Innovating in U.S. energy consumer data collections

Three programs, three models for change

For

IEA G20 End-Use and Energy Efficiency Metrics Initiative

Session 3 – Developments in traditional data collection methods

November 21, 2019 | Paris

By

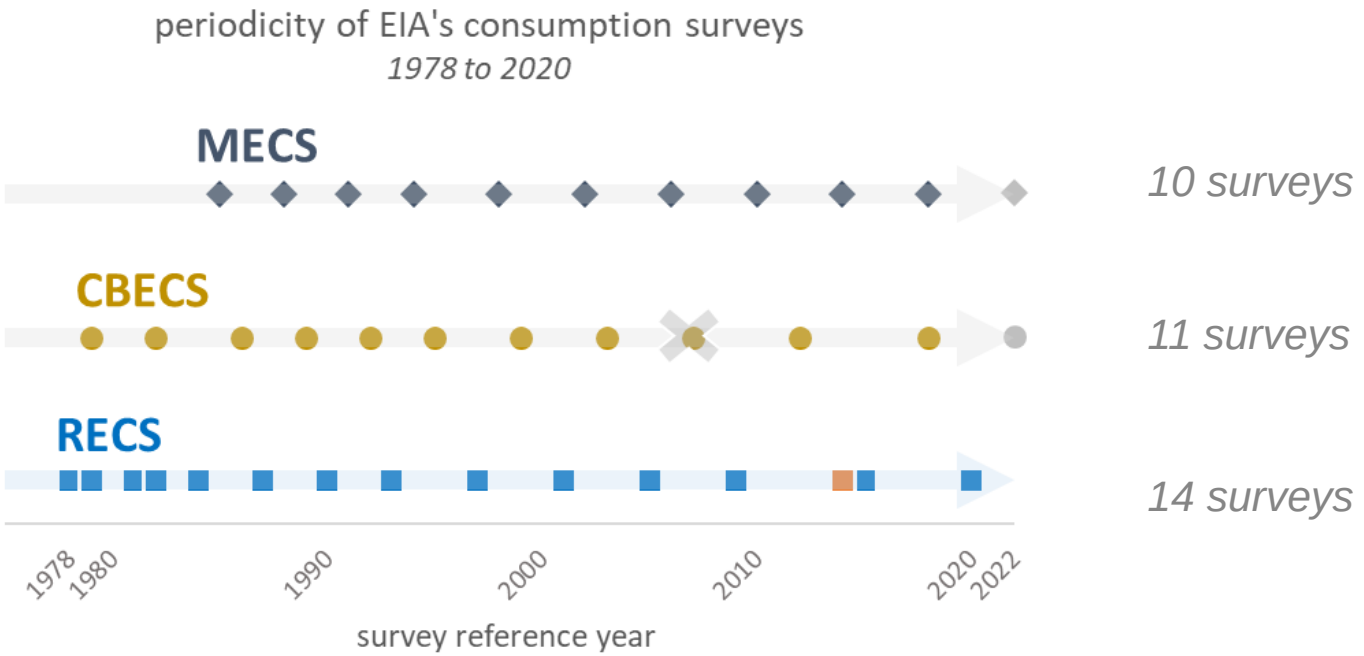
Eileen O'Brien, Program Lead

EIA is trusted. The 1977 law has the strongest provisions for independence of any statistical agency in the U.S.

> ''The Administrator shall not be required to obtain the approval of any other officer or employee of the Department in connection with the collection or analysis of any information; nor shall the Administrator be required, prior to publication, to obtain the approval of any other officer or employee of the United States with respect to the substance of any statistical or forecasting technical reports which he has prepared in accordance with law.''[2]

[1]Department of Energy Organization Act, Public Law 95-91, August 4, 1977; Sec. 205(a)(2).

With the award of the 2020 RECS, EIA will have conducted 35 consumption surveys in 40 years



Why innovate when…

- The statistical and survey methods are proven, best practices, and produce high quality at low risk; and

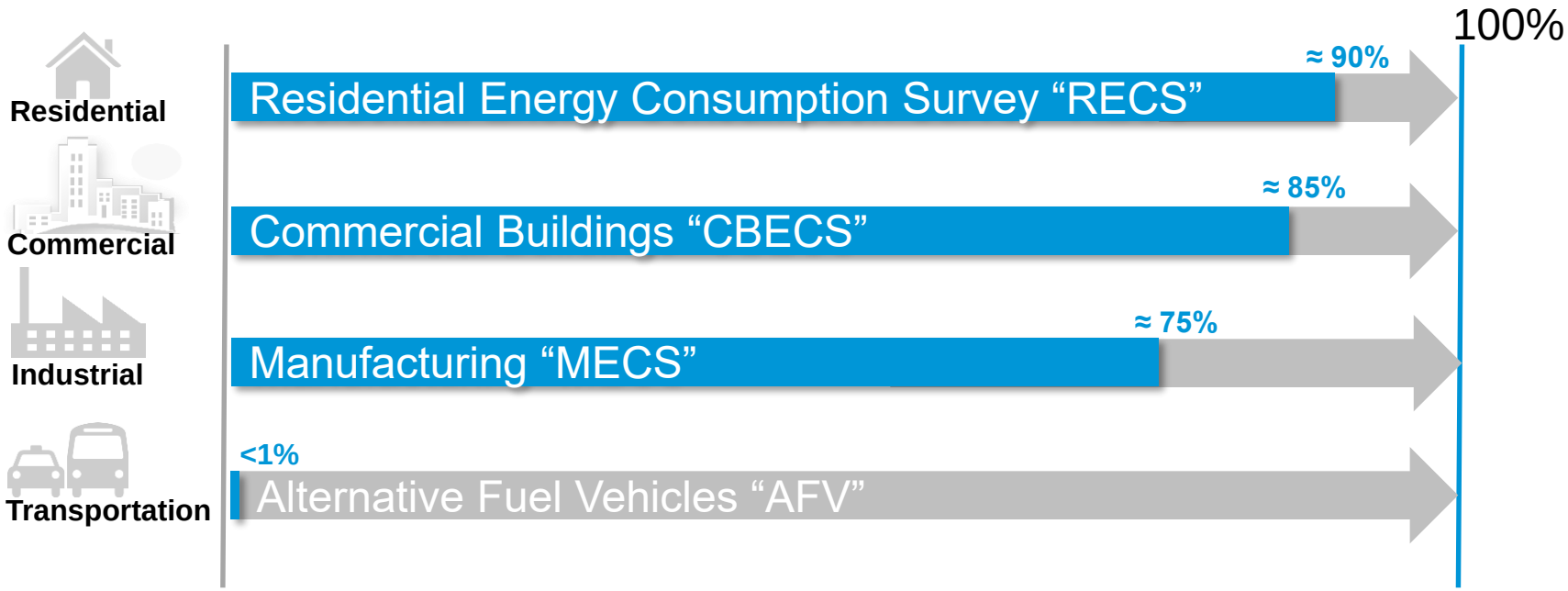
- We have limited resources and staff to devote to major program updates?

1. Potential to fill data gaps, add value to the program

2. Strong recommendations to update our methods

3. Continuity under rising costs, uncertain budgets

The consumption surveys cover varying shares of demand within sectors. New approaches may help fill data gaps.

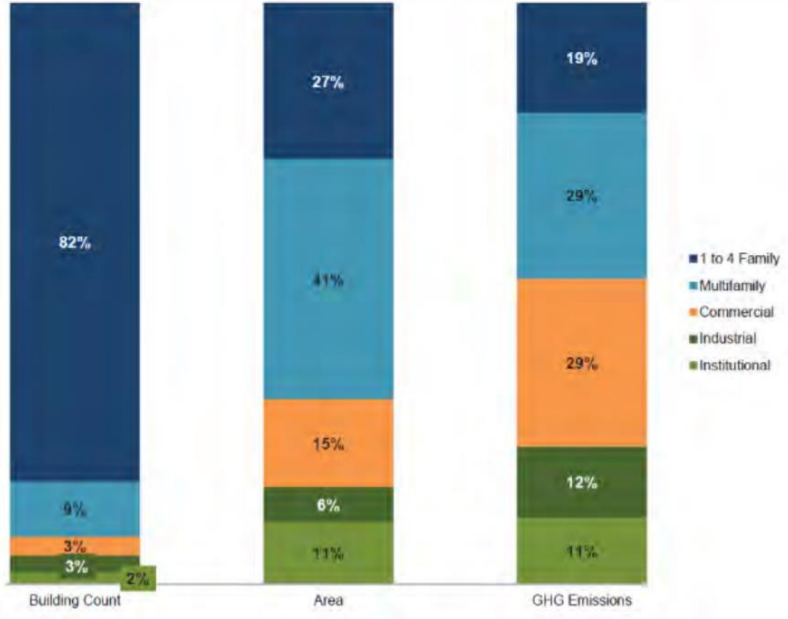


**These are approximations of site energy for major fuels consumed estimated by the consumption surveys, over primary energy plus electricity retail sales (MER Tables 2.2-2.5), referred to as ‘net energy’. This does not account for energy losses in commercial and industrial CHP and electricity-only plants and energy losses from fuel ethanol production.*

Differing impacts of data gaps: cities and multifamily buildings.

Key Findings about Building Energy Use

- **Greatest absolute number** of buildings: 1-4 family homes
- **Greatest share of GHG emissions**: Commercial and multifamily buildings



Source: PLUTO and 2015 GHG Inventory

Cassidy, R et al., "NYC's High Performance Retrofit Program and Resources", BuildingEnergy NYC 2019, September 26, 2019.
http://nesea.org/session/nycs-high-performance-retrofit-program-and-resources

EIA has acted on most of the recommendations of the National Academy's 2012 report[1]

- Rotating sample panels to improve timeliness
- ✓ Multimode data collections, including web
- ✓ Data processes for timelier data release
- ✓ Bigger samples sizes for more publishable data
- Research data centers for independent researchers
- Longitudinal component to provide estimates of change
- ✓ Capacity to collect and use smart meter data
- ✓ Regular updates to end uses
- ✓ Capacity to charge electric vehicles
- ✓ A 'whole-building' supplement for RECS
- ✓ Short- and long- forms to increase content and manage burden
- ✓ Regular updates to survey content and wording
- ✓ Administrative lists to construct sampling frames
- ✓ Centralized data collection for large building portfolios
- ✓ Collaboration with energy suppliers to improve efficiency
- ✓ Administrative data sources for program potential (local benchmarking, square footage)
- ✓ Use of Energy auditors on a subsample
- ✓ Regular interviewer debriefings to improve content

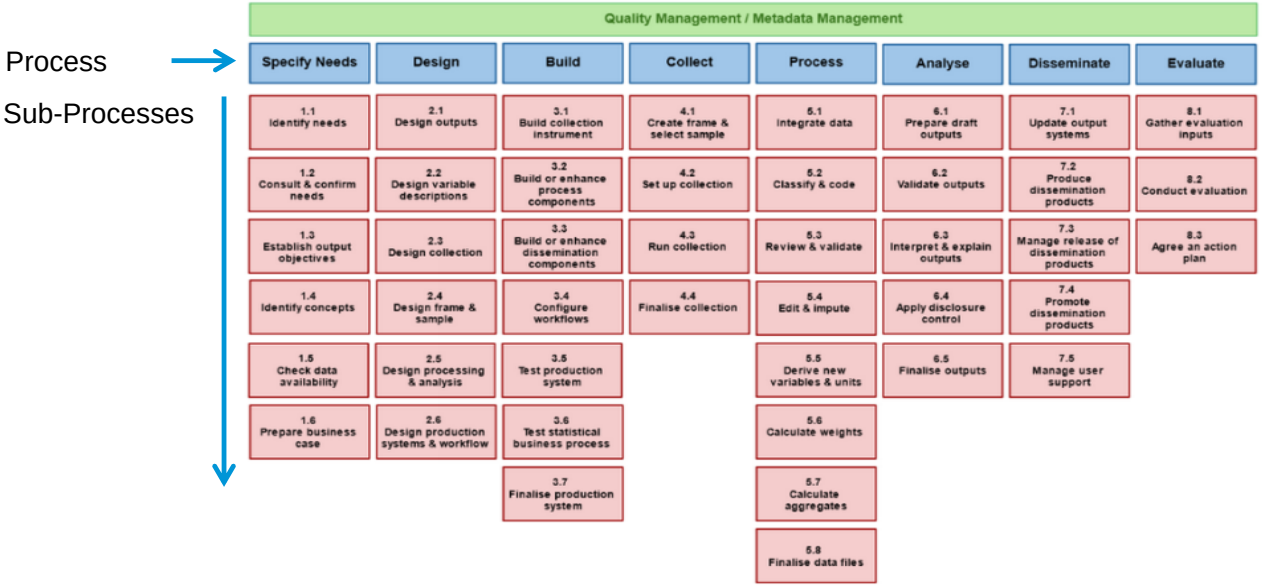
[1] *National Research Council. 2012. Effective Tracking of Building Energy Use: Improving the Commercial Buildings and Residential Energy Consumption Surveys. Washington, DC: NAP. https://doi.org/10.17226/13360.*

What did we update, how did we do it?

Best Practices at EIA: GSBPM

Benefits

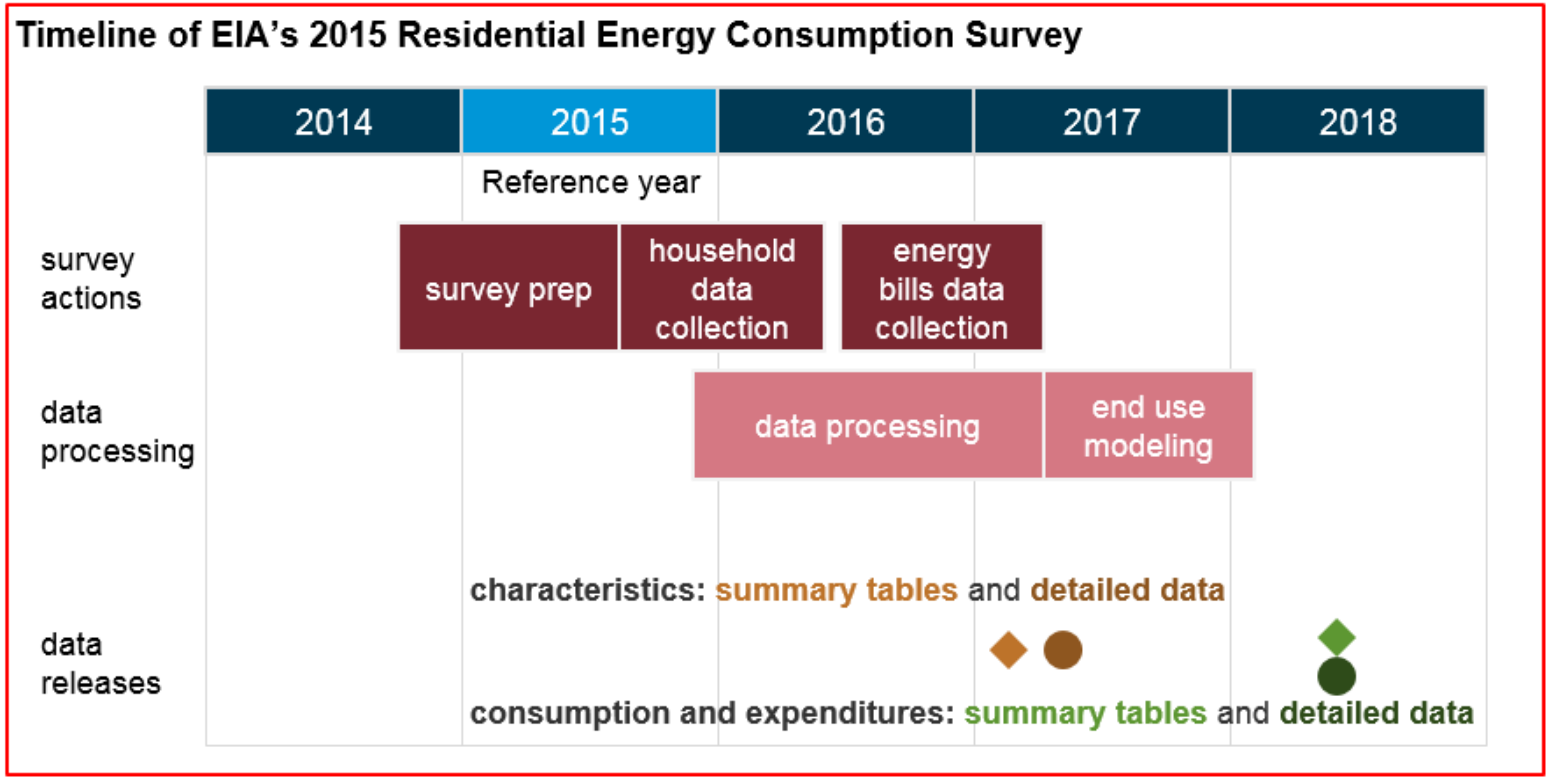
- ✓ Defines, describes and maps statistical processes in a coherent way
- ✓ Standardizes process terminology
- ✓ Compares / benchmarks processes within and between organizations
- ✓ Identifies synergies between processes
- ✓ Informs decisions on systems architectures and organization of resources
- ✓ Matrix Approach - there are many possible paths, including iterative loops within and between phases



EIA is now relying on multiple data collection modes for the consumption surveys. Example: CBECS

Mode	1979	1983	1986	1989	1992	1995	1999	2003	2007	2012	2018
Paper and pencil (PAPI)	X	X	X	X	X						
Telephone (CATI)							X			X	X
In-person (CAPI)						X		X	X	X	X
Web											X

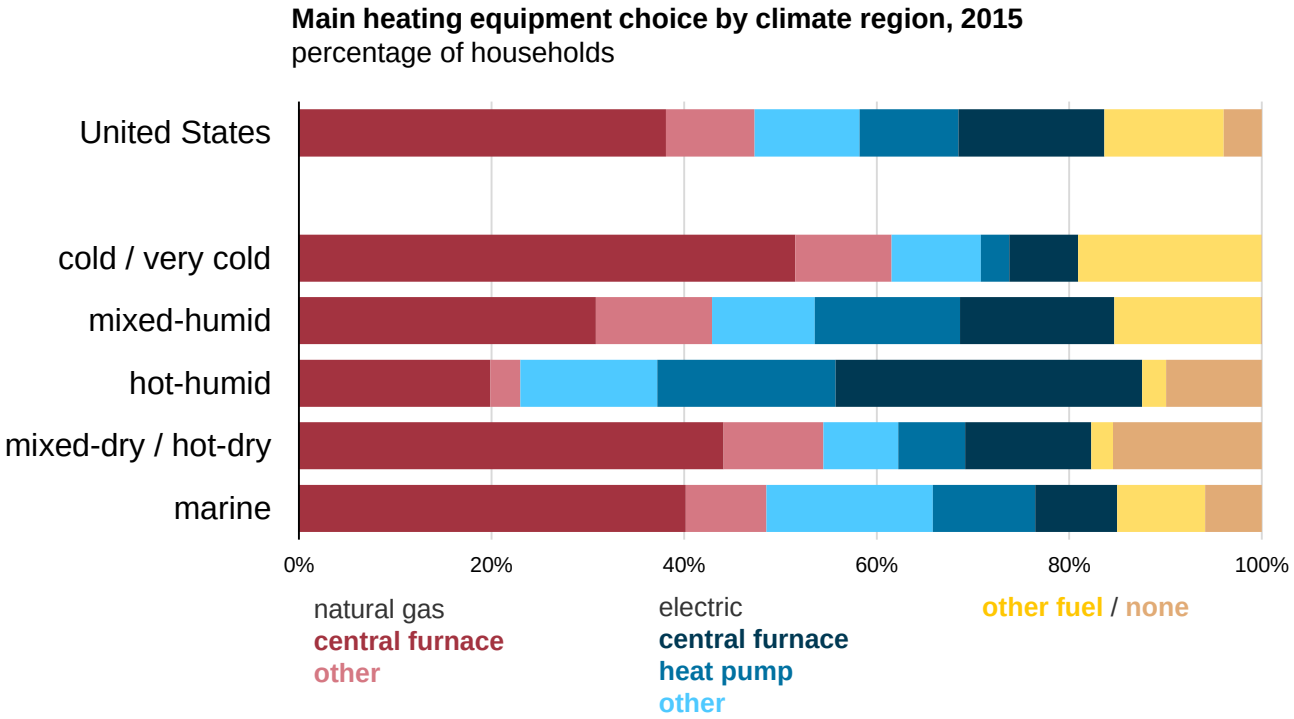
The final 2015 RECS estimates were the culmination of nearly 4 years of planning, collection, and modeling

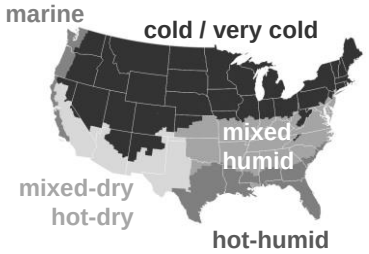


The 2020 RECS design will use web/mail strategies tested and deployed during the 2015 RECS

- 3 pilot tests of Web/mail proved those modes are viable options for future RECS
- Despite lower response rates, Web/mail responds are representative of the nation's households
- Web/mail data quality is comparable to traditional in-person response data quality

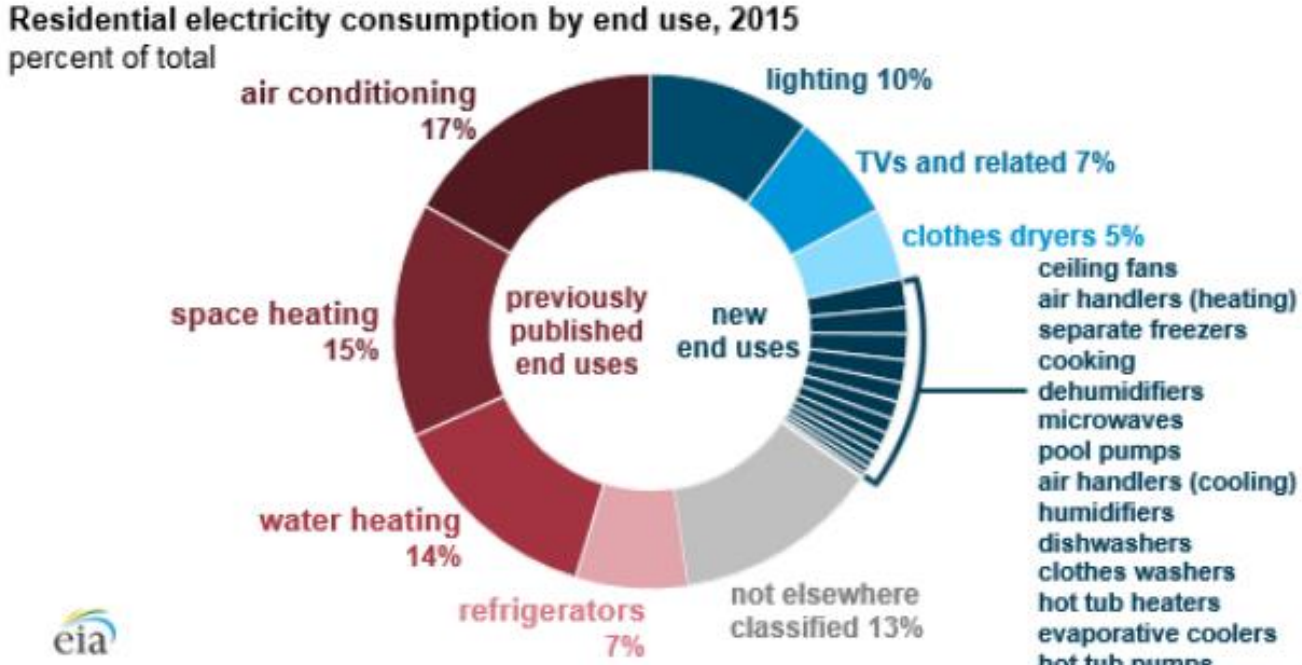
Heating equipment types vary across climate regions; capturing that variety is critical for end uses estimation.





Source: EIA, 2015 Residential Energy Consumption Survey

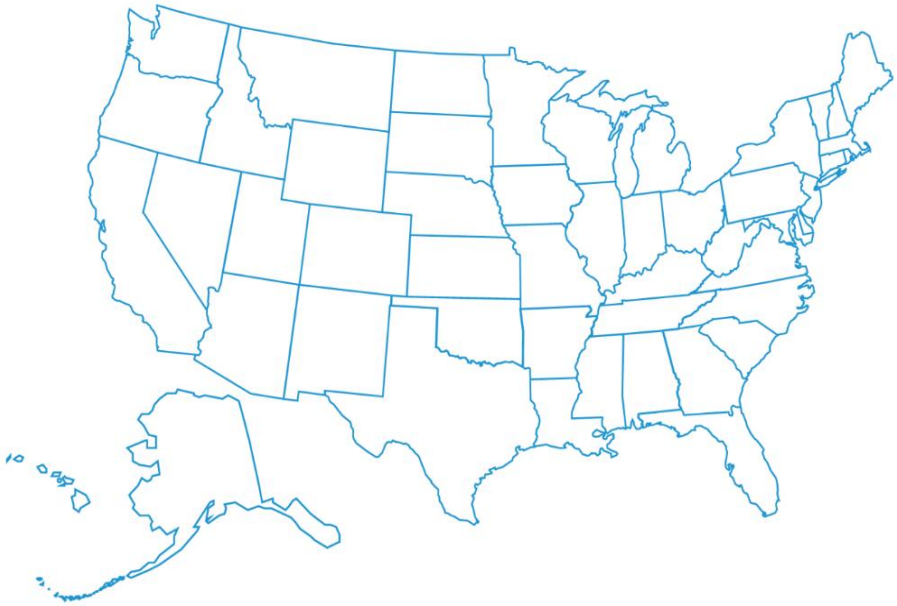
EIA expanded end-use estimation program in 2015



Source: U.S. Energy Information Administration, 2015 Residential Energy Consumption Survey

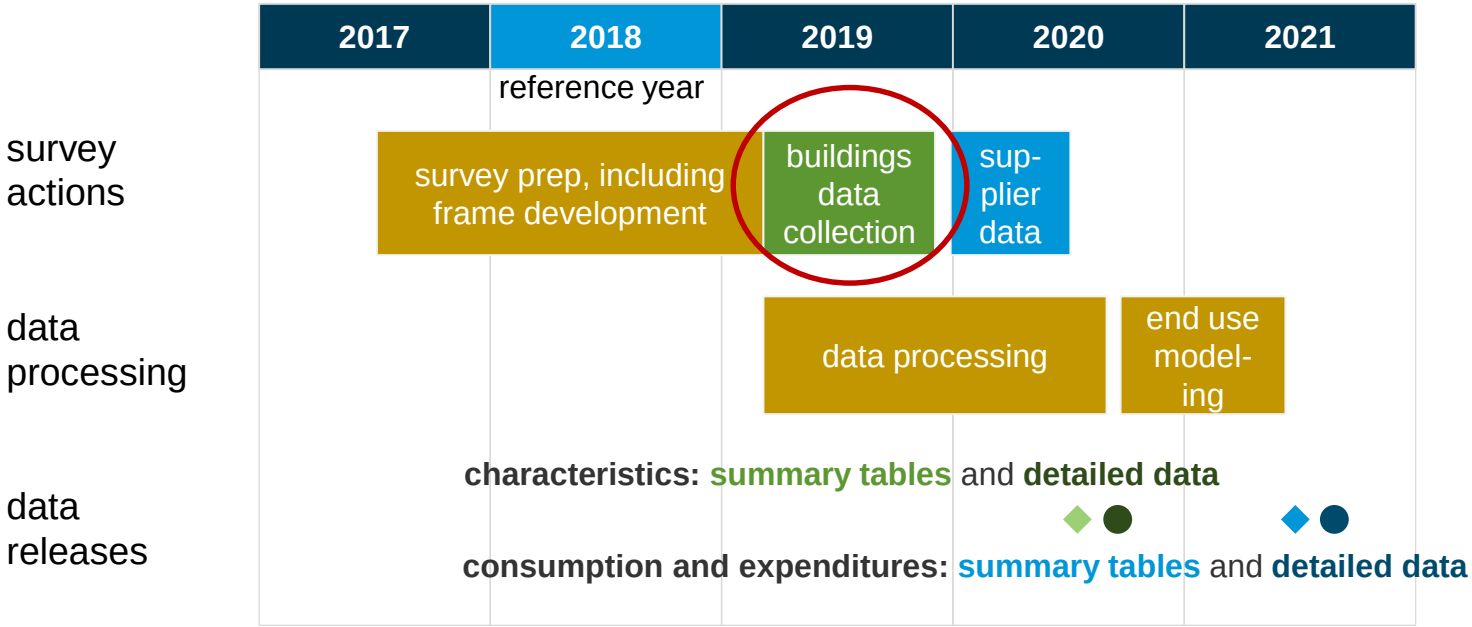
2020 RECS to include largest sample ever

- Key household energy metrics for all 50 states and DC
- Targeting responses from about 20,000 households
- Greater precision for ALL estimates
- Analysis of emerging topics
  - Electric vehicles
  - Electrification
  - Smart devices

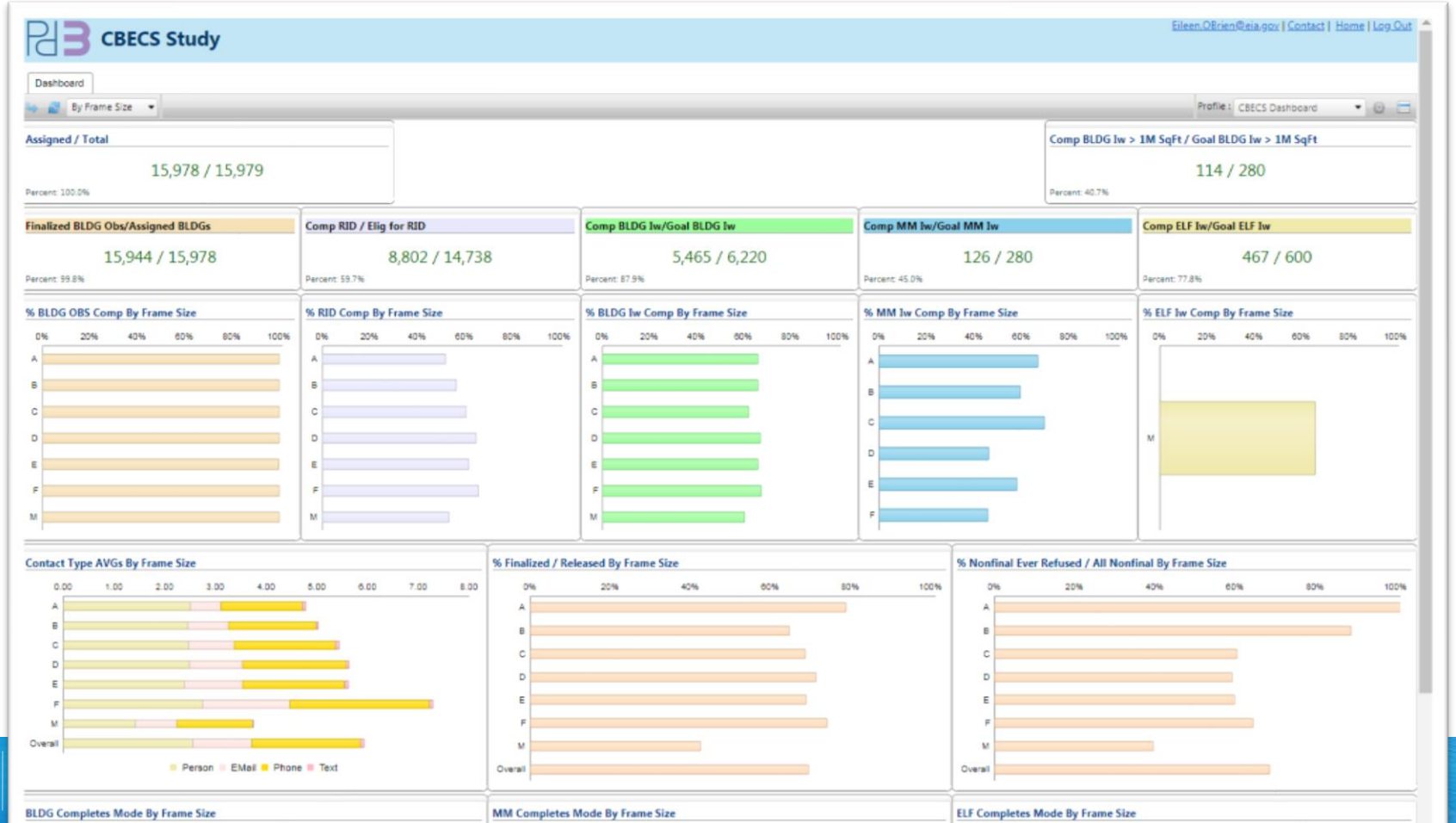


Data collection is *in progress* for the 2018 CBECS

Planned timeline for the 2018 CBECS

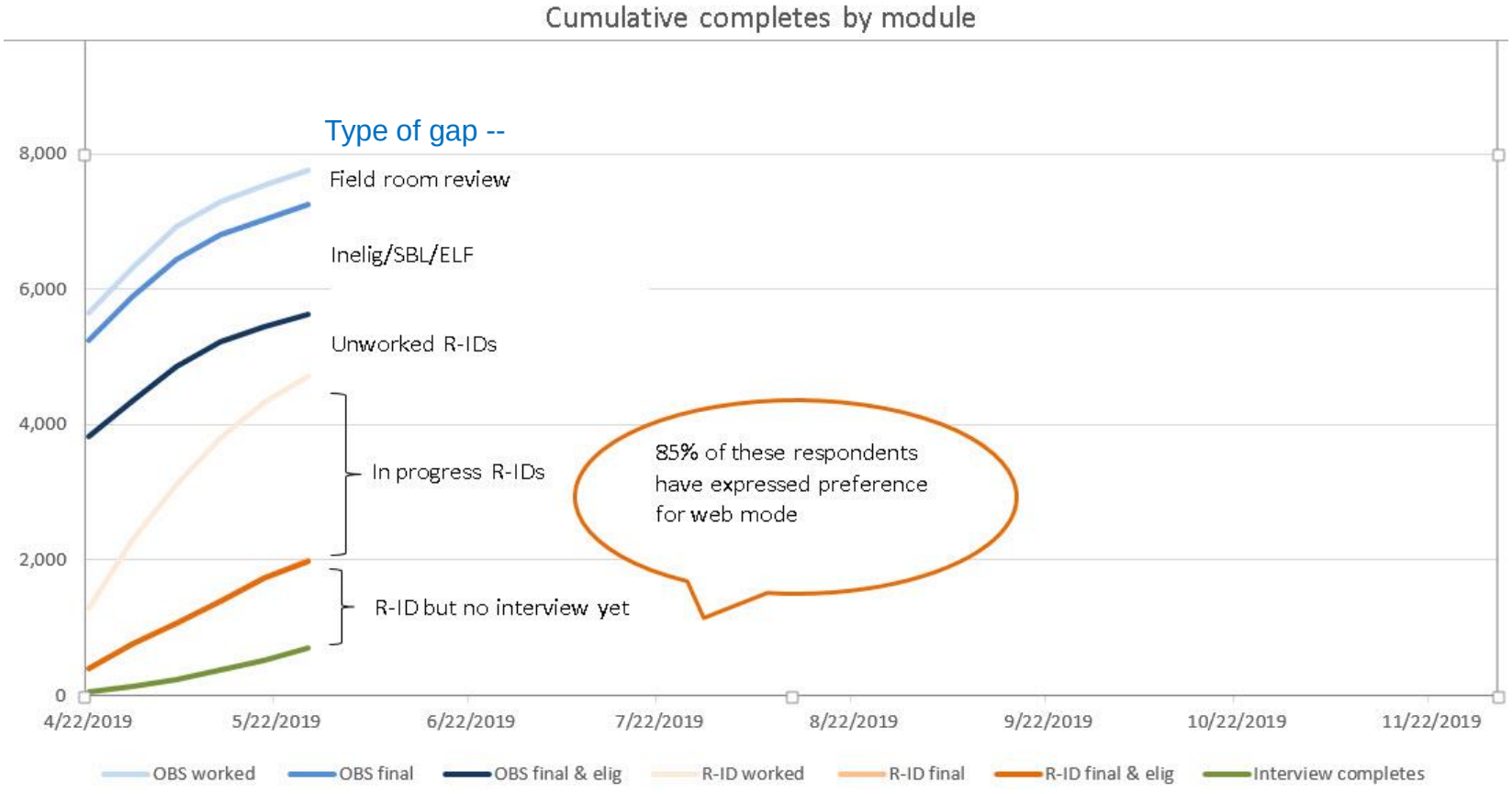


Real-time field production dashboard for CBECS



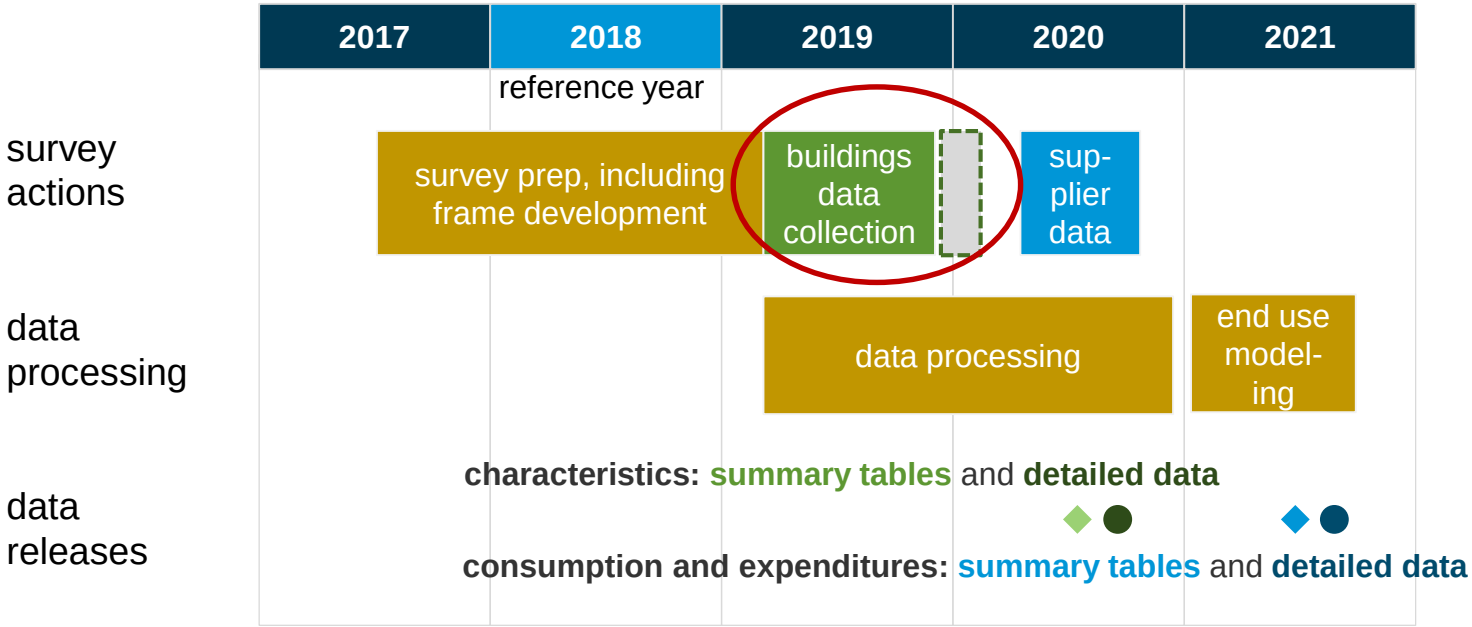
Changes in our business processes allow us to monitor production and costs more closely

Largest gap, most difficulty is in finding a respondent



The field period has been extended to increase response

Planned timeline for the 2018 CBECS



Response rate goals

- The overall response rate goal for the 2018 CBECS is at least 64 percent. The response rate goals by sampling frame are shown below.

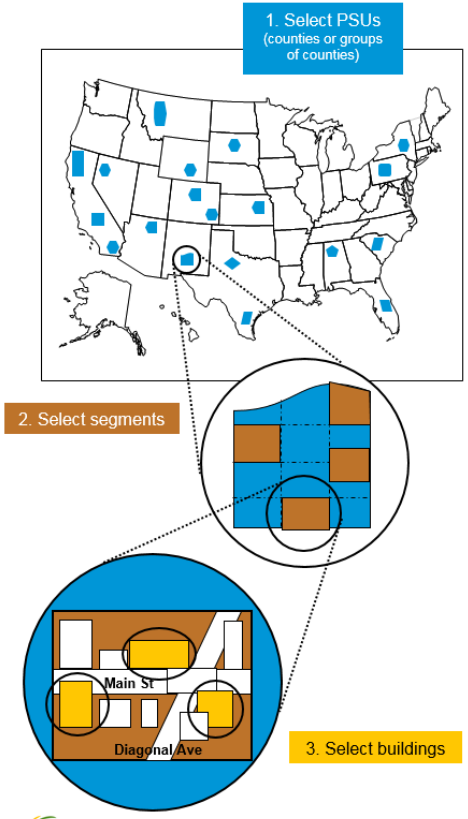
	Airport	Federal	College	Hospital	CPL	Area	Overall
Response Rate (AAPOR RR1)	95%	85%	90%	75%	65%	62%	64%
Completed Interviews	21	129	179	238	1813	5641	8021

CBECS uses a multi-stage area probability sample design supplemented with lists of large buildings

- No comprehensive frame of commercial buildings exists, so we have to build one for CBECS.

- Most of the frame elements come from the “area frame”, which has 3 levels of selection:
  - Primary sampling units (PSUs): counties or groups of counties
  - Segments within selected PSUs: Census block groups
  - Buildings within selected segments

- List frame of large buildings > 200,000 square feet
  - Sampled at PSU level
  - 5 sources: hospitals, colleges, GSA buildings, airports, and other large buildings (i.e. hotels, offices, malls, etc.)



Field listings of buildings were used to construct the area sample frame for all previous rounds of CBECS

- Trained field staff walk or drive every street in the segment, record key data about every “commercial” building:
  - Address
  - Square footage category
  - Building use
  - Number of floors

A	501 – 10,000 sq ft
B	10,001 – 25,000 sq ft
C	25,001 – 50,000 sq ft
D	50,001 – 100,000 sq ft
E	100,001 – 200,000 sq ft
F	200,001+ sq ft

- Field listing is expensive and time-consuming

- Estimating square footage is a complex, necessary task– building size is highly correlated with energy use

The development of a Virtual Listing System (VLS) is a new alternative to field listing

- EIA's 2018 CBECS data collection contractor developed the Virtual Listing System (VLS), a fully integrated, custom web-based system for observing, cataloging, and documenting CBECS-eligible buildings.

- The VLS frame development uses Google Maps
  - 2-D and 3-D aerial images
  - Google StreetView

- Customizations enhanced the quality of the frame
  - Integrated alternate sources of commercial building information
  - Added a "building footprint" that stores the relevant building characteristics
  - The technology included the ability to automatically calculate square footage
  - Easy to keep track of which blocks are complete so no blocks are missed

Square footage calculation

- Done automatically: footprint area x number of floors

- VLS allowed lister to override automatic calculation for special situations

EIA conducted an evaluation of virtual listing methods for constructing an area probability sample frame for buildings

- Over a 2 month period in 2018, traditional field listing and virtual listing was done on 59 segments

- Listings from both methods compared which buildings were found--
  - In both methods
  - In the field not VLS
  - In VLS but not the field

- Compared hours worked, minutes per building listed, costs

Coverage comparison – VLS found more buildings

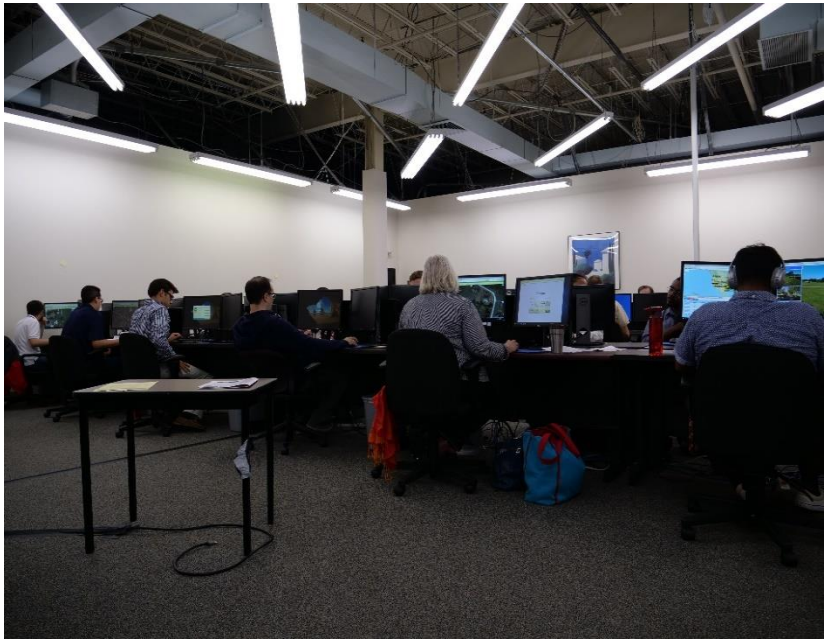
- 69% of the total buildings listed were found with both methods
- 20% were only found using the VLS methods
- 11% were only found with traditional on-the-ground field listing
- VLS found more buildings in every building size and use category
- Smaller buildings and warehouse/vacant buildings most likely to be missed by both methods

Comparing cost/effort, the VLS is much less expensive

- Field listing hours worked were more than double hours for virtually listing
- Field listing averaged 15.3 minutes versus 6.3 minutes per building in VLS
- There were zero travel costs for VLS methods
- It was easier to supervise VLS listers; all VLS listers worked in a central location. Questions were addressed answered in real-time.
- Automatic calculation of square footage provided removed the ‘guess work’ of listing by providing more structured and consistent building size estimates

VLS was used for most of the 2018 CBECS area frame

- Based on the test of 59 segments, VLS was deemed ready for production
- Most of the remaining segments (700+) were listed with VLS, except 26 segments with poor satellite/Street View coverage
- Listing was completed in summer 2018 by college students enrolled in GIS programs
- 225,000 building listed in the VLS area frame; 16,000 buildings selected for the sample overall from all frames



Future plans for U.S. EIA's energy consumption data program

We're proofing methods to push flexibility in periodicity, granularity, topical coverage, and specificity. Why?

- Energy end uses are changing rapidly (e.g., EVs, "smart" devices, LED, and displacement, substitution, upgrades, "electrify and decarbonize")
- Changes in energy consuming behavior create challenges in demand-side management, pricing, and reliability (e.g., peaks early in day and at night)
- Changes in technology, incentives, and/or capacities to manage daily peaks in demand (integrated systems, user interaction, messaging…)
- Energy sources are diversifying behind the meter (back-up power, battery storage, and onsite generation, e.g., PV)

Future directions

- Modernize sources and methods for the commercial buildings data program
- Produce residential demand data products and datasets for 50 states and possibly Puerto Rico
- Expand end-use estimation into more areas, identify and integrate objective data sources from new digital technologies
- Fill data gaps—e.g., multifamily buildings, commercial building tenant energy use
- Biennial MECS at the national level

EIA is busy! Key Milestone Events, Fiscal Year 2017 to 2021

Program	FY16 (actual)	FY17 (act., proj)	FY18	FY19	FY20	FY21
MECS	Post-collection Processing	Publication (2014 data)	Planning/Prep for 2018 collection	Fieldwork	Post-collection Processing	Publication (2018 data)
RECS	Fieldwork (pilots, CAPI)	Fieldwork (suppliers)	Publication (2015 data)	Planning/Prep for 2020 collection	Fieldwork (Households)	Fieldwork (Suppliers), Post-collection processing
CBECS	No activity	Award, Begin 2018 project	Planning/Prep for 2018 collection	Fieldwork (Buildings)	Fieldwork (suppliers), Post-collection processing	Publication (2018 data)

A few references

Commercial Buildings Energy Consumption Survey (CBECS) methods

- Lewis, K. and J. Michaels, 2019. “How Were Buildings Selected for the 2018 CBECS?” https://www.eia.gov/consumption/commercial/reports/2018/methodology/sampling.php.
- Lewis, K. and J. Michaels, 2019. “Virtual listing in the 2018 CBECS”, U.S. Energy Information Administration, https://www.eia.gov/consumption/commercial/reports/2018/methodology/vls.php.
- “2012 CBECS Survey Data Methods”, U.S. Energy Information Administration, https://www.eia.gov/consumption/commercial/data/2012/index.php?view=methodology.

Residential Energy Consumption Survey (RECS) methods

- Lewis, K., 2019. “2015 RECS Square Footage Methodology”, U.S. Energy Information Administration, https://www.eia.gov/consumption/residential/reports/2015/squarefootage/
- “2015 RECS Survey Data Methods”, U.S. Energy Information Administration, https://www.eia.gov/consumption/residential/data/2015/index.php?view=methodology.
- “RECS 2015 Consumption and Expenditures Technical Documentation Summary”, U.S. Energy Information Administration, https://www.eia.gov/consumption/residential/reports/2015/methodology/pdf/2015C&EMethodology.pdf

Other

- National Research Council. 2012. *Effective Tracking of Building Energy Use: Improving the Commercial Buildings and Residential Energy Consumption Surveys*. Washington, DC: The National Academies Press. https://doi.org/10.17226/13360.
- Address based sampling Harter, R. et al., 2016. “Address-based Sampling. American Association for Public Opinion Research Task Force Report” https://www.aapor.org/Education-Resources/Reports/Address-based-Sampling.aspx
- Olson, K. et al., 2019. “Transitions from Telephone Surveys to Self-Administered and Mixed-Mode Surveys. American Association for Public Opinion Research Task Force Report”, https://www.aapor.org/Education-Resources/Reports/Transitions-from-Telephone-Surveys-to-Self-Adminis.aspx