Assessing Measurement Error in the CE

Presented at the CE Survey Methods Symposium
Roger Tourangeau, Brandon Kopp, Lucilla Tan, and Scott Fricker
July 16, 2013
Introduction

• Goal is to develop specific measure that can be used on an ongoing basis to track measurement error in the Consumer Expenditure Survey (CE) over time

• As a practical matter, most of the measures proposed track overall error in the CE, not just measurement error
Background

• Many methods have been used to assess error in the CE, each with their strengths and weaknesses

• We recommend an multi-method-indicators (MMI) approach that consists into three main categories:
  — Internal indicators (based solely on CE data or paradata)
  — External indicators (compare estimates from the CE to an external data source)
  — Record check studies
Internal Indicators

• Some are based solely on CE data or information about the data collection process—Internal indicators

• Examples
  — Compare Interview Survey with Diary Survey
  — Compare across waves or across groups (for example, the proportion of CE respondents consulting bills or other records during the Interview Survey)
  — Develop latent class models based on several types of data/paradata
  — Use multi-level models to identify item or R characteristics associated with error

• Weakness—Can’t really estimate the magnitude of the errors (for example, proportion of Rs consulting records); these are indirect indicators
External Indicators

• Comparison to external data sources

• Two main external sources
  — Personal Consumption Expenditures from NIPA (National Income and Product Accounts)
  — Compare CE estimates with other surveys (e.g., MEPS, PSID, RECS)

• Weakness—Although PCE is covers many categories and a lot of work has gone into establishing “concordance” of PCE/CE categories, errors in PCE are not well established; not clear external benchmarks are really more accurate than the CE
Record Check/Validation Studies

• Compare CE reports to actual bills or other records

• In principle, this is a good method but it has many practical difficulties
  — Burdensome; likely to produce high rates of unit and item nonresponse
  — Some types of purchase unlikely to generate records
  — Recent feasibility study by Geisen and colleagues—Rs produced records for 36 percent of reported purchases
Selection Criteria for Internal Indicators

• **Sources of error**: Should give some indication of magnitude of different types of reporting error (forgetting, conditioning, satisficing)

• **Relation to error**: Should be predictive of level of error

• **Availability/stability**: Should tap data that will be available over time

• **Utility for improving items or survey procedures**: Should help identify problems that can be fixed
An Illustrative Set of Internal Indicators

• In the Diary Study, interviewer assessments of the diary keeper’s level of diligence in recording entries before pickup versus data collected by recall (recall error);

• The ratio of the number of entries in diary week one and diary week two (conditioning);

• The percentage of respondents who use records during the Quarterly Interviews (recall error);

• The length of interview (satisficing?);

• The average number of contact attempts needed to complete Quarterly Interviews (reluctance).
Selection Criteria for External Indicators

• Cover a range of categories, including some that differ in the likely availability of records;

• Include both regular (e.g., rent/mortgage, utilities) and irregular (e.g., clothing) expenditures;

• Include both large and small expenditures;

• Focus on categories in which the external source uses a definition that is reasonably consistent with the CE definition.
An Illustrative Set of External Indicators

• Comparisons with other surveys
  — ACS estimates for rent (6.1%) and mortgage (6.4%);
  — ACS estimates for utilities and fuel (7.5%);
  — Residential Energy Consumption Survey (RECS) estimates for utilities and fuel (7.5%);
  — MEPS estimates for hospitalization and health insurance (Healthcare 6.7%);
  — MEPS estimates for medical and health;
  — PSID estimates for medical and health.
An Illustrative Set of External Indicators

• Comparisons with the PCE
  — Household appliances (major and small appliances 0.6%);
  — Rent (6.1%) and utilities (7.5%);
  — Food purchased offsite (Food away 5.3%);
  — Women’s and girl’s clothing (1.5%);
  — Men’s and boy’s clothing (0.8%).
Record Check/Validation Study

• Use many of the same categories as in external Indicators to assess convergence

• An illustrative set:
  — Women’s and girl’s clothing;
  — Men’s and boy’s clothing;
  — Rent and utilities;
  — Food purchased offsite; and
  — Hospitalization and health insurance
Topics for Pre-Implementation Research

• Test protocols for obtaining records for more expenditure categories and for a higher percentage of survey reports
  — Should we try to collect records for all expenses or only for select categories?
  — What expenditure categories and what types of records raise privacy concerns?

• Determine the sample size for a records validation study

• Attempt to access respondents’ electronic records more effectively

• Develop improved methods to measure under- and overreporting of expenditures (as opposed to amounts)
Topics for Ongoing Research

• Examine interrelationships among indicators
  — Are the indicators unidimensional or multidimensional?
  — Factor analyze internal and external indicator values over time
  — Other approaches (like LCA) may be useful

• Ongoing research to identify more effective internal indicators will be needed, especially if the CE survey design changes.
Additional Considerations

• **Cost:** What inputs are needed to develop each MMI component?

• **Duration for development:** How long will the development efforts take?

• **Applicability:** Is the component applicable only to the current CE design or will it remain applicable to other designs?

• **Periodicity:** How often can the indicators be tracked?
Summary

• No one approach is perfect

• We recommend building on past efforts

• Develop a time series with multiple indicators
  — Internal indicators
  — External indicators
  — These are both inexpensive
  — Still, given the flaws, they should be supplemented with periodic (but regular) record check studies
  — Have overlapping expenditure categories to assess convergence across methods