

# Using Multiple Data Sources and Methods to Improve Estimates in Surveys

Clyde Tucker

# Contents

- Highlights from each paper
- Discussion relevant to CE

# Macklin Paper

- Uses subsample from a previously completed larger survey (NHIS) to conduct a more intensive analysis
- Specifically, collects additional data from the household (including details of health insurance), medical care providers for these households, Medicare records of the households, and Medicaid and Insurance information
- A number of methodological studies have resulted

# O'Brien Paper

- Using establishment surveys of rental agents and energy suppliers to supplement a household survey on residential energy consumption (RECS) and improve estimates of energy use
- Collect consumption and expenditures from linked suppliers with mandatory survey (large companies dominate) that uses multiple modes
- Can conduct methodological studies to measure response errors, improve questionnaire design and tailor the mode

# Schenker and Van Parsons Paper

- Combines data from NHIS, National Nursing Home Survey and NHANES to improve estimates from NHIS
- Combines NHIS with BRFSS to obtain small area estimates and improve BRFSS estimates
- Uses linkages of NCHS data to other Federal data sources for more in-depth causal analysis of health problems
- A number of technical difficulties noted

# Discussion with Relevance to CE Surveys

- **CAUTIONS:**

- All examples combinations of either directly collected Federal survey data (either household or establishment surveys) or other Federal sources of administrative data—no use of commercial databases.
- Two of the papers report their efforts have made possible a number of methodological studies.
- Only one, the Schenker and Van Parsons paper, provides detailed information on improvement of estimates.

# Discussion Continued

- A number of problems associated with using multiple sources are noted, including mode and questionnaire differences, sampling and weighting incompatibilities, privacy and confidentiality issues, linkage difficulties, increased agency efforts, and data sharing hurdles.
- No actual information on additional costs was provided.

# Discussion Continued

- **OPPORTUNITIES**

- CE should ask the authors for more details from the methodological studies they cite, especially concerning improvement in estimates
- CE should ask for estimates of the amount of increased effort and cost of the various methods used
- CE should ask for more information on the specifics of technical problems and how they were overcome

# Discussion Continued

- CE should consider collecting less data with the larger survey and use subsamples for more intensive data collection, including reconciliation of expenditure reports and using a subsample for Diary
- CE should consider using directly collected establishment surveys to supplement household data collection
- CE should search for other Federal data sources of either supplemental survey data or administrative data sets and extend the search to commercial databases

# Discussion Continued

- CE should consider using multiple modes and new technology (such as scanners) in the more intensive data collections with the subsamples
- CE should carry out pilots and small field studies to evaluate these methodologies over time using a phased introduction of the ones that prove to be the best