- Detection of quality variation
  - A country reproduced their weights after conducting the Basic NRBA by using stronger weighting variables
  - For countries with a low RR, the potential for bias was reduced to a comparable level due to their strong weighting variables
  - Easier cases scored lower on average, however, countries worked their cases thoroughly according to standards and reduced bias by getting the tough cases

# Recommended Improvements

- Sampling
  - Evaluate the quality of sampling frames
  - QC sample selection forms sign off before proceeding
  - Automated checks in QC sample monitoring forms
  - Require use of Response Rate ToolKit
  - Improve upon stratification and sorting scheme from field test
- SDIF: Program to conduct extensive edit checks prior to submitting
- Weighting
  - Consortium to conduct weighting
  - Review the set of weighting variables used by other countries
- NRBA: Consortium to conduct Extended NRBA



# Summary

- Detection of quality variation in terms of bias and variance
  - Through forms and files
    - Sample selection, sample monitoring, SDIF, weighting, NRBA
  - Face-to-face discussions or phone discussions help as well.
- Examples...
  - Substitution?
  - Stop at 5000?
  - Vacants left out of the address listing
  - Sometimes easier to communicate
- In general, we concluded that there was a fairly comparable level of quality across countries, citing attempts and results in reducing quality variation