

FAM SAMPLING WORKSHOP
WESTMINSTER CATERING AND CONFERENCE CENTER
WESTMINSTER, MARYLAND

TUESDAY, MAY 25, 1999

08:45 - 09:15

Bias Issues on Data Collection

- Importance of pre-testing surveys
- Minimizing non-response

BIAS ISSUES IN DATA COLLECTION

Accuracy, validity (non-sampling error)

- # *Mean Square Error* describes total survey error
 - Mathematically, variance of the survey estimate plus the square of the bias
 - MSE =
- # Survey and data processing biases
- # Sample bias (a form of non-sampling error)

Survey and Data Processing Biases

- # Conceptual errors
- # Invalid questions
- # Respondent errors
 - misunderstanding
 - lying
 - recall errors (in retrospective surveys)
- # Interviewer errors
 - not asking questions correctly
 - mis-recording answers
 - making up answers
- # Non-response

Survey and Data Processing Biases *continued*

- # Coding errors
- # Data entry mistakes
- # Computer programming errors
- # Mis-application of estimation formulas
- # Analytical mis-interpretations

Sampling Biases

Not to be confused with sampling variance , or standard errors

Failure to use probability methods at *every* stage of selection

- often occurs at the next to last stage of selection when shortcut techniques are used to avoid listing households

- example is the cluster survey method used in the EPI program of WHO

Coverage error

- occurs most often when sampling frame is old

- squatter areas on the fringes of cities may be candidates for under-representation

- areas of heavy new construction since the frame was developed are troublesome

Sampling Biases *continued*

Purposeful exclusions

- generally, hard-to-enumerate populations; examples:
 - entire communities difficult to access because of terrain or internal security problems
 - transients
 - boat people
 - nomadic groups, especially those with no fixed residence
 - ethnic minorities that are more or less disenfranchised

Failure to adhere to sample design in field

- example is when interviewer selects households that are easy to enumerate in lieu of the ones designated for selection

Remedies for Non-sampling Errors

Questionnaire development and design

- use of standard protocols, instruments and proven modules wherever possible

Careful pre-testing

- any new question or set of questions, essential
- questions altered from standard instruments, recommended
- small, focus samples may be used for pre-testing
- purpose is to validate the concepts and wording in advance of survey
- pre-test used to refine concepts and correct detailed wording

Remedies for Non-sampling Errors *continued*

Use of probability sampling techniques at each stage of selection

Proper up-dating of old sample frames

Careful, intensive training of interviewers

- mock interviews
- re-training when necessary
- special attention to sampling aspects (listing of households, recording proper identification codes)
- special attention to problems of respondent cooperation and response rate
 - dealing with refusals
 - making call-backs
 - explaining purpose and confidentiality

Processing checks and verification plans

Issue of Non-response

Particularly insidious form of bias in surveys, if greater than a few percentage points

Extra efforts may be necessary to combat it

- intensified interviewer training
- publicity campaigns
- intercession by field supervisors

Issue of Non-response *continued*

Non-response of 10 -15 percent or more can ruin the probability nature of a survey and its unbiasedness

Methods of adjustment are inadequate

- most adjustment methods assume that non-responding population is distributed in same way as respondents

- one form of adjustment is to Aweight \cong the survey data to account for non-response, but it does not remove the bias

- another is to impute values for item non- +
++ ° 0
 ° 0
 +
° ++