

A Tale of Two Surveys: Learning from the Application of Address-based Sampling in Federal Surveys



*Federal Committee on Statistical Methodology Conference
November 4, 2013 | Washington, DC*

By

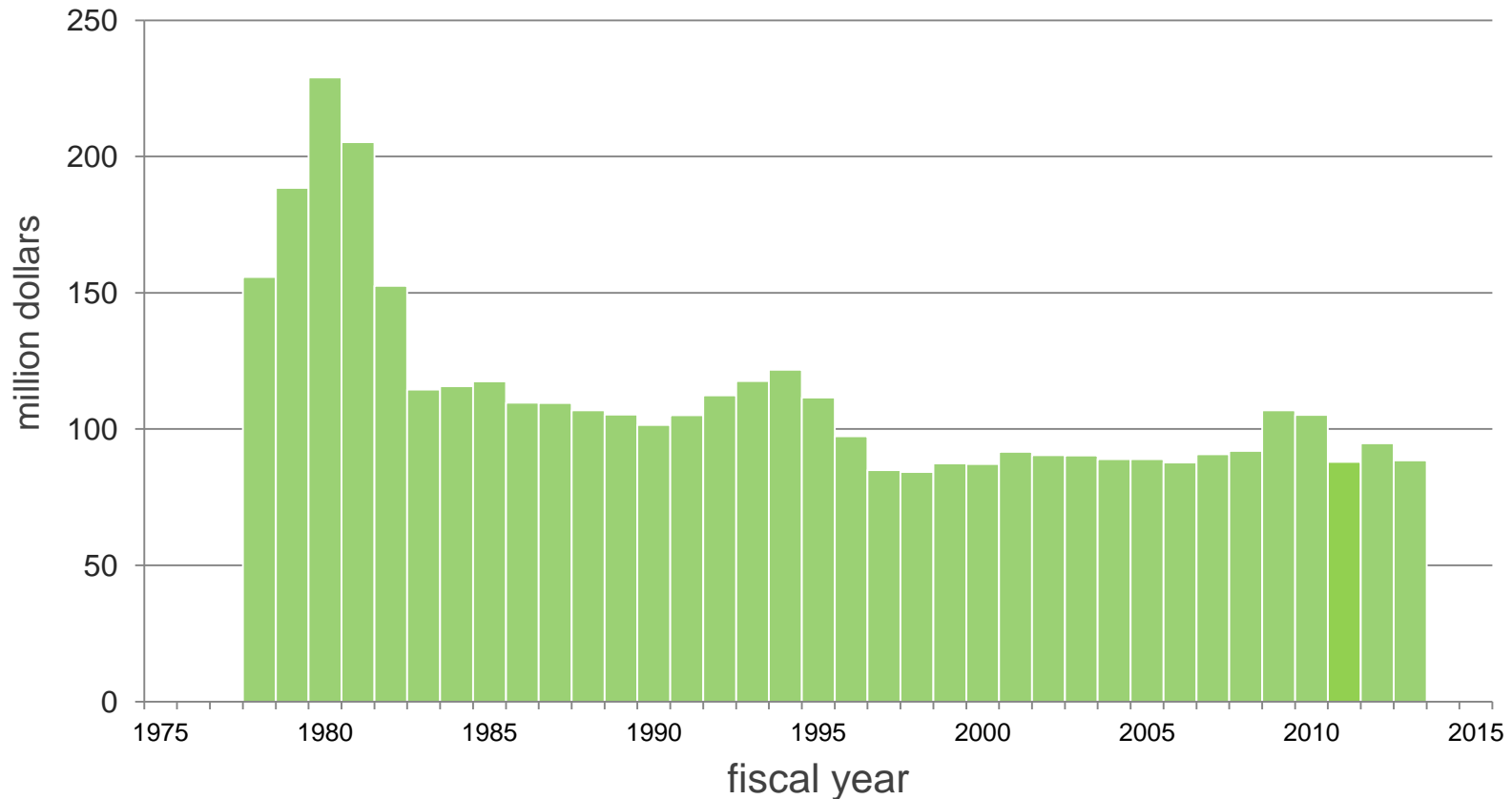
Eileen M. O'Brien, U.S. EIA

Overview

- Factors that led us to consider ABS methods in two federal surveys at U.S. EIA
- Challenges introduced
- Operational efforts to understand and refine the methods
- Suggestions for future ABS research and practices to better support surveys with unique mandates and design issues

EIA runs two complex periodic surveys of energy consumers against this budget profile

Agency funding (real dols, 2007 base)

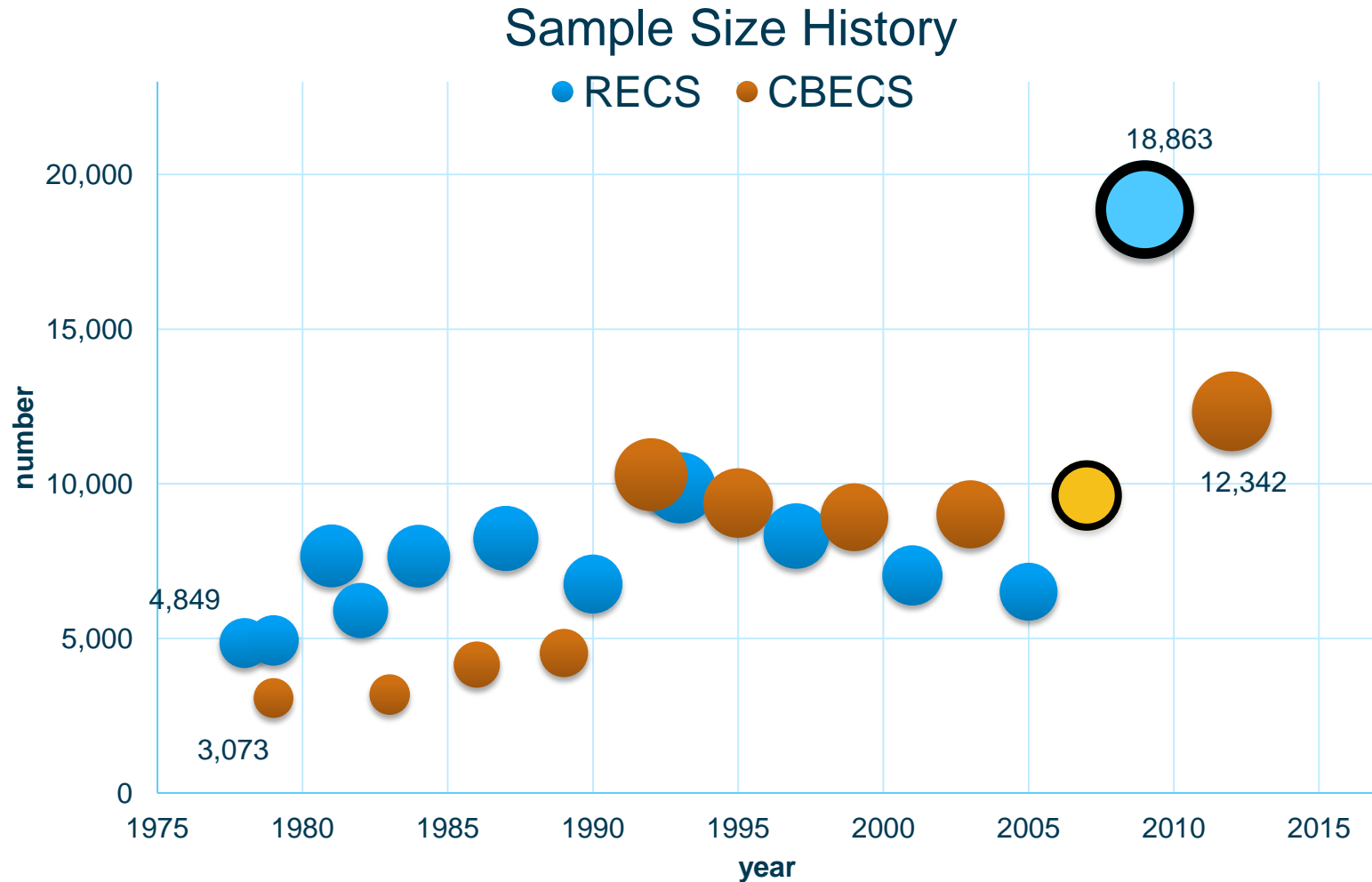


Source: OMB, Statistical Programs of the United States Government

Basic design features of RECS & CBECS

- Both use multi-stage area probability sample designs based on frames built by EIA
 - CBECS also uses a multiple *frame approach*—a deduplicated, hybrid frame of the area listed elements and special lists of large, complex buildings
- Both begin with in-person, voluntary CAPI field interviews
 - RECS contacts householders in housing units occupied as a primary residence
 - CBECS contacts key informants in “commercial buildings”
- Both have complex sample selection and enumeration stages
 - To reduce coverage biases, CBECS administers a special screener to every sampled building; RECS surveys Rental Agents for some apartment units
- Each is followed with a mandatory survey of energy suppliers

Over time, rising survey costs have eaten away at sample size, periodicity of surveys



What motivated us to consider ABS in these surveys?

- Investments in these sample frames are large relative to their frequency of use
- Best practices are moving the field to an ABS standard for frame updates
- The ABS frames are considerably better than 5-10 years ago for this purpose; it was time to try it

Question: Would ABS cost less in RECS than traditional listing methods for new construction?

Challenges introduced using ABS in RECS

- Great for creating *new* segments for the RECS sample expansion in 2009
- In hybrid segments, updated area segments from prior round for new construction using ABS lists. Deduplicating non-standard area addresses with ABS remains a problem.
- Forced a late decision to populate ‘hybrid’ segments only with ABS addresses.
- Some “rural” segments were still listed. There were typically segments where the ABS count varied substantially from modeled census counts.

Additional challenges introduced by ABS to CBECS

- CBECS studies the commercial sector— every thing that is not industrial, residential, or agricultural. There is no single source for a frame.
- The unit of analysis is a “building”, a structural concept to which energy systems are applied. “Building” is not an economic, financial, functional, or organizational concept.
- Not only are there no lists of buildings, but interviewers must get this concept right over several stages (area listing, prescreening, and interviewing stages) as must respondents.

Question: Use of ABS lists were likely to increase duplicates in the sample, but how much? Could it be managed?

Sample characteristics vary against ABS lists in ways that affect its utility for updates

Sample Characteristic	Residential (RECS)	Commercial (CBECS)
Unit of analysis	Housing unit occupied as a primary residence	“building”
Relationship of postal address to unit of analysis in survey	Generally 1 to 1	Highly variable for large, complex multi-tenant buildings, campuses and strip malls
Ease of removing duplicates introduced by ABS updates	High effort	High effort
Likelihood many duplicates would occur in selected sample	High, given limited time relative to survey start to complete deduplication tasks	Expected to be low, but highly correlated with complex buildings; recommended special field procedures be enacted
Impact of duplicates in field sample	Small, because we abandoned the hybrid approach and only used ABS address in segments updated for new construction. Coverage error is now, however, correlated with characteristics and age of segments.	Many duplicates retained in field sample; and field and processing procedures were insufficient to remove remaining duplicates. Attempted to remove post data collection.

Efforts to understand and refine use of ABS

- For RECS, EIA funded a record matching study to understand matching quality: a few findings...
 - DSF was excellent at adding addresses. This increased the number of non-occupied units to be handled by the field. Trade-off, however, seemed reasonable relative to costs that would have been incurred by relying only on traditional listing methods for the large RECS 2009 sample expansion.
 - The decision rules for traditional listing vs. ABS is complicated by the difficulty in matching addresses of different vintages and standards
 - Listers were best at fixing misclassified/misidentified units than missing units
- For CBECS, EIA's contractor took several steps:
 - BEFORE: Ran a small pretest of field procedures to understand sample quality after deduplication. While much duplication was detectable and removed, the report recommended added field and data processing procedures be developed and monitored.
 - AFTER: In editing and review stages, more duplication remained than expected. The contractor conducted an extensive, additional post-hoc review to detect and remove duplicates, and attempted to address the resultant weighting issues.

Recommendation: monitor these survey and sample performance metrics more

- Survey performance metrics relative to prior rounds
 - Response rate and its components, e.g., eligibility rate, unknown eligibility rates, rates for known ineligibles (vacant, seasonal, demolished units), etc.
 - Contact rates, cooperation rate, refusal rates
 - Completion rates between areas where frame may be known to perform better/worse relative to previous methods
- Sample performance metrics compared to prior designs
 - Sum of base weights; major shifts in weighting adjustments from initial base weights to final weights
 - Distribution of incoming sample relative to expected values in key estimation cells, unweighted and *weighted*, especially for any subpopulations
 - Sample efficiency (number of completes/sample drawn)

Conclusions

- Using ABS makes the most sense where the unit of analysis associates well with that on the list;
 - CBECS is not ready for ABS, or rather the field collection and data processes are not
- ABS should draw down the cost of making updates to RECS' sample frame; it might ultimately replace it
- ABS forces consideration of new modes for RECS, which was also expressed in a National Academy report on these EIA surveys

Ideas for future ABS research to support surveys like RECS & CBECS

- What is the coverage error between ABS and listed samples?
- Who might be reachable by mail in surveys of buildings?
- Record linkage seemed easier with ABS addresses in RECS 2009? Are there more opportunities for linkages?
- ABS methods open RECS up to new modes for precontacting households, data collection and follow-up
- More research through pilots, prototypes, and embedded research to support adjustments. Post hoc analysis is informative, but...