U.S. Energy Information Administration

Improving the Quality and Scope of EIA Data

July 2011

Independent Statistics & Analysis

www.eia.gov

U.S. Department of Energy
Washington, DC 20585

This report was prepared by the U.S. Energy Information Administration (EIA), the statistical and analytical agency within the U.S. Department of Energy. By law, EIA's data, analyses, and forecasts are independent of approval by any other officer or employee of the United States Government. The views in this report therefore should not be construed as representing those of the Department of Energy or other Federal agencies.

Contents

Executive Summary

Section 805(a) of the Energy Independence and Security Act of 2007 (EISA), Public Law 110-1401, requires the U.S. Energy Information Administration (EIA) to establish a five-year plan to enhance "the quality and scope of its data collection necessary to ensure that the scope, accuracy, and timeliness of the information needed for efficient functioning of energy markets and related financial operations."

In establishing the plan under section 805(a), EIA is to pay particular attention to six areas of data:

1. data series terminated because of budget constraints;
2. data on demand response;
3. timely data series of State-level information;
4. improvements in the area of oil and gas data;
5. improvements in data on solid byproducts from coal-based energy-producing facilities; and
6. the ability to meet applicable deadlines under Federal law (including regulations) to provide data required by Congress.

This report is in response to section 805(b) of EISA which calls on EIA to submit to Congress the plan established under subsection (a), "including a description of any improvements needed to enhance the ability of the Administrator to collect and process energy information in a manner consistent with the needs of energy markets."

EIA Background

EIA is the statistical and analytical agency within the U.S. Department of Energy. EIA collects, analyzes, and disseminates independent and impartial energy information to promote sound policymaking, efficient markets, and public understanding of energy and its interaction with the economy and the environment. EIA is the Nation's premier source of energy information and, by law, its data, analyses, and forecasts are independent of approval by any other officer or employee of the United States Government.

The Department of Energy Organization Act of 1977 established EIA as the primary Federal Government authority on energy statistics and analysis, building upon systems and organizations first established in 1974 following the oil market disruption of 1973. EIA conducts a comprehensive data collection program that covers the full spectrum of energy sources, end uses, and energy flows. EIA also prepares informative energy analyses, monthly short-term forecasts of energy market trends, and long-term U.S. and international energy outlooks. EIA disseminates its data, analyses, and other products primarily through its website and customer contact center. Major users of EIA's work products include the Congress, Federal and State government, industry, academia, financial institutions, news media, and the public. By law, EIA's products are prepared independently of Administration policy considerations.

The Federal Energy Administration (FEA) Act of 1974, Public Law 93-275, as amended, and the Department of Energy (DOE) Organization Act of 1977, Public Law 95-91, as amended, are the main laws providing EIA with authority for its mandatory data collection and analytic activities. The majority of EIA surveys and analysis products are based on the general mandates in these two laws; however, there are some surveys specifically mandated by law. A complete list of surveys and their collection authority can be found in Appendix B.

Recent Changes to Data Collection Program and Priorities

The energy field is frequently changing, and EIA adjusts its data collection activities to capture the information necessary to fulfill its mission and statutory obligations. Over time, EIA has eliminated a few surveys, added several surveys, and consolidated and streamlined others.

EIA has eliminated the following surveys in recent years:

- EIA-412 *Annual Electric Industry Financial* Report was terminated in 2005 due to resource constraints.

[1] For the text of EISA section 805, see Appendix A.

- EIA-6A *Annual Coal Distribution Report* was terminated in 2008, with a few critical items relocated to other coal survey forms, resulting in cost savings and reduced respondent burden.
- EIA-876 *Residential Transportation Energy Consumption Survey* was discontinued after the 1994 data year due to resource constraints.

EIA has a set of surveys in place that collect detailed data on end-use energy consumption: the EIA-457 A-G, *Residential Energy Consumption Survey* (RECS); EIA-871 A-F, *Commercial Buildings Energy Consumption Survey* (CBECS); and the EIA-846, *Manufacturing Energy Consumption Sur*vey (MECS). These surveys are required under Section 205 of the DOE Organization Act, which calls for the MECS to be a biennial survey and for the RECS and CBECS to be triennial. The RECS and CBECS are labor-intensive surveys conducted through visits to a sample of residential and commercial buildings throughout the country by contracted field staff and thus expensive to conduct. Because of resource constraints, the RECS, CBECS, and MECS are currently on a quadrennial (every four years) schedule.

EIA has initiated several new surveys during the last 10 years to capture new developments in energy markets or respond to specific requests, including:

- EIA-912 *Weekly Underground Natural Gas Survey* began in 2002 at the request of the Secretary of Energy when a comparable survey conducted by a trade association was phased out due to liability concerns. These data were designated a Principal Economic Indicator in 2008.
- EIA-914 *Monthly Natural Gas Production Report* was initiated in 2005 to improve the quality of our monthly natural gas data by obtaining data directly from the operators. After this survey was established, EIA discontinued the monthly survey that had obtained this information from State agencies (the EIA-895M Monthly Quantity and Value of Natural Gas Report).
- EIA-757 *Natural Gas Processing Plant Survey* was started in 2008 to collect baseline plant level information on the capacity, status, and operations of natural gas processing plants every three years. Updated information will be obtained during periods of supply disruption in areas affected by an emergency, such as a hurricane.
- EIA-22M *Monthly Biodiesel Production Survey* began in 2009 to collect monthly data on production, prices and feedstocks, as specified in section 1508 of the Energy Policy Act of 2005.
- EIA-809 *Weekly Oxygenate Report* began in 2010 to provide weekly data on ethanol production and stocks to meet the requirements of section 1508 of the Energy Policy Act of 2005. EIA also added data on ethanol to several other petroleum supply weekly and monthly surveys.

In addition, EIA has consolidated and streamlined several surveys. Notably, in 2008 five surveys were combined into the new EIA-923 *Power Plant Operations Report*. EIA has expanded other surveys, including the EIA-851A *Domestic Uranium Production Report (Annual),* which now also collects uranium reserves information.

For the next five years and beyond, within the context of limited resources, EIA has several areas of emphasis for its data collection and analysis program:

- *Improve the coverage, quality, and integration of* ***energy statistics***. EIA is committed to improving the coverage, quality, and integration of energy demand, supply, and market data. EIA's aging information technology infrastructure is poorly adapted for keeping up with the changing information needs of policymakers, the broader energy industry and its associated markets. All aspects of developing and disseminating data have been affected, from maintaining survey frames (i.e., lists of possible respondents), to collecting and processing data, to analyzing the data once acquired, to providing information to the public. EIA is currently undergoing a modernization effort to fill critical data gaps; update the statistical techniques used in data collection; better protect the integrity of data collected, processed, and published; assure documentation of data processing decisions; and reduce lifecycle development and operating costs for EIA's statistical programs.
- *Strengthen the relevance and breadth of* ***energy analysis***. A multiyear project is underway to replace the components of the National Energy Modeling System to improve EIA's ability to assess and project supply, demand, and technology trends affecting U.S. and world energy markets. In addition, EIA's Energy and Financial Markets Initiative, launched in September 2009, aims to improve energy market

transparency, support sound policy and efficient markets, and increase public understanding — activities that are central to EIA's mission. EIA's traditional coverage of physical fundamentals such as energy consumption, production, inventories, spare production capacity, and geopolitical risks continues to be essential, but, moving forward, EIA is also assessing other influences, such as speculation, hedging, investment, and exchange rates, as it seeks to fully understand energy price movements.

- *Enhance the **communication** of EIA's work to diverse audiences*. EIA is continually working toward increasing the accessibility, relevance, and understandability of our energy information. EIA's ongoing efforts include a major overhaul of our website to make our data and analyses easier to find and use. EIA is utilizing user assessment tools including usability testing, a "star-rating" system for web pages and customer surveys to collect feedback. EIA is actively participating in the Data.gov initiative to provide open access to datasets. EIA's energy education and public information initiative includes a number of award-winning products written in plain language, including Energy Explained, Energy Kids, and Energy in Brief.
- *Excel in our **management** of people, finances, and technology.* In October 2010, EIA changed its organizational structure to align with its four primary functional areas: statistics, analysis, communication, and management. The organizational change creates a more effective platform for meeting EIA's mission requirements; strengthens the information collection, analysis, reporting and dissemination functions; improves customer service; and will result in more efficient operations. EIA is also implementing significant improvements to procurement, budgeting, human resources, and information technology business processes.

The bulk of this report discusses recently completed, ongoing, and potential future improvements to aspects of EIA's program. It is preceded by a short description of data collections that have been terminated in recent years or proposed for termination, but subsequently retained.

1. Data Collections Terminated

EIA has eliminated the following surveys in recent years:

- EIA-412 *Annual Electric Industry Financial Report* was terminated in 2005 due to resource constraints.
- EIA-6A *Annual Coal Distribution Report* was terminated in 2008, with a few critical items relocated to other coal survey forms.
- EIA-876 *Residential Transportation Energy Consumption Survey* was discontinued after the 1994 data year due to resource constraints.

Two surveys were proposed for termination in 2006, but then retained. In 2006, EIA announced that it would suspend collection of the EIA-182, *Domestic Crude Oil First Purchase Report*, and EIA-856, *Monthly Foreign Crude Oil Acquisition Report*, due to resource constraints. EIA received numerous letters from Members of Congress and other major stakeholders requesting the continuation of these surveys in light of oil market developments and the importance of the information from these two surveys. Funding provided by Congress in 2007 allowed for the continuation of these surveys.

2. Improvements to Data Relevance and Reliability

2.1 Electricity, Coal, Nuclear, and Renewables

EIA collects and disseminates information from six electric power surveys, four coal industry surveys, and three renewable energy surveys. They include weekly, monthly, quarterly, and annual surveys.

2.1a Recently Completed Improvements

Consolidation of electric power surveys. In 2008 four surveys (*EIA-423 Monthly Cost and Quality of Fuels for Electric Plants Report, EIA-767 Steam-Electric Plant Operation and Design Report*, *EIA-906 Power Plant Report*, and *EIA-920 Combined Heat and Power Plant Report*) were merged into the new EIA-923 Power Plant Operations Report and the existing EIA-860 Annual Electric Generator Report.

Uranium production and reserves. The EIA-851A *Domestic Uranium Production Report (Annual)*, was expanded to collect uranium reserves information in addition to production.

Solid by-products from coal-based energy producing facilities. Section 805(a) of the Energy Independence and Security Act of 2007 (EISA) requires that EIA collect data on solid by-products from coal-based energy producing facilities, which were previously collected on the *Form EIA-767, Steam-Electric Plant Operation and Design Report.* These data were collected on that form through the 2005 data collection year and then terminated due to resource constraints. However, when EIA requested three-year approval of its electric power surveys from the Office of Management and Budget (OMB) in 2007, these data were added to the new Form EIA-923, *Power Plant Operations Report* and will continue to be collected.

2.1b Ongoing and Potential Future Improvements

Demand response data. Demand response has become an important strategy for reducing peak electricity demand, and its increased use would offset the need for additional electric generating capacity. EIA has been collecting some Demand Side Management (DSM) data for over 15 years, including information on peak load reductions, energy savings, and costs. In October 2010, OMB approved a three-year extension of EIA's electric power surveys, including an expansion of the DSM data collected on the Form EIA-861, *Annual Electric Power Industry Report*. EIA will now collect DSM information from all respondents regardless of size, instead of only from those with electricity sales of 150,000 Megawatthours or greater per year. EIA will also collect State and sector-level breakdowns of DSM costs, energy efficiency, and load management effects. The EIA-861 will also be collecting data on net metering by State and sector, including the capacity and technology type for net metering facilities.

Demand response data is also collected by the Federal Energy Regulatory Commission (FERC) as required by the Energy Policy Act of 2005. FERC's current demand response data collection instrument is its Form 731, *Demand Response/Time-Based Rate Programs and Advanced Metering Survey*. This is a voluntary survey, and the response rate in recent years has been just over 60 percent. FERC accounts for missing respondents by imputing based on data EIA collects on the EIA-861 survey. Any future expansion of demand side data that EIA collects will be coordinated with FERC in order to avoid duplication and lessen respondent burden.

Wholesale power data from non-FERC jurisdictional entities. Presently, the *Form EIA-861, Annual Electric Power Industry Report,* collects a limited amount of aggregate data on electricity sales for resale. This level of aggregation does not permit an assessment of wholesale market activities and pricing trends for specific types of power supply services and generation-related services associated with transmission. FERC collects these types of data in its *Electronic Quarterly Report (EQR)*. The 1,327 respondents to the EQR are all FERC-jurisdictional entities. However, municipal utilities, political subdivisions, Federal utilities, and cooperative utilities-- that accounted for nearly 76 percent of total sales for resale in 2006--are not required to submit an EQR. The collection of cooperative and public power data in a format similar or identical to the FERC EQR would fill the gap on bulk power trade and provide a more complete picture of wholesale market activity.

Transmission data. Currently, the Federal Energy Regulatory Commission and the Rural Utilities Service collects transmission-related transaction data (e.g., type of service, revenues, volumes) for a limited number of transmission service providers. In addition to the absence of a census survey of transmission service provid-

ers, the existing data are not systematically compiled by type of service or class of ownership or made publicly available. Collection of data on reliability, finance and investment, and transmission and wholesale power data would allow better evaluation of trends in the price and volumes of transmission service. Such data would be particularly useful to assess the impacts of proposed legislation and regulatory policies that could affect transmission investment and pricing.

Fossil fuel receipts from small electric power plants (under 50 megawatts). The *Form EIA-923, Power Plant Operations Report*, which came into use in 2008, collects fuel consumption for all types of combustible fuels, including both fossil and biomass, from all plants with capacity of 1 Megawatt (MW) or more. Coal and petroleum products stocks (excluding waste coal and waste oil) are also collected from all plants with capacity of 1 MW or more. However, the Form EIA-923 collects receipts of fossil fuels (coal, petroleum products, and natural gas) only from plants 50 MW and above. This includes 1,625 plants and accounts for about 97 percent of fossil fuel receipts. Currently, a statistical estimate is developed for the remainder of fossil fuel receipts (for plants under 50 MW) and is included in the data published in the *Electric Power Monthly* and *Electric Power Annual*. In order to complete the fossil fuel picture for the entire electric power industry, receipts would need to be collected from every facility over 1 MW that uses fossil fuels. This would match the method by which consumption of coal, natural gas, and petroleum products and stocks of coal and petroleum products are currently reported and it would approximately double the number of plants filing the form. Collecting these missing receipts would allow EIA to fully balance the stocks-receipts-consumption equation and ensure that the data are even more accurate. It would also give EIA customers access to the full set of data.

Data on coal preparation volumes, technology and costs. Currently, there is no national inventory of coal waste materials. Coal cleaning or preparation technologies are sometimes used to remove pollutants from coal before it is burned. Such data would provide information on coal preparation technology and capacity, amount of waste products produced, the disposal/use of waste products, and some of the costs incurred in the cleaning process.

Coal reserves estimates. EIA provides tonnage estimates by State of Recoverable Reserves and Demonstrated Reserve Base (DRB) using a reserve database established by the U.S. Bureau of Mines in the early 1900s. EIA has occasionally augmented the reserve database with additional data provided by some States, but the reserve data remain mostly the original U.S. Bureau of Mines data. The National Research Council (NRC), in its 2007 study *Coal Research and Development to Support National Energy Policy*, recommended a coordinated Federal-State-industry initiative to determine the magnitude and characteristics of the Nation's recoverable coal reserves, using modern mapping, coal characterization, and database technologies, with the goal of providing policy makers with a comprehensive accounting of national coal reserves within 10 years.

Up-to-date assessments of coal reserves are needed to support decisions concerning new technologies such as coal-to-liquids and coal-to-gas, as well as the conventional use of coal for electricity generation. EIA frequently is contacted by Federal, State, congressional, and industry groups seeking coal reserve data.

2.2 Petroleum, Biofuels, and Natural Gas

EIA information regarding prices, domestic reserves, production, imports, storage, distribution and consumption of oil, petroleum products, biofuels, natural gas and other associated liquids is collected through a comprehensive program consisting of 10 weekly surveys, 16 monthly surveys, 5 annual surveys, and one quadrennial survey. The number of respondents to particular surveys ranges from 50 to 20,000. Based on changing energy information needs, EIA has made numerous changes in the scope of its oil and natural gas data collections in the recent past, and plans to make more changes in the future.

2.2a Recently Completed Improvements

Monthly petroleum supply changes. Consistent with the changes made to account for fuel ethanol, in 2009 EIA consolidated the reporting of monthly terminal inventory data (EIA-811) with blending activity (Form EIA-815) and required all terminal activity to be reported on a site basis rather than on a region-wide aggregate basis. EIA expanded the product slate for blending operations to accommodate reclassification of products and added a full material balance (products received by the terminal, blended by the terminal, stored at the terminal, and leaving the terminal). Combined reporting of bulk terminal stocks and blending activity at the

site level in the context of a complete material balance allows reporting companies to determine more clearly whether or not bulk terminal reports are complete, consistent, and accurate and improves the ability of EIA staff to assist reporting companies in resolving questions and reporting issues relating to their bulk terminal data. Site-level data is also more useful to State and Federal officials charged with assessing and responding to energy emergencies.

Weekly petroleum supply changes. In 2010, EIA began collecting weekly U.S. and regional ethanol production and inventory data, using the new *EIA-809 Weekly Oxygenate Report,* and collecting weekly ethanol import data on the existing EIA-804 *Weekly Imports Report*. These data are published in EIA's *Weekly Petroleum Status Report (WPSR),* providing a more complete supply and disposition profile for fuel ethanol. Previously, fuel ethanol stocks and imports were estimated from monthly data and grouped in the "Other Oils" category. Incorporation of ethanol by expanding the underlying surveys has allowed EIA to better reflect shifting industry practices, such as blending finished products at terminals rather than at refineries. EIA now uses the weekly fuel ethanol data to calculate finished motor gasoline production, improving the overall quality and accuracy of the gasoline production data and the WPSR motor gasoline product supplied calculation. Beyond the addition of weekly ethanol data, EIA also added collections of asphalt and road oil, kerosene, natural gas plant liquids, and liquefied refinery gases data (in addition to propane/propylene which was previously collected). The additional data, collected on existing weekly surveys, significantly reduce the portion of the "Other Oils" category that EIA must estimate.

Biodiesel data. In 2009, the *EIA-22M Monthly Biodiesel Production Survey* began to collect monthly data on production, prices, and feedstocks, as called for in section 1508 of the Energy Policy Act of 2005.

Storage capacity data. In 2010, EIA began collecting working and shell storage capacity for crude oil and petroleum products semi-annually in March and September 2010 using existing Petroleum Supply Reporting System monthly survey forms. Storage capacity limitations are one factor that can affect the price of crude oil at major market hubs, potentially affecting linkages between physical energy markets and energy-based derivative contracts traded in financial markets. Storage capacity information is collected from refiners (EIA-810), terminal operators (EIA-815), pipeline operators (EIA-812), crude oil storage and pipeline operators (EIA-813), and ethanol producers (EIA-819). EIA began publishing this information in December 2010.

Natural gas production. In 2007, EIA eliminated its previous monthly production survey based on State agency data (EIA-895M *Monthly Quantity and Value of Natural Gas Report*) and moved to a new production survey of well operators (EIA-914). In 2010, the sampling and estimation processes were thoroughly reviewed and improved. EIA created additional quality assurance procedures to improve respondent data submissions. For quality purposes, EIA now publishes monthly comparisons to other data sources in addition to survey results.

Natural gas processing plant information. In 2008, EIA implemented a new form to provide natural gas processing plant information both to generate a baseline and provide operating status updates in emergency situations (EIA-757).gas supply

Shale natural gas production. In 2008, EIA added a shale gas category to its survey of annual production data (EIA-895, *Annual Quantity and Value of Natural Gas Report*) to address a significant shift in domestic gas production patterns.

Annual survey of domestic oil and natural gas reserves. Categories for shale gas and tight fields were added in 2006 to EIA's annual survey of U.S. oil and gas reserves (EIA-23). Separate shale gas reserves were published for the first time in 2009.

Natural gas distributor sendout. In 2010, EIA added total sendout to its monthly survey of natural gas distributors (EIA-857) to better reflect actual deliveries of natural gas on a monthly basis. As discussed in articles posted on the EIA website, the availability of sendout data should significantly improve the quality of EIA's monthly natural gas balance. Incorporation of total sendout in the estimation procedure more closely aligns EIA's published supply and demand totals by reducing the monthly balancing item, which represents the difference between natural and disposition at the national level. Additional steps on this issue may involve further work with industrial and electric power sector-level consumption patterns and coverage.

Use of third-party natural gas data. EIA has also begun the integration of interstate pipeline nominations collected from pipeline websites by a third party (Bentek Energy Consulting) to its short-term data collections.

These data have been incorporated into some *Natural Gas Monthly* and *Natural Gas Annual* processes, and data will be used to monitor gas movements during hurricane seasons.

Natural gas marketers. Expansion of the natural gas marketers survey (EIA-910) to collect annual gas marketer sales to residential and commercial customers is being considered. This annual data would be collected from the same companies that currently report on the monthly Form EIA-910 in 12 States that have active retail choice programs. In addition to getting annual totals for the data currently reported monthly, the annual survey would ask for natural gas sales broken out by each local distribution company whose customers they supply gas. Currently, the monthly EIA-910 collects sales of natural gas at the State level.

The breakout of sales by local distribution company would improve the quality of the marketer data by identifying discrepancies found when directly comparing marketers sold volumes reported on the annual EIA-910 to those volumes delivered on behalf of marketers reported by local distribution companies on Form EIA-176, *Annual Report of Natural and Supplemental Gas Supply and Disposition*. As resources allow, EIA will consider adding more States to the EIA-910 as it is only currently surveying 12 of the 22 States with active residential retail choice programs.

2.2b Ongoing and Potential Future Improvements

Refinery outage information. EIA is developing requirements for a new survey designed to capture planned and unplanned refinery outages and reduced production of gasoline and distillate products resulting from the outages. EIA has decided to collect its own data on this topic, pending available resources, when it determined that commercial refinery outage data were missing too many planned outages until the outage was about to begin or actually underway. The commercial data also did not include estimates of production losses for major products. Section 804 of EISA requires EIA to collect that data if commercially available sources are not adequate.

Collection of crude oil and petroleum product movements. The petroleum industry has changed since EIA's surveys were designed to collect information on the movement of crude oil and petroleum products between Petroleum Administration for Defense Districts (PADDs). Data on the movements of crude oil and petroleum products by pipeline are reported by pipeline operators on the current petroleum supply survey forms, and movements by water are reported for shippers using tankers or barges. However, the current data collection program does not capture the inter-PADD movements by rail or inventories held in rail cars, both of which have increased over the last few years. Rail movements have become increasingly important for products such as propane, butane, fuel ethanol, and crude oil.

Petroleum product export models. EIA plans to revisit the techniques it uses to estimate product exports, since those estimates have sometimes been a source of error in estimating weekly product supplied. For example, our models did not capture the sharp rise in exports of distillate fuels that occurred in recent years, leading EIA to overstate product supplied to the U.S. market until exports were properly measured in international trade data after a lag of many months. EIA will review the structure and estimation methods used. If updated estimation approaches cannot provide significant improvement in accuracy, EIA will consider collecting product export data.

Monthly natural gas and crude oil production. EIA is exploring the addition of more States to its monthly natural gas production survey (EIA-914). The change would require a review of the sampling process and possibly different estimating procedures for some of the new States. EIA is also examining adding oil production to the survey. This would allow for more timely and accurate crude oil production data at a time when production of domestic onshore crude oil resources is changing rapidly with the development of the Bakken shale and the increased application of enhanced oil recovery techniques.

Oil production estimates. In 2010, EIA started a review and update of the Monthly Oil Production Update system and the Crude Oil Production System used in the monthly oil production estimate process. These systems are showing their age and are in need of updating. Improved estimation techniques could reduce the need for the expanded collection of oil production data through the EIA-914.

Annual natural gas production data. EIA will assess its annual gas production data survey (EIA-895, *Annual Quantity and Value of Natural Gas Report*), specifically (1) whether the instrument as designed collects the

most relevant production data; (2) other sources of production and component data for benchmarking and/or publication purposes; and (3) areas requiring special focus, e.g., vented and flared volumes.

Natural gas customer billing data. The EIA-857 monthly survey of natural gas distributors is being examined for further refinement, including more precise guidelines for reporting sector-level revenue data used to derive monthly prices. Currently respondents are given considerable latitude as to how they report these data. More consistent reporting will improve the quality of the derived monthly prices.

Natural gas price data. EIA will assess the quality of its price data, especially the prices at the city gate and to industrial users.

2.3 End-use Consumption

2.3a Recently Completed Improvements

Increased sample size for residential survey. The Residential Energy Consumption Survey (RECS) is in the process of compiling 2009 data on U.S. household and housing unit characteristics that was collected during 2010. Additional funding provided through American Recovery and Reinvestment Act (ARRA) has allowed for a significantly larger RECS sample size — roughly three times the historical base sample size. As a result, EIA will be providing more State-level data (16 States instead of four) and has added questions regarding participation in ARRA-related energy efficiency programs such as the purchase of Energy Star appliances or weatherization
activity.

2.3b Ongoing and Potential Future Improvements

Frequency of end-use surveys. EIA has a set of surveys in place that collect detailed data on end-use energy consumption: the *EIA-457 A-G, Residential Energy Consumption Survey (RECS); EIA-871 A-F, Commercial Buildings Energy Consumption Survey (CBECS);* and the *EIA-846, Manufacturing Energy Consumption Survey (MECS).* These are the only U.S. sector-specific energy consumption surveys that are both statistically reliable and comprehensive, and their data serve as major inputs into EIA's National Energy Modeling System models, which in turn are used to benchmark other government and non-government modeling systems. In addition, these surveys provide a benchmark of adoption of energy efficiency products and technologies. These surveys are required under Section 205 of the DOE Organization Act, which calls for the MECS to be a biennial survey and for the RECS and CBECS to be triennial. They are labor-intensive and thus expensive to conduct. Because of resource constraints, all three surveys are currently on a quadrennial schedule. The FY 2012 budget request included funding initiatives to increase the sample size of CBECS and to increase the frequency of MECS to every two years. The increased CBECS sample size would allow for more detailed characterization of building types.

Review of end-use surveys. At EIA's request, the Committee on National Statistics (CNSTAT) of the National Research Council convened a panel in September 2009 to conduct a comprehensive 30-month study of the CBECS and RECS. The panel's charge is to consider possible improvements to data quality, geographic coverage, relevance, and the timeliness of data releases.

Because the next CBECS survey is scheduled to be fielded in 2011, the panel did deliver some preliminary recommendations for our next survey cycle (see Appendix C, *Residential Energy Consumption Letter Report*). Recommendations under consideration for implementation include the following:

- Conduct a small-scale test in which similar data for the same building would be collected by energy auditors as well as interviewers to evaluate differences between the two approaches.
- Attach GPS coordinates to each sampled building to facilitate augmentation of the file with additional data sources such as detailed weather information.
- Evaluate how digital photographs of buildings and equipment in buildings, subject to relevant privacy concerns, can be useful for field operations and for later editing.

At the conclusion of the study, the panel will deliver its comprehensive report on the overall design and conduct of both the CBECS and RECS.

3. Improvements to Data Collection Processes

One of the largest challenges EIA faces with regard to the quality of its data is its processes and systems. EIA manages a patchwork of programs and databases in various programming languages and operating environments that has grown up over decades of inadequate investment in systems. As a consequence, operational expenses have become a much greater share of expenditures over the years, perpetuating a cycle of underinvestment. These systems have become increasingly expensive to maintain; cumbersome to change; and, without targeted, strategic investment, nearly impossible to update. Over the past few years, EIA has undertaken several initiatives to provide control of the processes and systems underlying its data management.

3.1 Recently Completed Improvements

Natural gas survey system and process upgrades. Recognizing the need to use common software to maximize data processing efficiencies, EIA's natural gas data program has made significant improvements to its underlying data management systems. For example, all network-based programs now utilize the same database software (Oracle), which streamlines maintenance, aids information flow between systems, and improves archival of historical data. Additionally, the processing system for the *EIA-912, Weekly Underground Natural Gas Report*, (designated a Principal Federal Economic Indicator) was recently upgraded to enhance functionality and reliability. To improve production efficiencies, software was also developed to track and display the status of data feeds and quality control reviews for the Natural Gas Monthly. Plans are underway to expand its usage to other publications, including the Natural Gas Annual.

Natural gas survey frames and sampling. Early in 2010, EIA completed a comprehensive review of the *Annual Report of Natural and Supplemental Gas Supply and* Disposition survey frame (for the EIA-176), which touches on most aspects of the natural gas industry and forms the sampling basis for the monthly consumption data (EIA-857). The frames review identified about 40 additional companies. EIA developed a repeatable protocol for subsequent frame reviews.

Petroleum frames and sampling. In 2010, EIA developed a web-based frames management system intended to house relevant company-level information (e.g., contact data, corporate structure/ interrelationships, and history) along with an automated identification number assignment and approval process for all petroleum marketing and petroleum supply surveys and for the survey on natural gas liquids proved reserves and production, the EIA-64A*, Annual Report of the Origin of Natural Gas Liquids Production.* The system provides the ability to compare respondent information across surveys and provides the foundation for specific sampling frames

Exit from the mainframe. In 2010, the last three survey systems residing on EIA's mainframe were removed, permitting decommissioning of the mainframe. This has reduced EIA's dependence on increasingly obsolete software, and will result in savings in licensing fees, electricity consumption, and use of space.

Manufacturing Energy Consumption Survey (MECS). The MECS had collected data by mailed forms through 2002, but the 2006 MECS used the U.S. Census Bureau's electronic data collection system. This system was used by 85 percent of the respondents in 2006, and will be utilized again in the 2010 survey. Features of the MECS electronic questionnaire included these:

- All respondents are eligible to use it, but will have a paper form available as a back-up for those unable or reluctant to use an Internet-based system;
- Data were encrypted and submitted electronically through a secure Internet connection, thus maintaining data integrity and eliminating the potential for keying errors;
- Industry classification (i.e., NAICS code) was used as a screener to customize the form for the specific respondent. That screening allowed the respondent to focus only on the energy sources most relevant to that industry.
- Check boxes or other means were used to give the respondent an opportunity to report in his or her most convenient units. Edit price checks were developed for each of the unit choices.

- The questionnaire was developed and formatted with screen input in mind, using the best available practices

EIA believes that the above features will lessen the response burden and improve data quality.

3.2 Ongoing and Potential Future Improvements

Weekly petroleum supply data collection. In 2009, EIA began collecting monthly petroleum supply data at the facility rather than the company level, allowing for more granular monitoring of data quality. Unfortunately, the current weekly petroleum data systems and processes cannot support the increased number of responses for facilities-based surveys. The sample size for weekly reporting would increase by approximately 500 additional responses, going from 1,140 responses to over 1,600 responses. EIA is examining alternative sampling methodologies that would reduce respondent burden by reducing the number of reports that must be collected in order to shift the *Weekly Petroleum Status Report* (WPSR) to a facilities-based collection. EIA is also assessing the WPSR process and system in detail to identify opportunities to improve and streamline it so that efficient and effective data quality can be accomplished within the weekly two-day processing schedule, including the possibility of rewriting the system in its entirety, and is examining options for use of a common internet data collection system to increase the speed of, and reduce errors from, the current, manually-oriented WPSR data collection process.

Natural gas frames and sampling. EIA will develop a frames management system to house all relevant company-level information (e.g., contact data, corporate structure/interrelationships) along with an automated ID assignment. There is currently no frames management repository. Instead, frame-related information resides within the various processing systems. New ID assignments currently are made via a labor-intensive, non-automated process.

Annual survey of domestic oil and natural gas reserves. EIA plans to review and modify the sampling procedure for its annual survey of U.S. oil and gas reserves (EIA-23) with the intent to reduce the burden on smaller respondents and improve coverage.

Natural gas storage capacity. EIA will explore the combination of its monthly and annual storage capacity surveys (EIA-191) into a single form, which could streamline processing, allow for more timely updates regarding storage capacity, and reduce respondent burden.

Internet survey management system. One of EIA's goals to improve data collection processes is to create an EIA-wide data collection tool. Currently, each program office conducts surveys using various collection methods such as manual, telephone, interview, email, secure file transfer, and internet data collection. For all appropriate web-based data collections, EIA is in the process of developing a common platform—the Integrated Survey Management System (ISMS)--for Internet data collection agency-wide. ISMS is designed as both an authoring tool (a program used to create or maintain Internet survey forms) and data collection tool, allowing the survey managers and staff to design the form and other processes at a lower cost than before. By standardizing across offices, EIA expects to lower development and maintenance costs for all of its major surveys, as well as improve survey quality at both the operational and management levels.

Consolidation of all data collection activities within a realigned organization. In October 2010, EIA changed its organizational structure to align with its four primary functional areas: statistics, analysis, communication, and management. The organizational change creates a more effective platform for meeting EIA's mission requirements; strengthens the information collection and processing, analysis, and dissemination functions; improves customer service; and will achieve more efficient operations. Under the new organizational structure, all data collection activities are under the leadership of an Assistant Administrator for Energy Statistics.

Improvements in resource management. Under the new organizational structure, all resource management activities are under the leadership of an Assistant Administrator for Resource and Technology Management. EIA will implement processes to improve the efficiency and the timeliness of EIA's human resource, contracting, and finance/budget efforts. EIA will continue to reduce skill gaps in mission-critical occupations by replacing vacated industry specialist positions with "core-series" professional specialists (industry economists, operations research analysts, mathematical statisticians, survey statisticians, and general engineers). The hiring process for entry-level and journeyman-level positions will be further streamlined to keep pace with attrition.

In addition, EIA will continue to upgrade the technical expertise of its contractor staff, and continue to pursue formal training/certification of its project managers, information technology specialists, and financial and human resources analysts. Where possible, EIA will consolidate requirements into less costly contract vehicles employing the most cost-efficient pricing formulas.

4. Improvement to Accessibility and Timeliness of State and Other Data

EIA is continually working toward increasing the accessibility, relevance, and understandability of our energy information. Under the new organizational structure, all communications activities are under the leadership of an Assistant Administrator for Communications. Our ongoing efforts include:

- A complete redesign of our website to make our data and analyses easier to find and use
- User assessment tools including usability testing, a "star-rating" system for web pages, and customer surveys to collect feedback
- Active participation in the Administration's Data.gov initiative to provide open access to datasets
- The energy education and literacy initiative that includes a number of products written in plain language, including Energy Explained, Energy Kids, and Energy In Brief

Website redesign project. In February 2011, EIA launched the first phase of its redesigned website featuring:

- Today in Energy, a new energy education product published every weekday highlighting current issues, topics, and data trends in short articles written in plain language
- New homepage and improved navigation to make it easier for customers to find the abundance of EIA information and data and better showcase the breadth and depth of EIA content
- Updated logo and dramatic new visual identity to help re-introduce EIA, its initiatives, and its programs

This is the first major redesign of EIA.gov in six years and the third since its inception in 1995. We know from our customer survey that more than 90 percent of users are satisfied with the quality of the information on our site, and that 80 percent said that they found what they were looking for either easily or with some effort. But we also know that with a site as large and complex as ours, people still have trouble finding what they need and completing their tasks. The new EIA.gov will improve the agency's capacity to meet the evolving needs of our diverse customers, enhance access to EIA information and data, present a unified and consistent image to both external customers and EIA employees, and better enable the agency to speak with one voice.

User assessment tools. EIA has employed a variety of user assessment tools to improve the accessibility and understandability of our energy information. Card sorting, a technique to determine appropriate categories and names for information, was used to create the navigation for the Energy Explained section of our website. EIA has been conducting usability testing on our website for many years to watch actual users try to find information. The website redesign project includes several rounds of usability testing, both in-person and remote testing. The annual web customer survey collects a large amount of feedback from customers about whether they are satisfied and can find what they are looking for, as well as providing suggestions for improvements. Our plain language/energy literacy project to provide more content for citizens in non-technical language was a direct result of feedback we heard from users on the survey. Additional customer input is gained through the use of a "star-rating" system that allows visitors

Data.gov participation. The Data.gov initiative increases the ability of the public to easily find, download, and use data sets that are generated and held by the Federal government. EIA has participated in this effort since the very beginning, and was one of the first agencies to have data posted on the site.

As of March 2011, EIA has over 70 datasets available on Data.gov. EIA's data have always been available to the public on our website, but with Data.gov, a wide variety of Federal datasets are all in one place so users can find them more easily and use them for building applications, conducting analyses, and performing research.

Energy education and literacy. In 2009, EIA launched Energy Explained and a redesign of its popular Energy Kids website. Energy Explained offers an encyclopedia of energy topics. The site explains where gasoline comes from, what determines the price of electricity, how much renewable energy the United States uses, and hundreds of other energy topics. It also features a user "star rating" and commenting system to facilitate audience feedback and engagement and builds in viral marketing techniques to further disseminate this product.

Energy Kids features more than 100 pages of engaging educational content for kids, parents, teachers, and other customers who need content written in plain language. Kids can learn about energy and challenge their brains

with energy Sudoku, crossword puzzles, and word searches. The site also features energy-related stories, hands-on activities, and research articles for the classroom. It has used audience polling to determine final designs for promotional materials for kids and "crowdsourcing" to solicit lesson plans from teachers and make them available on the site.

EIA's latest energy education offering is Today in Energy, launched in February 2011 on the EIA.gov homepage. Each business day, the new plain language feature focuses on timely energy topics and issues by highlighting intriguing energy data that the public might otherwise have missed. Creating this kind of topical and timely content on a daily basis is a significant improvement in the delivery of EIA's information to a broader audience. Today in Energy also allows visitors to actively engage with energy information and experts. Charts on the site are now more interactive, and enable users to modify and drill down when viewing them. Site visitors are also able to suggest topics to the Today in Energy team for future articles.

EIA also publishes the Energy in Brief series, which presents a number of short, printable articles on relevant energy issues. The articles explain important energy topics in concise, plain language and provide a valuable tool to help the public better understand these issues.

These education initiatives strive to engage a broader audience online and advance the public's energy literacy. In May 2010, the National Association of Government Communicators (NAGC) honored Energy Kids as the first recipient of its "Best in Show" award and as the first place recipient of awards for "Best Website" and "Shoestring Budget." In April 2010, Energy Explained earned a "ClearMark" Award of Excellence in the category of "Best Website" from the Center for Plain Language. An Energy Education team within EIA manages and updates the content, adding innovative features, and developing new products.

State data needs assessment. Section 805(d) of EISA required EIA to assess State-level data needs and submit to Congress a plan to address those needs. EIA completed this assessment in January 2009 (see Appendix D *State Energy Data Needs Assessment*.) That assessment outlines the consultations EIA undertook to determine the needs of the users of our State-level data and includes over 30 possible initiatives to close data gaps and address concerns.

Since that assessment was issued, EIA has worked to implement several of those initiatives. For instance, following discussions with staff of the Committee on Natural Resources of the U.S. House of Representatives, we added profiles for U.S. Territories to our State Energy Profile webpage in June 2009. EIA has also implemented some of the suggested mapping enhancements, including "mouse-over" label capability. Although we are not ready to implement user-generated mapping, we are exploring options for allowing more control, such as allowing user-specified layers of data in graphs.

In addition, EIA improved the timeliness of State-level data, releasing the final 2008 data in our State Energy Data System in 18 months, down from a lag of 23 months a few years previously. We are also developing a set of advance estimates of State consumption to be released nine months after the end of the reference year.

The Residential Energy Consumption Survey (RECS), as noted previously, collected 2009 data on U.S. household and housing unit characteristics, and is in the process preparing it for release. Additional funding provided through American Recovery and Reinvestment Act (ARRA) has allowed for a significantly larger RECS sample size — roughly three times the historical base sample size. As a result, EIA will be providing more State-level data (16 States instead of four).

EIA is committed to working closely with State governments in order to meet the evolving needs for more detailed energy data, such as expanded data on renewable energy. Having more types of EIA data available at the State level will be increasingly important as more States consider enactment of policies to boost energy efficiency or shift their electricity supply mix. EIA will continue to explore options for expanding our data sharing agreements in order to maximize data utility while minimizing costs and respondent burdens.

5. Improvements to Data Analysis

Under the new organizational structure, an Assistant Administrator for Energy Analysis will lead our short-term, long-term, and other energy analysis activities.

Energy and Financial Markets Initiative. Launched in September 2009, EIA's Energy and Financial Markets Initiative aims to improve energy market transparency, support sound policy and efficient markets, and increase public understanding — activities that are central to EIA's mission. EIA's traditional coverage of physical fundamentals such as energy consumption, production, inventories, spare production capacity, and geopolitical risks continues to be essential, but moving forward, EIA is also assessing other influences, such as speculation, hedging, investment, and exchange rates, as it seeks to fully understand energy price movements.

The Markets Initiative focuses on four main areas: (1) collection of critical information on factors affecting energy prices, (2) analysis through in-depth studies of energy market behavior, (3) outreach to solicit feedback from a broad range of experts on the interrelationship of energy and financial markets, and (4) coordination with other Federal and international agencies engaged in energy market information collection and analysis.

EIA's standard collection cycles of annual, monthly and some weekly data mean that much of the information EIA gathers comes after any possible market response to that information. Over time, with improving information technology, third-party vendors have developed valuable short-term information about energy market factors, based on publicly available scheduling information, regulatory posting requirements, and sources of business intelligence. As energy markets have developed over time, often with the help of improved information technology capabilities, access to more granular information has increased significantly, as has the possibility of a more detailed understanding of how market participants form and use prices. Examples of more complete data coverage include pipeline shipments of natural gas, bulk movement of electricity, and shipments of liquefied natural gas and oil in international waters. Contracting for this vendor-developed information and developing the associated reporting capabilities would allow EIA to develop the short-term view of energy market drivers needed for comprehensive market analysis. Finally, the information would enhance quality assurance efforts dealing with EIA's survey data by providing an alternate source of information to check for misreporting.

National Energy Modeling System. EIA is continuing its efforts to upgrade the aging National Energy Modeling System (NEMS). A new National Energy Model (NEM) will improve the ability to assess and forecast supply, demand, and technology trends affecting U.S. and world energy markets. EIA will improve representations of regional transportation energy markets and vehicle efficiency standards, as well as fuel supply and conversion, to meet the needs of the Administration, the Congress, and other customers for relevant, reliable, timely, and accessible energy projections and analyses

6. Conclusion

EIA is committed to improving the coverage, quality, and integration of energy demand, supply, and market data to meet the changing information needs of policymakers, the broader energy industry and its associated markets.

All aspects of developing and disseminating data are targets for improvement, from collecting and processing data, to analyzing the data once acquired, to providing information to the public. Within the context of its budgetary constraints, EIA is working to fill critical data gaps; update the statistical techniques used in data collection; better protect the integrity of data collected, processed and published; assure documentation of data processing decisions; and reduce lifecycle development and operating costs for EIA's statistical programs.

EIA is continually working toward increasing the accessibility, relevance, and understandability of our energy information. EIA's ongoing efforts include a redesign of our website to make our data and analyses easier to find and use. EIA is committed to working closely with State governments in order to meet the evolving needs for more detailed and timely energy data.

Efforts to strengthen the relevance and breadth of EIA's energy analysis include EIA's Energy and Financial Markets Initiative which aims to improve energy market transparency, support sound policy and efficient markets, and increase public understanding — activities that are central to EIA's mission. Analysis will also be improved through a multiyear project underway to replace the components of the National Energy Modeling System to improve EIA's ability to assess and project supply, demand, and technology trends affecting U.S. and world energy markets.

All of these efforts will take place within an organization that is now aligned with its four primary functional areas: statistics, analysis, communication, and management. The organizational change creates a more effective platform for meeting EIA's mission requirements; strengthens the information collection and processing, analysis, and dissemination functions; improves customer service; and will achieve more efficient operations.

Appendix A. Energy Independence and Security Act of 2007, Section 805

Energy Independence and Security Act of 2007, Public Law 110-140

SEC. 805. ASSESSMENT OF RESOURCES.

(a) 5-YEAR PLAN.—

(1) ESTABLISHMENT.—The Administrator of the Energy Information Administration (referred to in this section as the ''Administrator'') shall establish a 5-year plan to enhance the quality and scope of the data collection necessary to ensure the scope, accuracy, and timeliness of the information needed for efficient functioning of energy markets and related financial operations.

(2) REQUIREMENT.—In establishing the plan under paragraph (1), the Administrator shall pay particular attention to—

(A) data series terminated because of budget constraints;

(B) data on demand response;

(C) timely data series of State-level information;

(D) improvements in the area of oil and gas data;

(E) improvements in data on solid byproducts from coal-based energy-producing facilities; and

(F) the ability to meet applicable deadlines under Federal law (including regulations) to provide data required by Congress.

(b) SUBMISSION TO CONGRESS.—The Administrator shall submit to Congress the plan established under subsection (a), including a description of any improvements needed to enhance the ability of the Administrator to collect and process energy information in a manner consistent with the needs of energy markets.

(c) GUIDELINES.—

(1) IN GENERAL.—The Administrator shall—

(A) establish guidelines to ensure the quality, comparability, and scope of State energy data, including data on energy production and consumption by product and

sector and renewable and alternative sources, required to provide a comprehensive, accurate energy profile at the State level;

(B) share company-level data collected at the State level with each State involved, in a manner consistent with the legal authorities, confidentiality protections, and

stated uses in effect at the time the data were collected, subject to the condition that the State shall agree to reasonable requirements for use of the data, as the Administrator may require;

(C) assess any existing gaps in data obtained and compiled by the Energy Information Administration; and H. R. 6—231

(D) evaluate the most cost-effective ways to address any data quality and quantity issues in conjunction with State officials.

(2) CONSULTATION.—The Administrator shall consult with State officials and the Federal Energy Regulatory Commission on a regular basis in—

(A) establishing guidelines and determining the scope of State-level data under paragraph (1); and (B) exploring ways to address data needs and serve data uses.

(d) ASSESSMENT OF STATE DATA NEEDS.—Not later than 1 year after the date of enactment of this Act, the Administrator shall submit to Congress an assessment of State-level data needs, including a plan to address the needs.

(e) AUTHORIZATION OF APPROPRIATIONS.—In addition to any other amounts made available to the Administrator, there are authorized to be appropriated to the Administrator to carry out this section—

(1) $10,000,000 for fiscal year 2008;

(2) $10,000,000 for fiscal year 2009;

(3) $10,000,000 for fiscal year 2010;

(4) $15,000,000 for fiscal year 2011;

(5) $20,000,000 for fiscal year 2012; and

(6) such sums as are necessary for subsequent fiscal years

Appendix B. Legal Sources for EIA's Mandatory Data Collection Authority

There are two main laws that the Energy Information Administration relies upon for mandatory data collection authority for its surveys.

The first law is the **Federal Energy Administration (FEA) Act of 1974**, Public Law 93-275, as amended. The FEA Act created the first U.S. agency with the primary focus on energy and mandated it to collect, assemble, evaluate, and analyze energy information. It also provided FEA with data collection enforcement authority for gathering data from energy producing and major consuming firms. 15 U.S.C. §790a of the FEA Act mandated establishment of the National Energy Information System to

> ...contain such energy information as is necessary to carry out the Administration's statistical and forecasting activities...as is required to define and permit analysis of:
>
> - the institutional structure of the energy supply system including patterns of ownership and control of mineral fuel and non-mineral energy resources and the production, distribution, and marketing of mineral fuels and electricity;
> - the consumption of mineral fuels, non-mineral energy resources, and electricity by such classes, sectors, and regions as may be appropriate for the purposes of this Act;
> - the sensitivity of energy resource reserves, exploration, development, production, transportation, and consumption to economic factors, environmental constraints, technological improvements, and substitutability of alternate energy sources;
> - the comparability of energy information and statistics that are supplied by different sources;
> - industrial, labor, and regional impacts of changes in patterns of energy supply and consumption;
> - international aspects, economic and otherwise, of the evolving energy situation; and
> - long-term relationships between energy supply and consumption in the United States and world communities.

The **Department of Energy (DOE) Organization Act of 1977**, Public Law 95-91, created the Department of Energy. 42 U.S.C. 7135 of this law established the Energy Information Administration (EIA) to carry out a

> . . . central, comprehensive, and unified energy data and information program which will collect, evaluate, assemble, analyze, and disseminate data and information which is relevant to energy resource reserves, energy production, demand, and technology, and related economic and statistical information, or which is relevant to the adequacy of energy resources to meet demands in the near and longer term future for the Nation's economic and social needs.

The majority of EIA surveys are based on the general mandates in the two laws set forth above. However, there are some surveys specifically mandated by law. Below is a listing of current surveys and corresponding collection authority:

- EIA-1, **Weekly Coal Monitoring Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-3, **Quarterly Coal Consumption and Quality Report**, Manufacturing and Transformation/Processing Coal Plants and Commercial and Institutional Coal Users - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-5, **Quarterly Coal Consumption and Quality Report, Coke Plants** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-7A, **Coal Production and Preparation Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(b), 5(a), 52
- EIA-8A, **Coal Stocks Report – Annual** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-14, **Refiners' Monthly Cost Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52; Public Law 94-163 (Energy Policy and Conservation Act), Sec. 507

- EIA- 22M and EIA-22S, **Monthly Biodiesel Production Survey** - Section 503 of the Energy Policy Act of 1992 (EPACT1992) requires DOE to estimate consumption of alternative and replacement fuels annually and Section 1508 of the Energy Policy Act of 2005 (EPACT2005) directs EIA to conduct a survey of "renewable fuels demand in the motor fuels market," which includes biodiesel. Section 1508 specifies that EIA shall collect monthly data, on a national and regional basis for fuel production, quantity of fuel blended, fuel imports, fuel consumption, fuel market prices, and any other data the EIA determines are necessary to achieve the purpose of this section
- EIA-23L and EIA-23S, **Annual Survey of Domestic Oil and Gas Reserves** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 15, 5(a), 5(b), 52
- EIA-28, **Financial Reporting System** - Section 205(h) of the DOE Organization Act 42 U.S.C. 7135(h).
- EIA-63A and EIA -63B, **Annual Solar Thermal Collector Manufacturers Survey** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(b), 5(a), 52
- EIA-176, **Annual Report of Natural and Supplemental Gas Supply and Disposition** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-182, **Domestic Crude Oil First Purchase Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52; Public Law 94-163 (Energy Policy and Conservation Act), Sec. 507
- EIA-191A, **Annual Underground Gas Storage Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-191M, **Monthly Underground Gas Storage Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-411, **Coordinated Bulk Power Supply Program Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-457A-G, **Residential Energy Consumption Survey** - Section 205(k) of the DOE Organization Act (the Act calls for a triennial survey; however, this survey is done quadrennially due to resource constraints).
- EIA-757, **Natural Gas Processing Plant Survey** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b)
- EIA-782A, **Refiners'/Gas Plant Operators' Monthly Petroleum Product Sales Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52; Public Law 94-163 (Energy Policy and Conservation Act), Sec. 507; Public Law 102-486 (Energy Policy Act of 1992), Sec. 407(a)(3)
- EIA-782B, **Resellers'/Retailers' Monthly Petroleum Product Sales Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52; Public Law 94-163 (Energy Policy and Conservation Act), Sec. 507; Public Law 102-486 (Energy Policy Act of 1992), Sec. 407(a)(3)
- EIA-782C, **Monthly Report of Prime Supplier Sales of Petroleum Products Sold for Local Consumption** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52; Public Law 94-163 (Energy Policy and Conservation Act), Sec. 507
- EIA-800, **Weekly Refinery and Fractionator Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-801, **Weekly Bulk Terminal Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-802, **Weekly Product Pipeline Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-803, **Weekly Crude Oil Stocks Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52

- EIA-804, **Weekly Imports Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-805, **Weekly Terminal Blenders Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-809, **Weekly Oxygenate Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-810, **Monthly Refinery Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-812, **Monthly Product Pipeline Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-813, **Monthly Crude Oil Report** -Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-814, **Monthly Imports Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-815, **Monthly Terminal Blenders Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-816, **Monthly Natural Gas Liquids Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-817, **Monthly Tanker and Barge Movement Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-819, **Monthly Oxygenate Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-820, **Annual Refinery Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-821, **Annual Fuel Oil and Kerosene Sales Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52; Public Law 94-163 (Energy Policy and Conservation Act), Sec. 507
- EIA-826, **Monthly Electric Utility Sales and Revenue Report with State Distributions** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(b), 5(a), 52
- EIA-846A-C, **Manufacturing Energy Consumption Survey** - Section 205(i) of the DOE Organization Act 42 U.S.C. 7135(i) (the Act calls for a biennial survey; however, this survey is done quadrennially due to resource constraints).
- EIA-851A, **Domestic Uranium Production Report (Annual)** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-851Q, **Domestic Uranium Production Report (Quarterly)** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-856, **Monthly Foreign Crude Oil Acquisition Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52; Public Law 94-163 (Energy Policy and Conservation Act), Sec. 252, 253, 254 and 507
- EIA-857, **Monthly Report of Natural Gas Purchases and Deliveries to Consumers** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-858, **Uranium Marketing Annual Survey** - Section 1015 of the Energy Policy Act of 1992.
- EIA-860, **Annual Electric Generator Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52

- EIA-860M, **Monthly Update to the Annual Electric Generator Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-861, **Annual Electric Power Industry Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-863, **Petroleum Product Sales Identification Survey** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52; Public Law 94-163 (Energy Policy and Conservation Act), Sec. 507
- EIA-871A-F, **Commercial Buildings Energy Consumption Survey**- Section 205(k) of the DOE Organization Act (the Act calls for a triennial survey; however, this survey is done quadrennially due to resource constraints).
- EIA-877, **Winter Heating Fuels Telephone Survey** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(b), 5(a), 52; Public Law 94-163 (Energy Policy and Conservation Act), Sec. 507
- EIA-878, **Motor Gasoline Price Survey** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(b), 5(a), 52; Public Law 94-163 (Energy Policy and Conservation Act), Sec. 507
- EIA-882T, **Generic Clearance for Questionnaire Testing, Evaluation, and Research** - Public Law 93-275 (Federal Energy Administration Act of 1974), 13(a), 5(b), 5(a), 52
- EIA-886, **Annual Survey of Alternative Fueled Vehicle Suppliers and Users** - Section 503(b) of the Energy Policy Act of 1992.
- DOE-887, **DOE Customer Surveys** - Executive Order 12862
- EIA-888, **On-Highway Diesel Fuel Price Survey** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52; Public Law 94-163 (Energy Policy and Conservation Act), Sec. 507
- EIA-895, **Annual Quantity and Value of Natural Gas Production Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), 13(a), 5(b), 5(a), 52
- EIA-902, **Annual Geothermal Heat Pump Manufacturers Survey** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-910, **Monthly Natural Gas Marketers Survey** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-912, **Weekly Underground Natural Gas Storage Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-914, **Monthly Natural Gas Production Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-923, **Power Plant Operations Report** - Public Law 93-275 (Federal Energy Administration Act of 1974), Sec. 13(b), 5(a), 5(b), 52
- EIA-1605 and 1605EZ, **Voluntary Reporting of Greenhouse Gases** - Section 1605(b) of the Energy Policy Act of 1992.

Appendix C. Residential Energy Consumption Letter Report

THE NATIONAL ACADEMIES
Advisers to the Nation on Science, Engineering, and Medicine
Committee on National Statistics
Division of Behavioral and Social Sciences and Education

The Keck Center
500 Fifth Street, NW
Washington, DC 20001
Phone: 202 334 3096
Fax: 202 334 3751
www.national-academies.org/cnstat

May 18, 2010

Dr. Richard Newell
Administrator
U.S. Energy Information Administration
1000 Independence Avenue, SW
Washington, DC 20585

Dear Dr. Newell:

At the request of the Energy Information Administration (EIA), the Committee on National Statistics (CNSTAT) of the National Research Council convened a panel to conduct a comprehensive 30-month study of the Commercial Buildings Energy Consumption Survey (CBECS) and Residential Energy Consumption Survey (RECS). Many of the design and operational procedures for the CBECS and RECS were developed in the 1970s and 1980s, and resource limitations during much of the time since then have prevented EIA from making significant changes to the survey methodology or operations. With the possibility of additional funding available in the next few years, EIA asked the National Research Council to conduct a comprehensive review to assess how the CBECS and RECS can take advantage of recent developments in survey methods and to ensure the relevance of the data for meeting increased user needs in the next decade and beyond. The panel's charge is to consider possible improvements to data quality, geographic coverage, relevance, and the timeliness of data releases.

Because plans for the upcoming 2011 round of CBECS must be finalized in the near future, the panel was charged to comment as soon as possible on design and data collection options that would enable the upcoming round of this survey to better support U.S. Department of Energy program information needs, reduce respondent burden, and increase the quality and timeliness of the data. This letter responds to that request, and is limited in scope to discussing issues that the panel believes are realistic to consider in the timeframe leading up to the 2011 data collection. At the conclusion of the study, the panel will deliver its comprehensive report on the overall design and conduct of both CBECS and RECS.

At the first meeting of the panel on February 1-2, 2010, EIA staff discussed preparations for the 2011 CBECS and indicated that work will begin on the 2011 CBECS sample design in the summer of 2010. Thus, any changes to this round of the data collection would have to be evaluated before then. EIA staff also informed the panel that the 2011 CBECS is anticipated to have more funding than it has had in the past. The panel also learned in those discussions that EIA has relatively little empirical data on how well the current design and procedures are

working in comparison with approaches that have been tried in the past and that EIA has not conducted an analysis of options considered but not pursued. Based on the factors described above, the panel's overarching recommendation is to invest some of the currently available additional funding in research that will enable EIA to make future decisions based on empirical evidence about what is most likely to improve geographic coverage, data quality and relevance, while controlling costs. The panel's specific recommendations for research as part of the 2011 CBECS are described below.

BACKGROUND ON THE CBECS

The CBECS is a survey of commercial buildings in the United States, mandated by Congress to provide comprehensive information about energy use in commercial buildings. In addition to energy consumption and expenditure data, the survey collects information about building characteristics, such as energy source, physical structure, equipment used, and activities performed, which provides researchers with detailed information about commercial sector energy use and how it relates to building characteristics. The CBECS is the only national source of these data, and is used for energy forecasting, program development, and policy development.

The survey collects information from a sample of commercial buildings in the United States, and it is administered in two data-collection stages: a Building Characteristics Survey and an Energy Suppliers Survey. During the first stage of the data collection, interviewers visit the buildings selected into the sample and ask a representative of the building, such as the building's owner, manager, or other knowledgeable person to complete the survey. During the second stage of the data collection, the energy suppliers of buildings that were not able to provide adequate information in the first stage are contacted to obtain actual usage and expenditure data from the supplier's records.

SAMPLING FRAME

There is little comprehensive information about the stock of commercial buildings in the United States, and EIA indicated that the lack of a comprehensive national list of commercial buildings or another suitable source from which to select a sample of buildings to interview is one of the major challenges for the CBECS data collection. Because no complete list of buildings is available to use as a sampling frame, EIA builds a new area probability sampling frame for the CBECS on a decennial basis. The frame is based on field listings of commercial buildings within specified geographic areas. This sampling frame is updated between each data collection. However, field listings are resource intensive and relying on sources that are not comprehensive for updating the sampling frame leads to coverage problems.

The CBECS sample design has undergone numerous revisions over the years, as EIA has attempted to address the cost and coverage issues, but most rounds of the CBECS have relied on a combination of an area frame and a list frame, based on existing lists of commercial buildings from a variety of sources and added at the second stage of the area frame sample. The primary sampling units have been counties or groups of counties, within which smaller geographic areas were randomly selected. All commercial buildings were listed and stratified within these smaller areas, and then a sample of buildings was randomly selected from each stratum. This approach

was supplemented with information from existing building lists from other sources to ensure adequate representation of buildings that were of special interest because of their size or type of primary activity.

For the 2007 administration of the CBECS, the 2003 sampling frame had to be updated. At the recommendation of the data collection contractor, the National Opinion Research Center (NORC), the update was based on a U.S. Postal Service (USPS) Delivery Sequence File (DSF) purchased from a vendor licensed by USPS. The DSF is USPS's list of all delivery points in the United States. Using the DSF for updating meant that this list had to be matched to the addresses in the second-stage area frame and the duplicates removed. NORC reported that the unduplication turned out to be a major challenge, in part because of imprecise address records.

As EIA is aware, another major redesign of the CBECS sampling frame could be very productive, but due to the limited time and resources available, this is neither feasible nor recommended for the 2011 data collection. However, leading up to and during the 2011 CBECS, alternative approaches to building a good second-stage sampling frame should be the focus of EIA research, particularly the availability of administrative records and lists. As the EIA staff indicated, and the panel concurs, a sampling frame based on administrative records may have to completely or partially replace the second-stage area frame in the future because of the high costs associated with field listing. Although EIA has considered the use of more lists throughout the years, research on this should continue because the availability of sources of data is constantly evolving, particularly with more information becoming available on the Internet.

For the 2011 CBECS, the most practical approach is to perform another round of updating of the sampling frame using the DSF. Even though unduplication proved to be a challenge when the DSF was first used in 2007, presumably the bulk of the work has now been done, and the 2011 frame can be updated by simply matching the new addresses to the address files used in 2007. We assume that a 2003-2007 cohort of listings is available for use in the 2011 sample based on the matching and updating performed in preparation for the 2007 data collection. As was done in previous years, this approach would have to be supplemented with lists from other sources to assure adequate representation of buildings of special interest.

As an example of such a supplemental source, we recommend exploring the usefulness of local government databases that are available online, such as county property tax databases, some of which include information on square footage and heat source. Two available online databases of which we are aware are those of the Allegheny County Office of Property Assessments in Pennsylvania and of the King County Government in Washington. Although such databases are not universally available online, and their use would undoubtedly present some standardization challenges, their usefulness should be evaluated for two purposes: as a source for a sampling frame and for the possible use of some of the data that are now collected through interviews. Other possible data sources are discussed in the next section, although all require further research to evaluate them.

In rural areas, the DSF often includes only rural route or post office box numbers and so tends not to be very useful. EIA should evaluate information available from vendors who specialize in providing address data to fill these types of gaps. If these sources are found to be inadequate,

field listing may still be necessary. Alternatively, half-open interval updating could be considered, if relisting is deemed too inefficient because of the scattered nature of rural areas. This technique involves updating only new or missed units within a small geographic area (an "interval" associated with an address in the sample). In areas where buildings are scattered in unusual ways, half-open interval updating may be difficult to carry out accurately, but the accuracy of the approach in this particular context could be evaluated as part of the 2011 CBECS. For example, in addition to performing half-open interval updating in the rural sample segments, relisting could also be carried out in a subset of these segments to compare the outcome of the two techniques in terms of the number of listings identified and the number that would end up being added to the 2011 CBECS. The relisting could be performed by experienced listers or supervisors to minimize the costs associated with training for these types of assignments.

Further research is needed to understand the quality and future potential of the DSF. In addition to evaluating the performance of the DSF in comparison with other sources for a sampling frame, the panel recommends adding a question to the CBECS questionnaire to better understand the extent to which there is overlap between street addresses and the addresses where the building occupants receive their mail. For example, one challenge is that the DSF contains business-level entries, rather than building-level entries. Furthermore, some businesses have their mail delivered somewhere other than the street address (for example, to a post office box).

SUPPLEMENTARY DATA SOURCES

In addition to considering existing administrative records as an alternative source for a sampling frame, the panel recommends evaluating these records as potential sources for substantive data that could possibly replace an on-site interview at the building's location or could provide additional data for modeling or to conduct new analyses. Relying on data from other sources may become more of a necessity as it becomes increasingly expensive to maintain high response rates, even if an ideal sampling frame of commercial buildings were available. Although gathering and combining data from a variety of administrative records can also be resource intensive, the costs may go down as such data become more widely available, especially online.

Such other sources may also provide higher quality data. For example, EIA staff have expressed concerns regarding the difficulties associated with collecting data about the technical topics covered in the CBECS survey. Neither the typical interviewer nor the typical respondent is particularly knowledgeable about many of the items in the questionnaire, and this raises the question whether there are other sources that could provide better quality data.

Because the CBECS is currently the most comprehensive data available on the energy consumption of commercial buildings in the United States, there is no "gold standard" against which the quality of the survey or other potential data sources can be evaluated. However, a variety of other sources exist and can provide at least partial data. Comparing the data from several of these sources will help EIA begin to understand the relative advantages associated with each and the optimal ways of combining information from different sources. Ultimately, conducting in-person interviews for at least a subset of the sample may be necessary for

validation purposes, if the research indicates that these interviews produce the highest quality data.

In the rest of this section we discuss some additional sources of data that should be explored and validated at this stage, even if none of them, by itself, represents a realistic replacement for the building interviews. The panel is aware that EIA has considered the use of a variety of administrative records over the years. This option should be revisited periodically as the costs and data quality benefits associated with integrating these data sources evolve, and the 2011 CBECS presents an opportunity to carry out this research. Although different sources may be available for different types of buildings, for the purposes of evaluating them the overlap should be maximized to the extent possible. In other words, a subset of the buildings should be selected for this research and, for these buildings, data should be gathered from all of the available data sources being evaluated. The overlap will be particularly important with the building audits, as discussed below.

Building Audits

We understand that EIA has considered involving professional energy auditors to collect building data instead of relying on interviewers, but there has never been sufficient funding to implement this approach (except in the form of a small study in the past that was not conducted in conjunction with any of the CBECS data collections). We recommend testing the use of auditors on a small scale in the 2011 CBECS to better understand the costs and to collect data that can be used to assess the quality of other data sources. The data collected by the auditors would also be useful for evaluating some of the current back-end procedures, such as data editing, or the regression model used to identify outliers and to initiate a supplier follow-up survey.

For a subset of the buildings we think it would be useful if the same data were collected by both interviewers and auditors to allow the evaluation of the differences between these two in-person data collection approaches, in addition to comparing them to information collected from other data sources. Even if geographic representation cannot be achieved due to cost considerations, to the extent possible the test should include buildings of different sizes and with different characteristics. The data collection should be performed by professional energy auditors, who would carry out their work around the same time as the other data collection efforts relevant to a particular building, and without knowledge of any data already collected or available about the buildings from other sources.

Online Research

The panel recommends selecting a small subset of the buildings in the 2011 CBECS sample and investigating the information that can be obtained about them solely from the Internet. This could be set up in the form of a pilot test involving a small number of buildings (for example, 10 large, 10 medium, and 10 small buildings). The results will provide EIA with a better understanding of what types of data are available online in terms of both quality and extent. If this research is scheduled before the beginning of the data collection, the insights gained could be useful in fine-tuning the data collection instruments and sample design for the 2011 CBECS,

but this type of research could be carried out at any time. Once the 2011 CBECS data are available, examining the consistency between the information available from a variety of Internet sources and the data collected through the current methods, as well as from building audits, will provide valuable information about data quality.

LEED and Energy Star Certified Buildings

Data quality can also be assessed by taking advantage of the information that is publicly available online about buildings that have received leadership in energy and environmental design (LEED) or energy star certification. Comparing the data collected through the building interviews about a subsample or all of the buildings that have such certification to the data submitted as part of the certification process for the same buildings can also contribute to a better understanding of possible data quality issues. Naturally, certified buildings are a specific subset of the CBECS sample, and their characteristics are not representative of the entire population of interest for the survey. However, examining any discrepancies in the data available about these buildings can improve EIA's overall sense of the quality of the data and also identify potential areas of concern. The comparison could even be performed on data that have already been collected through a previous round of CBECS to inform the 2011 design.

Data from Energy Suppliers

The CBECS includes an energy supplier survey for about half of the CBECS buildings in the sample. The survey is initiated in cases where the energy usage and cost information cannot be obtained through a building interview or if the data obtained through the building interview are flagged as out of the expected range based on a regression model developed by EIA. To evaluate the quality of the data obtained from the building interviews, as well as the regression model used to identify out of range responses, the next round of the CBECS should include an energy supplier follow-up for more than just the problem cases currently included. For example, the 2011 CBECS could collect supplier data for a random sample of cases that provided usage and consumption data that were deemed valid during the building interview.

Given the increasing interest in time-of-use, hourly, and real-time energy use data, the panel recommends collecting this type of information as well, where available. EIA could aim to collect hourly data or time-of-use data (along with rates) from a random sample of the suppliers contacted for a follow-up interview, all suppliers who are contacted for a follow-up interview, or a random sample of the suppliers for buildings for which interviews were also conducted.

In addition, it is possible to specifically identify a few buildings with real-time metering and explore the data available from this source. The goal, again, would be to start establishing a framework for integrating this type of data into future surveys, understanding what type of information can be collected, and fine-tuning the procedures for working with a variety of (often reluctant) energy suppliers.

Digital Photographs

EIA has considered the use of digital cameras in the past, and this idea should be revisited, at least as a one-time research effort. A test could be accomplished by either providing a small number of the 2011 CBECS interviewers with cameras or in a separate operation from the 2011 data collection, if the latter approach is deemed more cost-effective because of training and operational considerations.

Photographs may provide valuable basic information about buildings. EIA's definition of a building does not always correspond to a respondent's definition of a building, and it is often left up to the interviewer to clarify the definition and come to an understanding with the respondent about what is meant by a building for the purposes of the interview. In addition, EIA staff indicated to the panel that the buildings of interest tend to use a wide range of specialized equipment related to building activity, and respondents' abilities to describe the equipment vary. EIA should evaluate the extent to which the pictures of the buildings and critical equipment, including nameplates, are useful in the data cleaning and editing process (for example, for reconciling ambiguous or questionable entries) and whether their use could translate into cost savings in comparison to the current procedures. EIA should also investigate the privacy and confidentiality concerns and regulations that may be applicable to the potential use of cameras in this context, even if the pictures are only used for data cleaning and editing.

Geocoding

EIA should examine the costs and benefits of supplementing the data available about each case in the sample with the geographic coordinates of the building's address. Geocodes could be added to the sample in house or during the data collection process. The former approach would probably be less precise, so capturing this information during the field work would be preferable if the interviewers can be equipped with the necessary devices at a reasonable cost. Adding geospatial information to each of the cases in the sample will enable researchers to conduct additional analyses of the CBECS data. For example, EIA currently integrates weather data from the National Oceanic and Atmospheric Administration (NOAA) into some of its analyses, and recording the building's proximity to the closest weather station would expand the analytic possibilities. Again, EIA should conduct research on any potential confidentiality concerns related to the use of this type of data and whether there are ways of appending additional geographic information to the data while maintaining confidentiality.

Other Data Sources

The panel recommends evaluating other existing data sources that EIA has considered in the past, as well as the breadth and consistency of information that could be obtained from local governments. The availability of more funding for the 2011 CBECS than has been available in the past provides a unique opportunity to carry out research that can inform future decisions about the design of the survey. Even if the information available from the various data sources is limited in scope, the recommended research can provide valuable feedback about the quality of the self-reported building data and identify options for integrating a variety of data sources in the future.

DATA COLLECTION

Process

The panel understands that EIA staff participate in all interviewer training, but even more active involvement may be necessary to share the study's goals and communicate how the quality of the data determines their usefulness. In addition, EIA staff members are also best qualified to conduct training on topics and concepts that are complicated, as a result of a long institutional history, such as the definition of a building and of a qualified respondent.

Additional resources should be invested in analyzing the characteristics of the field operations and in identifying opportunities for increased efficiency. EIA should review any information available from the data collection contractor regarding the amount of time spent on cases of various types (such as buildings with different characteristics, respondents with different backgrounds, etc.). If the case level contact history is not recorded in sufficient detail, efforts should be made to capture this information in the future. In addition, EIA should ask the data collection contractor to schedule debriefings with the interviewers soon after the beginning of the field period, and EIA staff should attend these debriefings to better understand how interviewers spend their time in the field, what types of cases are presenting the biggest challenges, and why. A detailed analysis of the time allocation should reveal whether there are subsets of cases that require a disproportionately large amount of time to complete and whether the effort is justified in the context of data needs and statistical techniques available to compensate for missing information.

EIA should also work closely with the data collection contractor to review the procedures used to select the best respondent for the building interviews and identify opportunities to streamline this process. Again, debriefings with interviewers can provide invaluable feedback that can help fine-tune the process and contribute to the development of new interviewer protocols. More efficient procedures for identifying a qualified respondent can not only reduce costs, but also address some of the concerns related to the technical nature of the questions. The qualitative feedback from the interviewers can then be further examined with an analysis of the quantitative responses by respondent type to identify possible differences in data quality. In other words, it is possible that most of the questions are not "too technical" if posed to the right respondent.

Additional activities for which analyzing existing data could identify opportunities for increased efficiency include the handling of partial interviews, both in terms of the field operations and from the perspective of data editing. Given that the CBECS interview is relatively long (with an estimate of 30 to 45 minutes provided to respondents), interviewer debriefings could reveal new strategies for approaching buildings and asking for appointment times. It would also be useful to understand whether there is a pattern to at what point the partial interviews end and whether the order of the items in the questionnaire could be rearranged to make the partial interviews more useful to EIA for either weighting or imputation.

Interviewers can be a good source of background and contextual information on questions that are difficult to administer, especially on whether particular questions are leading to partial interviews or possible data quality concerns. Discussions with the interviewers could represent

the beginning of a close examination of the questionnaire that has evolved with a face-to-face administration in mind and may need revising or simplifying to accommodate different future modes of data collection, as described in the next section.

Modes

Although EIA has considered the use of other modes of data collection, CBECS data are still collected primarily by in-person interviewing. In part because applying the CBECS definitions to determine the boundaries of a building is not always a straightforward task, as discussed above, EIA has continued to rely on face-to-face interviews. Identifying the most appropriate respondent is another task that is thought to benefit from the presence of an interviewer. Interviewers also carry hard-copy "show cards" that list the answer options for specific items and can be handed to the respondent to assist with answering questions that may otherwise be too difficult to remember if only read by an interviewer. In addition, one of the roles of the interviewers is to scan utility bills if they are available.

To prevent declines in the response rates and to limit costs, EIA will have to revisit the use of other modes of data collection, particularly the possibility of a multimode approach, with at least a portion of the interviews being conducted online. Transitioning at least a subset of the buildings to the web will free up some resources in the long run, which then can be allocated to the more complex cases and possibly invested into increasing the sample size. Although collecting this type of data on the web will present some methodological challenges, the panel believes that these challenges can be addressed and that web data collection may also represent some methodological advantages, in addition to the likely cost savings.

One possible approach that should be explored is to divide the sample into buildings that can be relatively easily transitioned to a web administration and buildings with more complicated characteristics that may benefit from interviewer administration. It may also be necessary to treat large buildings differently from smaller ones. The review of the case histories and the interviewer debriefings described above will be helpful in beginning to identify the building types for which data collection is fairly straightforward.

Until reliable auxiliary data sources can be integrated into the data collection process, a first in-person visit to each building will still be useful. During this visit, interviewers should follow a protocol developed by EIA to determine whether a second in-person visit is necessary (as is currently done) or whether the building is a good candidate for a web interview. Given the concerns related to the definition of a building, the decision of whether a case can be transitioned to the web will likely have to depend in part on whether the definition seems straightforward, as it would be, for example, for a small, standalone building occupied by one business. Resources should be invested in testing ways of communicating the definition of a building through a self-administered format, in anticipation of possibly being able to transition more and more complex buildings to web administration in the future.

The logistics of the best way to collect contact information for a web survey would have to be explored. Possible options include obtaining the information during the first visit or by telephone. Sometimes information on how to access a web survey is included in a hard-copy

advance letter mailed to respondents, even though this is less ideal than an e-mail invitation because it requires respondents to manually enter the web address of the survey and the login information. Since contacting respondents by mail may be the only option if an e-mail address cannot be obtained, investigating the extent to which building addresses and the mail delivery addresses overlap (as recommended above) will also be useful for this purpose.

There is no question that identifying the best respondent for completing the interview is crucial in the case of the CBECS, but exploring ways of accomplishing this without involving an interviewer should be examined. It is possible that a web option could in fact contribute to more interviews being conducted with qualified respondents. In some cases, it may be easier to forward a questionnaire to the right person than to locate him or her in a building and arrange an interview. In-person interviewers may also have an incentive to complete an interview as quickly as possible by settling for a willing respondent rather than pursuing the most appropriate one.

If a questionnaire is available on the web, it is also easier for several respondents to collaborate, each completing the sections he or she is most knowledgeable about. In addition, a web option could result in more complete data because it gives respondents the option to obtain information for questions they are not sure about and resume the survey later. Naturally, if respondents stop or forward a survey, there is a risk that they will not return to complete it, so an extensive follow-up effort is likely to be necessary. However, when a topic is too technical for many respondents, such as is the case of the CBECS, this kind of follow-up could make a significant difference in data quality.

The CBECS questionnaire relies very heavily on show cards, which is another reason why the survey is administered face to face. However, the use of show cards raises the concern of order effects, especially because many of the show cards have a large number of answer options listed, making it difficult for respondents to focus equally on all of them. For example, the show card listing the answer options for the primary activity in the building contains 16 items. A respondent for a building with multiple activities may be tempted to select the first one that is applicable as the “primary” activity instead of carefully reviewing the entire list. A web questionnaire would make it easier to restructure these questions into layered sets of items, with fewer answer options, or to reduce the possibility of primacy effects with the use of innovative methods, such as the animated presentation of response choices or an eye-catching emphasis on the end of the list.

The additional funding available for the 2011 CBECS represents an opportunity to test various ways of asking questions that EIA has identified as problematic because of their technical nature. For example, different approaches to obtaining the square footage information from respondents can be tested in the form of a split-sample experiment. Deconstructing this kind of an item into a series of questions would introduce complex skip patterns, but it would be easy to implement on the web without increasing the cognitive burden on respondents. Web administration can integrate various aids and tools for respondents, such as definitions or diagrams that can pop up if a respondent seems to be having trouble with a question or requests help. The interviewer debriefings described above will be useful in pinpointing specific questions that could benefit from a different approach and whether web administration is a promising option to pursue.

As is always the case with self-administered surveys, providing respondents with an e-mail address and toll-free telephone number they can use if they have questions may be valuable. The staff accessible through these means should be able to provide assistance related to the technical topics in the questionnaire, as well as answer to questions specific to the web administration.

Finally, when evaluating the implications of transitioning to a mixed mode administration, options for collecting the utility bills that are currently collected during the interview should also be considered. Some respondents may be able to easily upload an electronic copy of their bills through the questionnaire website, and this possibility should be investigated. Asking respondents to mail a copy of their utility bills would probably not be cost-effective because extensive follow-up would likely be necessary. The options should be assessed in the context of the research conducted to evaluate the possibility of increased reliance on supplier data.

The ideal time for beginning to explore the feasibility of transitioning some of the sample to web administration and conducting experiments on question wording is in parallel with the other data validation efforts, especially the involvement of the energy auditors. This timing will allow EIA to allocate some of the funding currently available to cover the cost of the transition, and it will provide an opportunity to take maximum advantage of the data collected from different sources. An analysis of the data collected from different sources can also guide decisions on whether the in-person interviews can be used in the future to calibrate the data collected through the web. All of the research should keep long-term plans in mind, such as the characteristics of the web as a data collection mode, even if the integration of web interviews is not realistic for the 2011 CBECS.

DATA RELEASES

The panel learned that the schedule of the data releases is a major concern to users who would like to see the lag between the data collection and release date reduced. EIA has been working on taking greater advantage of the Internet to facilitate data distribution.[1] We note that introducing a web option during the data collection stage can, in the long run, reduce the time necessary for preparing the files for release by reducing data editing and cleaning time.

The panel also recommends evaluating the possibility of eliminating some of the editing steps by reducing the number of editing rules or the number of variables edited or by focusing on cases that have the most impact on the estimation. Many data users will not only appreciate a shorter lag between data collection and release, but may also prefer access to data with fewer edits.

SUMMARY

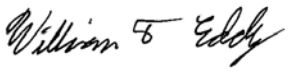
The 2011 CBECS presents an opportunity to conduct research that can guide the redesign of the survey on the basis of empirical data about the most cost effective approach for collecting valid and reliable information about the energy consumption of commercial buildings. This letter

[1]Energy Information Administration. (2009). *State Energy Data Needs Assessment*. Report SR-EMEU(2009)01. Washington, DC: U.S. Department of Energy.

report of the panel outlines a variety of research topics that seem most promising to pursue before or as part of the 2011 CBECS data collection. EIA should focus its efforts on (1) evaluating the availability and quality of alternative data sources that could assist with sampling frame development and potentially provide substantive data, and (2) developing a strategy for transitioning some of the interviews to a web-based data collection mode. This research will inform a possible future redesign of the sampling methodology and revisions to the data collection procedures that could be considered for subsequent rounds of the CBECS.

We hope this letter and our recommendations are helpful to you in planning the 2011 CBECS.

Sincerely,



William F. Eddy
Chair

APPENDIX A

Panel on Redesigning the Commercial and Residential Energy Consumption Surveys of the Energy Information Administration

WILLIAM F. EDDY (Chair), Department of Statistics, Carnegie Mellon University
MARILYN A. BROWN, School of Public Policy, Georgia Institute of Technology
FREDERICK CONRAD, Institute for Social Research, University of Michigan
DON A. DILLMAN, Social and Economic Sciences Research Center, Washington State University
DWIGHT K. FRENCH, Energy Consumption Division, Energy Information Administration (retired)
JACK G. GAMBINO, Household Survey Methods Division, Statistics Canada
CLARK W. GELLINGS, Electric Power Research Institute
JANE F. GENTLEMAN, National Center for Health Statistics
DAVID G. HUNGERFORD, California Energy Commission
PHILLIP S. KOTT, RTI International
NINA S-N LAM, Department of Geography and Anthropology, Louisiana State University
ALAN K. MEIER, Lawrence Berkeley National Laboratory
MICHAEL M. MEYER, Google

KRISZTINA MARTON, *Study Director*
MICHAEL COHEN, *Senior Program Officer*
NANCY KIRKENDALL, *Senior Program Officer*
AGNES GASKIN, *Administrative Assistant*

APPENDIX B

Panel Charge

The Energy Information Administration asked the National Research Council of the National Academies to conduct a comprehensive 30-month study of the Commercial Buildings Energy Consumption Survey (CBECS) and Residential Energy Consumption Survey (RECS). The panel's charge is to consider for these two surveys how to improve data quality, geographic coverage, timeliness of data releases, and relevance of the data for meeting user needs. The panel's work will include a review of survey design, frequency, and scope options, survey practice and operations, and the role that auxiliary data could play in improving survey coverage and editing and imputation methods.

The panel was asked to issue a letter report by spring 2010 that comments on design and data collection options for the 2010 CBECS to enable it to support Department of Energy program information needs, reduce respondent burden, and increase the quality and timeliness of the data. The panel will issue a final report at the conclusion of a 24-month study that makes recommendations for the design and conduct of CBECS and RECS and the dissemination of CBECS and RECS data for the next decade and beyond, including consideration of the level of resources likely to be required in comparison with the current survey program.

APPENDIX C

Glossary

Area probability sample	a sample generated by dividing a geographic area into a number of smaller areas, and then sampling from a subset of these areas
Coverage error	bias resulting from the omission of units from the sampling frame
Energy star certification	a joint program of the U.S. Environmental Protection Agency and the U.S. Department of Energy that certifies energy-efficient products
Geocoding	the process of appending geographic identifiers (codes or coordinates) to an address
LEED certification	a building certification system providing verification that a building was designed and built according to set of “green” standards
List sample	a sample generated from a sampling frame that exists in a list form
Multistage sampling	a sampling process involving several stages, in which units at each subsequent stage are subsampled from previously selected larger units
Sampling frame	the set of units from which the sample is selected
Sampling units	the individual units selected from the sampling frame
Show card	an interviewing aid consisting of a paper version of answer options or definitions associated with questionnaire items and used during an in-person interview when the questions are read to the respondent, and may be too difficult to understand or remember without a visual aid (also referred to as hand cards or flash cards)
Stratified sample	a sampling technique that involves dividing the sampling frame into distinct subgroups of similar units, and then selecting a separate sample from each of the subgroups

Appendix D. State Energy Data Needs Assessment

SR-EMEU/2009-01

State Energy Data Needs Assessment
January 2009

Energy Information Administration
U.S. Department of Energy
Washington, DC 20585

Contents

Executive Summary

Section 805(d) of the Energy Independence and Security Act of 2007 (EISA), Public Law 110-140, requires the Energy Information Administration to assess State-level energy data needs and submit to Congress a plan to address those needs (see Appendix A). In response, this report identifies gaps in EIA's current State-level energy data programs based on stakeholder outreach and internal deliberation, and outlines 34 possible initiatives to close those gaps.

To help users understand in ballpark terms what various enhancements to State-level data might cost, the report provides some information on the potential costs of the possible initiatives. It is not, however, intended as a budget document. In particular, it does not address the prioritization among the different possible initiatives, or their priority relative to improvements in EIA's national or regional data and analysis programs. Moreover, in many cases, the initiatives discussed in this report, such as expanding the sample size for the end-use energy consumption surveys, could be implemented on varying scales and with possible opportunities for cost-sharing, both of which could significantly impact actual implementation costs.

EIA Background. The Energy Information Administration (EIA), created by the Congress in 1977, is the statistical agency of the U.S. Department of Energy. EIA conducts a comprehensive data collection program that covers the full spectrum of energy sources and energy flows, generates short- and long-term domestic and international energy forecasts, and performs analyses on a wide range of energy topics. EIA's policy-neutral energy data and information are designed to meet the needs of Government, industry, and the public for the purpose of promoting sound policy decision-making, efficient markets, and public understanding. EIA disseminates its data products, analyses, reports, and services primarily through its Web site and telephone contact center. Major users of EIA's work products include the Congress, Federal and State government, industry, academia, financial institutions, news media and the public. By law, EIA's products are prepared independently of Administration policy considerations.

Report Organization. The report is organized as follows. Chapter 1 reviews the process EIA undertook to gather information from stakeholders and customers, consistent with the strong emphasis on consultation in EISA Section 805, and from EIA staff and managers. The process was informed by the direction in EISA Section 805(c) to ensure "the quality, comparability, and scope of State energy data, including data on energy production and consumption by product and sector and renewable and alternative sources, required to provide a comprehensive, accurate energy profile at the State level. . . ." Chapters 2 through 4 cover three major categories of gaps: the future of EIA's integrated State energy data programs, the quality and scope of EIA's end-use consumption data, and the quality and scope of data obtained from EIA's surveys of energy suppliers and markets. Each chapter describes user needs and develops initiatives to address the identified data gaps. Chapter 5 describes EIA's current efforts and proposed plans regarding communication, collaboration, and outreach to State energy data users and discusses opportunities and constraints related to the sharing of company-level energy data.

Integrated State Energy Programs. EIA currently collects and publishes data at different geographic levels depending on the type of fuel or energy source (e.g. petroleum, natural gas, coal, uranium, renewables, and electricity) and the type of data being reported, such as prices, reserves, supply, distribution, trade and marketing, consumption, sales, and end use. Only data relevant to a given State are collected and reported for that State because not all States have production plants, refineries, storage facilities, pipelines, ports, etc. Additionally, insufficient sample sizes can prohibit State-level aggregations, and some data are more easily interpreted if they are reported by regions or districts that are associated with established production, marketing, or transmission boundaries. While EIA posts State-level data throughout its Web site, EIA has two State energy data compilation programs to assist data users. The

State Energy Profiles (SEPs) is a data portal to all of EIA's State-level data and fuel-specific State energy profiles, essentially a one-stop shop for users interested in a specific State's energy picture. The *State Energy Data System (SEDS)* fulfills an additional need by providing annual *estimates* of State-level consumption, prices, and expenditures for those seeking time-series data on key energy variables. Both *SEPs* and *SEDS* (which rely essentially on EIA data) would benefit from improvements to EIA's State-level data collection programs but would also benefit from investments aimed specifically at improving the quality of SEDS estimates and reducing the time lag between estimation and release. *SEPs* investments could streamline data integration, add on mapping features, and improve timeliness of release.

End-Use Consumption Surveys. A wide range of stakeholders consulted prior to the writing of this report expressed strong interest in a variety of enhancements to EIA's surveys of energy consumption in end-use sectors, including larger sample sizes to increase the availability of State-level data, more frequent end-use surveys, faster processing and issuance of end-use survey data, and expansion of the end-use survey program to additional sectors, such as agriculture and household transportation energy. Options for improving the consumption survey data programs include expanding the current survey designs to cover up to all 50 States. An increase sample size and/or the addition of even a limited number of States would be a vast improvement over current programs because it would permit more complex analysis of key indicators of energy use, publication of more building types, and more accuracy for secondary uses of the data by other Federal agencies. Additional options aimed at improving end-use data and estimates focus on restoring and adding new end–use surveys, conducting feasibility studies to explore alternative methods to update end-use estimates, and decreasing the time between when data are collected and released.

Supplier Surveys: Data Quality and Scope. In addition to data gaps associated with end-use data, stakeholders identified needs associated with EIA's supplier surveys. EIA's supplier surveys cover all aspects of the energy supply picture (petroleum, natural gas, coal, renewables and electric power) and are published at national, regional, and State levels, depending on the survey, the sample, and confidentiality requirements. Stakeholders and customers indicated that EIA supplier data are important for State-level policymaking as well as for responding to emergencies and understanding markets within and affecting their States, and they emphasized the importance of maintaining quality in EIA data. Taking the feedback we received from stakeholders over the years, as well as from workshop participants and others during our 2008 outreach efforts, EIA developed several data quality options aimed at reducing statistical error in EIA surveys. The options include increasing sample sizes to provide better statistics, improving survey coverage by updating frames (the list of survey respondents), and adding new surveys to obtain data not currently collected. Current statistical data edits, nonresponse follow-up methods, and other internal EIA quality checks are covered by strict adherence to EIA statistical standards and controls.

Communication, Collaboration, and Outreach. One of the requirements of Section 805, as noted previously, is "to ensure the quality, comparability, and scope of State energy data, including data on energy production and consumption by product and sector and renewable and alternative sources, required to provide a comprehensive, accurate energy profile at the State level. . . ." EIA's stakeholders and customers note that it takes Federal initiatives to provide data that are comparable across States. Consequently, several options discussed in the report are aimed at expanding EIA's interaction with the States, as well as the U.S. territories, and include hosting conferences and workshops to assess data needs, and options aimed at using collaborative, Web-based tools to share information.

Resource Considerations. EIA's 2009 budget request of $110,595,000 supports primary data collection, data processing and dissemination, short- and medium-term forecasting and economic and industry

analysis. Costs (both start-up and ongoing) associated with each of the 34 possible initiatives to address State data gaps are *preliminary* estimates and are presented in variety of ways depending on the proposal (all costs include Federal staff and contractor support). For example, some costs are reported as increments to annual costs as presented in EIA's 2009 budget request. This is the case for initiatives that address weekly, monthly, or annual surveys and/or initiatives that are improvements to data quality. For improvements to the end-use (consumption) surveys, which are conducted every 4 years, EIA presents preliminary cost estimates in terms of increments to the projected budget for the next upcoming consumption survey cycle. EIA's end-use consumption survey program is currently funded on a levelized basis over the 4-year cycle required to complete the current set of 3 surveys; annual incremental resource requirements after an initial transition period can be estimated as the cycle cost divided by the cycle length. For initiatives aimed at increasing the frequency of the consumption surveys (to every 2 or 3 years), the costs are per-cycle costs, noting that this also translates into larger *annual* budgets to fund shorter cycles on a levelized basis. In addition, some proposals identify needs that require coordination and resource commitments among statistical agencies. Consequently, the costs of the proposals in this report cannot easily be added together for a grand total. More accurate budget numbers, both start-up costs and per-survey cycle costs, would require more detailed assessments. Furthermore, EIA's budget is part of the broader Federal budget, and EIA's funding must be weighed against other Federal priorities.

In some cases, improvements to EIA's State energy data can be undertaken for a relatively small cost, particularly improvements associated with supplier survey data quality, State data integration programs, and outreach. Expanding the consumption survey programs to include more or all States would be much more costly. For example, EIA currently projects that the 2009 RECS, the 2010 MECS, and 2011 CBECS will cost roughly $28,000,000 to complete (or about $7,000,000 on an annual basis). Simply expanding the RECS by 50 percent more sample units would cost about $3,000,000 more per 4-year RECS cycle. Similarly, increasing the sample size for the quadrennial CBECS by 50 percent would cost about $7,000,000 more than what EIA currently projects for the 2011 CBECS.

Compiling this report has helped EIA look carefully at gaps in its State-level data collection efforts and develop options to address them. As it is unlikely that EIA would receive the budget to support all the initiatives in this report, it is important for EIA, along with its stakeholders, to consider State energy data gaps within the context of gaps in EIA's other data and analysis programs to be sure funds are directed at the most critical needs. It is equally important to seriously consider options to the expansion or improvement of existing State energy programs. The inclusion of several feasibility studies in this report recognizes the value of fleshing out some of these options.

An old management adage that still holds true today in a variety of energy and other contexts states that you can't manage what you don't measure. Given the current interest in energy and energy efficiency issues at the State level, as well as at the Federal level, it is not surprising that there is considerable interest in more, and more timely, State-level energy data. EIA can respond to this interest in several ways. First, EIA can continue to seek to operate as efficiently as possible to allow the use of a portion of our existing resources to address new priorities, including State-level data. Second, EIA can propose reallocating resources away from existing EIA activities towards work that is judged to have a higher priority. EIA has done this on several occasions over the past 5 years, and in many cases has met strong resistance by users (and/or their representatives) of the existing data proposed for elimination. A third option would be to seek additional resources through the budget process. Decision makers in the Administration and the Congress could then consider our request within the context of two overarching questions. First, what overall level of resources should be devoted to the energy mission of the

Department of Energy, including, but not limited to EIA? Second, what portion of overall energy mission resources should be devoted to energy data programs at EIA?

Chapter 1. Consultations with Customers, Stakeholders, and EIA Staff

Stakeholder Outreach

EIA's first step in responding to the requirements of section 805(d) was to form the State Data Assessment Team consisting of six EIA analysts who conducted the necessary research, gathered input from stakeholders[1], developed options to fill data gaps, and compiled this report. Over the past 15 years, the number of people and organizations seeking information from EIA has increased significantly. Additionally, their interests in and needs for different types of energy information have constantly changed and expanded**.** EIA's formal stakeholder outreach related to the preparation of this report began with a presentation by the Deputy EIA Administrator to the National Conference of State Legislatures at its July 2008 Legislative Summit in New Orleans. During September and October, the State Data Assessment Team consulted with additional stakeholders. Team members held meetings and telephone conferences with representatives of the National Association of State Energy Officials (NASEO) and the Northeast-Midwest Institute. They attended NASEO's Annual Meeting in September 2008 in Overland Park, Kansas, where they met informally with State officials and gave a presentation that explained EIA's work on Section 805. NASEO members expressed their need to for greater accuracy in EIA's data, more access to company-level data, and more data that would help States track progress towards meeting energy-related goals. For example, two of the goals established by the Environmental Protection Agency's (EPA) *National Action Plan for Energy Efficiency Vision for 2025*[2] require the establishment of mechanisms for measuring and evaluating a State's progress in achieving greater energy efficiency. States are requesting State-level data on energy consumption that would assist them in measuring their progress toward reaching this goal.

In October 2008, EIA held a workshop inviting attendees to share their perspectives on the scope, accuracy, quality, comparability, and timeliness of EIA's State energy data. Announcements about the workshop were sent to two of EIA's extensive list serves (State Energy Data and Energy Consumption) and to the NASEO Winter Fuels Conference mailing list. Recipients of the announcements represented a wide range of organizations, including the:

- Federal Energy Regulatory Commission
- Environmental Protection Agency
- U.S. Departments of Transportation, Commerce, Labor, Treasury, and Health and Human Services
- State Energy Officials (in all 50 States and the U.S. Territories)

---

[1] The U.S. Government Accountability Office defines a "stakeholder" as an individual or group with an interest in the success of an organization in delivering intended results and maintaining the viability of the organization's products and services. Stakeholders influence programs, products, and services. Examples include congressional members and staff, representatives of central management and oversight entities such as the Office of Management and Budget; and representatives of key interest groups, including the organization's customers and interested members of the public. EIA's major customer groups include: Federal, State, and local governments; academic and research communities; businesses and industry; foreign governments and international organizations; news media; financial institutions; and the general public.

[2] National Action Plan for Energy Efficiency (2008). *National Action Plan for Energy Efficiency Vision for 2025: A Framework for Change.* <www.epa.gov/eeactionplan>

- National Conference of State Legislatures
- National Governors Association
- National Association of Regulatory Utility Commissioners
- American Council for an Energy Efficient Economy, the Alliance to Save Energy, the National Association of Home Builders, and
- ExxonMobil, Amerigas, and other energy companies.

Over 80 people registered for the workshop, which was attended by representatives of the Federal and several State governments, including the California Energy Commission and the New York State Energy Research and Development Authority (NYSERDA), non-profit organizations, and private companies. Several workshop participants also provided written comments. Workshop topics included the availability of measures of renewable energy consumption; electricity grid reliability; data on distributed generation, plug-in vehicles, and other new technologies; end-use data on appliances such as gas clothes dryers; and fuel prices for home heating, among many other topics.

EIA In-House Assessments

The State Data Assessment Team initially reviewed EIA's State Data Directory http://tonto.eia.doe.gov/state/SEP_MorePrices.cfm , which helped serve as a guide when considering data gaps. Because EIA's analysts are continually responding to customers, they are uniquely aware of EIA's data gaps, so the Team then canvassed EIA management and analysts to identify gaps and propose options to address the gaps. EIA staff addressed topics such as survey data quality and coverage, State-level data estimation procedures, and information dissemination. Lastly, the Team reviewed public and Congressional requests for State-level information received by EIA during the first 5 months of 2008.

Results of the Assessment

After analyzing all of the feedback received from external and internal stakeholders, the State Data Assessment Team categorized State data needs into four main categories:

1) "Integrated State Energy Data Programs," which covers improvements to the dissemination of data and estimates via the *State Energy Profiles (SEPs)* and the *State Energy Data System (SEDS)*;
2) "End-Use Consumption Surveys: Data Quality and Scope," which covers data needs to increase the comprehensiveness of EIA's existing and proposed end-use consumption data collection activities;
3) "Supplier Surveys: Data Quality and Scope," which covers data needs associated with increasing the comprehensiveness of EIA's existing supply-side surveys; and
4) "Communication, Collaboration, and Outreach," which covers actions EIA could undertake to consult and collaborate with stakeholders.

The specific energy data needs identified by EIA and options for addressing them are covered in the next four chapters of this report.

Chapter 2. Integrated State Energy Data Programs

EIA receives comments and questions about State-level data every day. Some users are looking for comprehensive datasets, such as total energy consumption for all 50 States (and U.S. territories) for the past 10 years. Other users are looking for quick access to a comprehensive profile of their State's current energy situation. All users expect to find timely, technically accurate, and complete State-level data on the EIA web site. One participant in EIA's October 2008 State Energy Workshop told EIA, "We need the data for use in making policy. Energy data makes the markets work." Another participant said, "We rely on EIA's data–it's absolutely critical. It's the only source that's reliable State to State."

EIA has two State energy data compilation programs (referred to as "integrated energy" programs) that cover all of the major sources and uses of energy: the *State Energy Profiles (SEPs)* and the *State Energy Data System (SEDS)*. Both programs seek to address the needs of a wide variety of data users. The SEPs program was launched in December 2006 as an innovative web application to improve user access to and understanding of EIA's State-level data. It is a continually updated web-based portal to all of EIA's State energy data; it also includes individual profiles that present key facts and statistics about State energy markets and industries. Because SEPs is a compilation program, it can only include available data. Expanding the scope of SEPs to include, for example, more State-level or U.S. territory end-use data, would require expansion in EIA's data collection programs, discussed in other chapters of this report.

The *SEDS* has provided annual *estimates* of State-level consumption, prices, and expenditures since 1978. In the past few years, an intensive effort has led to significant improvements in timeliness and scope. EIA has begun posting updated estimates for individual energy sources as soon as they are processed. In addition, the time lag from the close of the data year to the release date of SEDS integrated totals (all energy sources) has been shortened from 34 months in October 2006 to 23 months in November 2008, with an ultimate goal of an 18-month time lag in the current system. In addition, SEDS began providing State-level production estimates in 2008 for all energy sources and States for 1960 forward.

In 2009, EIA will more closely align the programs to gain efficiencies and to provide users with an up-to-date, comprehensive, and accurate State Energy Profile for each State as quickly as possible after the data become available.

Improved Integrated State Energy Data Products

- *Improve Timeliness of the SEDS.* The SEDS provides annual estimates of consumption, prices, expenditures, and production for all energy sources and States. Many stakeholders, including the Northeast-Midwest Coalition and the National Association of State Energy Officials, have expressed concern about the timeliness of SEDS data, which are used as inputs for forecasting and for the estimation of greenhouse gas emissions. State administrative and energy officials also use the data for planning and analysis. The timeliness of SEDS depends on the availability of annual data from various EIA program offices and external sources. It would be difficult to improve timeliness to less than 18 months after the data year with the current schedule of data inputs into SEDS, but it would be possible to use estimation procedures to compute a set of *preliminary* SEDS estimates by using preliminary data from monthly and quarterly surveys. Using model-based estimates would allow the first set of SEDS data (for example, coal consumption) to be released 6 months after the end of the data year, and total energy estimates to be released in 15 months.

Initiative 2.1. Start-up Cost: $300,000; Annual Operating Cost $300,000 (increment over EIA's 2009 budget request).

- *Improve Quality of the SEDS.* In order to evaluate total energy use in the SEDS, a complete set of State-level consumption and price data for all energy sources must be compiled. There are gaps, however, in the data inputs needed to calculate total energy use. In some cases, source data come solely from non-EIA sources, and EIA has no control over the timeliness or quality of those exogenous data sources. For example, EIA uses asphalt and road oil consumption data from the Asphalt Institute and liquefied petroleum gas by end-use sector from the American Petroleum Institute. EIA has no control over the data quality, data collection, or publication schedule for those data. In other cases, there are no data available from any source. For example, there are no consumption data for lubricants used in the industrial or transportation sectors. When data are not available from EIA or non-EIA sources, they must be estimated. If the missing data could be filled in with EIA survey data, the quality of the SEDS data would be improved. It is unclear, however, what resources would be required to obtain all of the missing data. Therefore, EIA could conduct research to define the level of effort required to obtain the missing data. *Initiative 2.2. Start-up Cost: $130,000; Annual Operating Cost: $130,000 (increment over EIA's 2009 budget).*

- *Create a State Data Application Programming Interface (API).* Upgrading EIA's databases would also make it possible to implement a State Data Application Programming Interface (API), a computer model that makes it easy to access and exchange EIA's State energy data. An API would allow Federal and State agencies, financial markets, research institutions, analysts, and others to quickly access EIA State data and process it into new services beyond the scope and resources of EIA. Such an interface would allow EIA to give customers direct access to EIA data for a variety of purposes. EIA would be able to deliver State energy data in multiple formats in the timeliest manner possible. An API is also the first step to developing a global energy navigator–a database management tool that links multiple databases through a single interface. *Initiative 2.3. Start-up Cost: $500,000; Annual Operating Cost: $194,000 (increment over EIA's 2009 budget request).*

- *Develop System to Support User-Generated State Energy Maps*. Computer system upgrades would help make it possible for users to create their own State energy data maps. Congressional staffers, members of the press, and energy analysts routinely call EIA to request high-resolution copies of EIA's State energy maps, which display energy infrastructure sites such as electric power plants, transmission lines, natural gas pipeline flow and hubs, oil ports and refineries, and renewable energy potential at a glance. The maps are unique in that they display geospatial data for several different major energy sources. Currently, EIA's mapping environment does not support the specific geospatial analysis needs of some users. EIA is frequently asked but unable to provide maps showing crude oil and petroleum product pipelines. The addition of biomass potential and more detailed data on solar, wind, and geothermal potential are other common requests. A more robust computing environment would make it possible to include the large amounts of data required to display detailed oil pipelines, energy potentials, and other energy map elements. The analytical value of the maps would also be enhanced if they were published as a system that would allow users to generate their own maps, adding and subtracting map layers and creating customized, high-resolution maps suitable for analysis and publication. *Initiative 2.4. Start-up Cost: $330,000; Annual Operating Cost: $150,000 (increment over EIA's 2009 budget request).*

Chapter 3. End-Use Consumption Surveys: Data Quality and Scope

EIA's end-use consumption surveys were among the first fielded by EIA after its creation in 1977 and now cover energy end uses in three areas: the Commercial Buildings Energy Consumption Survey (CBECS) covers the commercial buildings sector, the Residential Energy Consumption Survey (RECS) covers the occupied housing portion of the residential housing sector, and the Manufacturing Energy Consumption Survey (MECS) covers the manufacturing share of the industrial sector.[3] These resource-intensive surveys are the only source of data for current estimates of energy end uses in homes and commercial buildings and for key industries within manufacturing. Due to limited resources, they are not currently meeting their statutory requirements for frequency and scope[4].

Two features limit the geographic level at which consumption data can be provided to the public: 1) sample sizes, which are driven by available resources and 2) data confidentiality laws, which require that EIA protect the identity of individual respondents and establishments. In practice, where firms make up a large share of their industry class or building type, they could be identified at lower levels of aggregation. However, EIA would be prohibited from publishing some statistics at the State level or would need to collapse disparate classes of data to protect its confidentiality. The same legislation, however, does allow certified agents of EIA to make limited use of non-public files for statistical purposes only.

Given current resources and confidentiality constraints, EIA is limited to publishing end-use data for the following geographic areas[5]:

- MECS: National and Census Region
- CBECS: National, Census Region, and Census Division
- RECS: National, Census Region, Census Division, and the four most populous States (California, Florida, New York, and Texas)

Stakeholder Needs

Stakeholders expect EIA to take the lead in providing energy consumption data that meet the quality and scope necessary to monitor topics related to climate, the environment, and energy security and they often request data to help evaluate energy programs and policies that are often written, funded, and implemented at the State level. They need more and new consumption data at lower levels of geography, more frequently, and with less lag time between the period of data collection and the release date. A Federal statistical program that can assess the value of a dollar invested in a particular program, technology, or system is a much broader and more complex data operation than EIA has ever run. Such a program would have profound resource implications for EIA.

Stakeholders report they need data for geographic areas at and/or below the State level–counties, metropolitan areas, or cities–to tie outcomes to specific programs. Stakeholders provide the following

[3] Although transportation represents about a third of domestic energy consumption, and residential transportation consumption about two-thirds of that, budget shortfalls have prevented EIA from collecting data on transportation for the past 15 years.

[4] Public Law 92-275 (Federal Energy Administration Action of 1974) and Public Law 95-91 (Department of Energy Organization Act).

[5] See Appendix B for a brief discussion on EIA geographic reporting levels.

arguments, among others, for larger sample sizes to improve data quality and to provide new estimates for smaller geographic aggregations:

- EPA reports that it needs a larger sample size for CBECS to produce energy performance benchmarks for more building types. EPA offers an online rating system called Portfolio Manager where commercial buildings can be rated for their energy consumption relative to similar buildings. This tool has been used for more than 78,000 buildings; about 5,600 have achieved an EnergyStar rating. Although this rating system has become the industry standard, EPA can only produce benchmarks for 10 broad building types. An increase in the CBECS sample size would be needed to produce them for more diverse building types and principal activities.

- The American Society of Heating, Refrigerating and Air Conditioning Engineers (ASHRAE) states that the CBECS current sample size is too small to evaluate the impact of critical building labeling programs like EnergyStar by building type *within* climate zones. They also believe that States and cities rely on CBECS information to develop rating systems for "comparable buildings." Because data at that level are imprecise or absent, ASHRAE is concerned about the accuracy, utility, and impact of the ratings.

- The National Renewable Energy Laboratory (NREL) reports it needs much larger RECS and CBECS sample sizes to perform necessary multivariate analyses to understand the adoption rates and impact of the new technologies, building designs, and energy-efficient equipment they test and promote.

- The National Association of Home Builders (NAHB) indicates it needs larger samples to estimate how much consumption is explained by householder behavior versus that which the builder can control. Although NAHB makes no direct request for State-level estimates, building codes are an important feature in that analysis because they are enacted at the State and local level.

- The American Council for an Energy-Efficient Economy (ACEEE) produces an annual State Energy Efficiency Scorecard, which ranks States according to their adoption and implementation of energy efficiency policies and programs. Without sufficient State-level sample sizes in EIA's consumption surveys, there is no accurate, direct link between State policies and consumer participation in these programs, which represent large, non-Federal program investments.

- The Department of Housing and Urban Development (HUD) and its local housing authorities use RECS data to calculate utility allowances for subsidized housing by States and localities. State-level estimates would provide a direct source of data closer to what is needed for program management.

- Where utility costs are bundled with rent by landlords, the Bureau of Labor Statistics (BLS) uses RECS data to allocate portions of total rent to housing or utilities. Rents and utility costs vary widely for smaller areas, so BLS would benefit from lower levels of data aggregation.

- The Department of Health and Human Services (HHS) uses RECS to support program needs of the Low Income Home Energy Assistance Program (LIHEAP). Improved State-level estimates would provide marked improvement in allocating funds in this multi-billion dollar grant program to States.

- ACEEE notes that because EIA summarizes consumption data by Census Division, EIA necessarily aggregates statistics in the Mountain Division across a huge area—from Montana on the Canadian border to Arizona on the Mexican border. Program evaluations are not served by combining data across such diverse States and climates. Without additional sampling, EIA could increase breakouts in the Mountain Division only by decreasing the sample in the Pacific Division, which would reduce the accuracy of current estimates for California.

- Stakeholders want EIA to balance sample allocations between new and existing building stock for the RECS and CBECS to improve the accuracy of comparisons among and between buildings of different ages. Although Federal or industry programs offer explicit guidance on energy-efficient building design and practices, program participation is often voluntary. This situation is less true for State and local governments where codes can be tied to building permits. Therefore, State-level data would provide more accurate analysis of the sources of differences in energy consumption for new construction.

- Numerous officials from a wide array of State program offices, consulting firms, utility companies, and equipment manufacturers contact EIA directly for data that are not available without large increases in survey sample sizes: State-level energy intensities within consumption sectors including breakouts by industry, building type or activity, fuel, end use, and equipment type; and 'typical' intensities for specific buildings types, industry codes, or residential classes.

There is a recurring dilemma when trying to optimize sample allocations across such overlapping geographic variables as physical geography and climate zones. A good sample can achieve sufficient coverage by States or by climate zones but not both without increasing the likelihood that some buildings or establishments could be individually identified. For reasons of data confidentiality, EIA will probably need to continue to aggregate some data to mask the location of particular buildings.

Options for Improving Consumption Data Surveys

The consumption surveys draw samples from large, heterogeneous populations, and the survey methodologies necessary to produce accurate results mirror that complexity. Therefore, for RECS and CBECS, EIA samples, lists, and enumerates units selected from area clusters at rates that have to balance sampling error across key characteristics: geography, building size and type or activity, main fuel used for space heating, and so on. Area-probability field studies are the most expensive but provide the best frame coverage for complex populations, the highest item and unit response rates, and the ability to collect key data in physical form (e.g., household square footage measurements and interviewer observations of buildings characteristics). MECS is considerably less expensive per sampled unit because it is drawn from a known list of establishments with routine updates maintained by the Census Bureau and utilized for other economic data collections, and can be conducted by mail and the Internet.

Options for improving the consumption survey data programs are described below. The first options expand the current survey designs to improve data quality without State-level estimates, other than for the four most populous States in the RECS. Programs with expanded survey designs would be a vast improvement over current programs because they would permit more complex analysis of key indicators of energy use, publication of more building types, and more accuracy for secondary uses of the data by other Federal agencies. Short of a 50-State estimates program (or a 50-States and U.S. territories program), also described below, an intermediate option within each of these designs is to add some but not all States. Proposals are also included that describe data collection needed to produce baseline

measures for 50 States, although such a program would require considerably more resources. Additional options focus on restoring and adding new end–use surveys, conducting feasibility studies to explore alternative methods to update end-use baselines, and decreasing the time between data collection and release.

For each initiative or proposal, EIA estimates preliminary costs, which are presented in different ways depending on the initiative. Costs are presented either in terms of increments to EIA's 2009 budget request (for wholly new initiatives); increments to the projected budget for an upcoming survey cycle (in the case of initiatives that increase an ongoing survey's sample size or data scope to improve data quality); or relative per-cycle costs (for initiatives that address frequency of surveys). More accurate budget numbers (both start-up costs and per-survey cycle costs) require knowing, *inter alia*, which States/Regions would be added, the homogeneity of key characteristics within them, and the level of accuracy desired.

Enhance Data Quality by Adding Data Elements and Increasing Sample

There are many ways in which EIA could improve the value and accuracy of consumption surveys. By far, the most important is to improve the accuracy of end-use estimates. Managing the sample sizes and allocations among key characteristics is equally important as adding questions to capture new phenomena that effect consumption patterns and levels. Changes in sample coverage below address two stakeholder needs: 1) produce sufficient sample size to be able to perform needed multivariate analyses (i.e., having the statistical power to understand the relative contribution of three or more characteristics) and 2) produce accurate State-level energy intensities and end-use estimates with which States can monitor and evaluate energy programs.

- *Obtain More Geographic Detail on Fuel and Nonfuel Uses of Fossil Fuels for Manufacturing* (MECS). Doubling the quadrennial MECS sample size to 31,000 sample units would allow EIA to provide estimates for manufacturing energy consumption for industry groups by Census Region, improve the statistical accuracy of national analyses, estimate energy efficiency in this sector, and calculate changes in carbon emissions over time that result from structural change. This initiative would serve essential missions of many Federal, State and industry energy, environmental and commercial interests. *Initiative 3.1. Cost per 4-year Cycle: $2,800,000 (increment over EIA's projected cost for 2010 MECS).*

- *Enhance the Quality of the RECS.* Increase the quadrennial RECS by 50 percent to 9,750 sample units and add questions to improve the accuracy of multivariate data analysis and end-use estimates. For example, collect data on: the share of remodeling that is done to incorporate energy efficient equipment, systems, and designs; the degree of compliance with building energy codes (and which version or source) for new construction; and building operation, because behavior can often explain more variation in consumption than can technology and equipment. Conduct a periodic sub-metering study on a subsample of RECS households to measure actual energy use by refrigerators, hot water heaters, televisions, and computers and their peripherals. Compare actual consumption by end use and behavior to estimates based on non-linear models developed by EIA staff. Such a comparison would help EIA adjust for bias that might occur in an infrequent survey that lags in accounting for new technology, standards, and growing plug loads. *Initiative 3.2. Cost per 4-year Cycle: $3,160,000 (increment over EIA's projected cost for 2009 RECS).*

- *Enhance the Quality of the CBECS.* Increase the sample size for the quadrennial CBECS by 50 percent to 17,250 sample cases (which include buildings and establishments) and target specific building types that are big energy users, such as data centers, laboratories, convention centers, and arenas. Add questions to improve the accuracy of multivariate data analysis and end-use estimates. For example, EIA could collect new data on: the degree of compliance with building energy codes (and which version or source) for new construction; building operation, to isolate energy management practices from fixed factors; and building type, to explain more variation in energy consumption, e.g., linear shelf feet rather than number of refrigerators for groceries, size of the eating area for food service, volume of transactions/sales for retail, number of beds or rooms for dorms/hotels/hospitals, and number of service bays for auto repair shops. EIA could conduct a quality study on a subsample of the CBECS buildings in detail and compare the end-use energy use of these buildings to the estimates developed by EIA non-linear models. EIA could validate the information collected during the field interviews and also provide an independent estimate of the energy consumption by end use. This information could lead to improvements in future CBECS questionnaires, as well as measuring the quality of current CBECS estimates. *Initiative 3.3. Cost per 4-year Cycle: $6,880,000 (increment over EIA's projected cost for 2011 CBECS).*

Restore and Add New Surveys to Fill Data Gaps

In addition to increasing the geographic scope of consumption surveys to provide a neutral source of data for Federal, State, and local energy policy, EIA could increase coverage across consumption sectors to improve forecasts of short- and long-term energy demand. Data gaps are sometimes addressed by increasing sample sizes to permit publication of more subclasses, or by adding questionnaire items or surveys. To increase the coverage and accuracy of the amount and sources of growing energy demand, particularly for electricity and petroleum-based fuels, EIA would need to restore and add data collections for transportation and agriculture. The transportation sector is going through a dynamic era of technological, fuel, and industry change in response to energy prices and global economic and climate concerns. Measuring transportation energy consumption poorly, infrequently, or not at all has implications across all sectors of the economy. With the emergence of biofuels, energy and food policy are now more closely linked.

- *Collect Data on End Uses of Energy by the Residential Transportation Sector.* Restore the Residential Transportation Energy Consumption Survey (RTECS), which was discontinued after the 1994 data year due to insufficient funds. EIA could expand the current RECS to include a follow-on study of residential transportation end uses, which account for two-thirds of the entire transportation sector and most of motor gasoline consumption. Many stakeholders value a reliable, policy-neutral source of data to understand on-road fuel economy, price elasticities, vehicle miles traveled, commuting behavior, and vehicle purchases relative to new energy policies and technologies. *Initiative 3.4. Start-up Cost: $1,000,000; Cost per 4-year Cycle: $4,048,000 (increment over EIA's 2009 budget request).*

- *Collect Data on the End Uses of Energy by the Non-Residential (Truck) Transportation Sector*. To increase coverage of the transportation sector, EIA could sponsor, in part, the collection of data on the physical and operational characteristics of the Nation's private and commercial truck populations. Until 2002, the Census Bureau conducted a truck survey with its quinquennial economic census. A similar truck survey, with additional questions about fuels used, energy end-uses, and costs would be beneficial. The series would produce national and State-level estimates

of the total number of trucks and their end uses. Expanding the survey to the State level could require data coordination and cost-sharing with Federal, State, and local agencies. These data could serve missions of Federal energy and transportation agencies by providing a comprehensive data set for assessing energy efficiency and the environmental impact of the Nation's truck fleet. Cost estimates only assume EIA's role adding fuel-related questions. *Initiative 3.5. $3,000,000 (increment over EIA's 2009 budget request).*

- *Collect Data on End Uses for the Agricultural Sector*. EIA currently collects data on industrial sector energy end uses only for the manufacturing portion. Adding an agricultural survey would improve the industry coverage in EIA's State Energy Profiles. For example, fuel rates vary significantly by scale and type of operation, but EIA now assumes that commercial rates prevail. An agricultural series would provide the only baseline to: measure opportunities for new energy-efficient technologies and practices; allow a more uniform evaluation of the impact on production agriculture of Federal and State energy policies, such as fuel tax abatements, efficiency incentives, and alternative fuel use; and produce estimates of the share of greenhouse gases resulting from different enterprises, production practices, and technologies. Two agencies would be served: the Department of Agriculture (USDA) could measure the farm-level response to changes in energy prices and supply (food security), and the Department of Energy could study the flow of crops into biofuels compared with other uses (energy security). Cost estimates only assume EIA's role in a new agricultural survey. *Initiative 3.6. Start-up Cost: $200,000; Cost per 4-year Cycle: $1,500,000 (increment over EIA's 2009 budget request).*

Produce End-Use Data for All 50 States

Developing and operating a 50-State consumption data program (or a 50-State and U.S. territories data program) would require significantly more resources than EIA's current program and significantly more than any of the proposals discussed so far. For example, a preliminary estimate for adding States to RECS would be $750,000 for each additional State. For the CBECS, each additional State would add about $1,200,000 to the total survey budget. State selection criteria would vary according to measurement goals, accuracy and confidentiality requirements, and costs relative to sampling efficiencies for other States. For example, if States were selected according to population or building rank, coverage would quickly increase for one Division in the Midwest[6] (Illinois, Ohio, and Michigan in the East North Central Division) and one in the South (Georgia and North Carolina in the South Atlantic Division). Population or the number of buildings as the main selection criteria would yield no data improvements for the Mountain or West North Central Division States and might not meet the goals of States with aggressive energy policies, stakeholders such as the ACEEE, and many Federal agencies with little or no data to monitor programs. The criteria for adding States would need to be carefully developed, with stakeholders' input, and clearly communicated.

- *Provide Residential Energy Consumption Data for All 50 States (maintaining current 4-year cycle)*. Increasing the RECS sample to cover 50 States would ensure that new (often State) policies can be monitored for their impact on fuel type used, intensities, and end uses. Accuracy would improve for national, regional, and division-level estimates so that analysts could isolate the effect of such factors as new efficiency standards, building technologies, and program participation from factors over which consumers have no control, such as weather. EIA would have sufficient sample counts to produce estimates for small appliances and home electronics–a growing portion of residential

[6] See Appendix B for a map of the U.S. Census Regions and Divisions.

consumption, a source of greenhouse gases, and an opportunity for technology change and innovation. Expanding the program to include all States would serve essential missions of many Federal, State, and energy industry, environmental, and commercial interests. *Initiative 3.7. Start-up Cost: up to $8,000,000; Cost per 4-year Cycle: up to $26,460,000 (incremental cost over EIA's projected cost for 2009 RECS).*

- *Provide Residential Transportation Data for All 50 States (maintaining current 4-year cycle).* If residential transportation data were deemed necessary to collect via a follow-on survey to the 50-State RECS design, it would require additional funds beyond funds required for the 50-State RECS. *Initiative 3.8. Start-up Cost: $1,000,000; Cost per 4-year Cycle: $15,660,000 (incremental cost over EIA's projected cost for 50-State RECS).*

- *Use Current Population Survey to Collect State-Level Residential Transportation Data*. Some State-level residential transportation data could be collected by adding questions to the Current Population Survey (CPS), such as the vehicle identification number and the current odometer reading for all vehicles held by members of the sampled household. The CPS, which is conducted by the Bureau of the Census for the Bureau of Labor Statistics, is a large monthly survey that would yield about 15,000 interviews per month. Estimates would permit State-level estimates of fuel use by month for large States and 6- or 12-month averages for smaller States. The data could potentially be combined with other social variables, such as household characteristics, employment, and income. *Initiative 3.9. Start-up Cost to EIA: up to $4,000,000 (over 3 years); Cost per Annual Cycle: up to $900,000 (incremental cost relative to EIA's 2009 budget request).*

- *Provide CBECS Data for All 50 States (maintaining current 4-year cycle).* Increasing the CBECS sample would allow EIA to publish State estimates for some major building types and dramatically increase the number of types that could be published at the national, regional, and division level. A larger sample would also allow EIA to undertake energy efficiency analysis; calculate changes in carbon emissions over time; and monitor the adoption of new building design, equipment technologies, and energy management tools and practices. *Initiative 3.10. Start-up Cost: up to $13,000,000; Cost per 4-year Cycle: up to $42,640,000 (incremental cost over EIA's projected cost for 2011 CBECS)*.

Increase Frequency of Energy End-Use Data

- *Collect End-Use Data More Frequently (maintaining current sample sizes).* Note that per-cycle costs are roughly the same regardless of whether the cycle is 2, 3 or 4 years but shorting the current cycle time would require additional resources on a per-year basis.

  — *Conduct the MECS biennially*. With the current 4-year cycle, EIA cannot accurately describe how energy consumption in manufacturing relates to changing energy market conditions, to the cost and availability of capital for investment in new technologies and energy management practices, and to structural shifts in demand for its products. Because energy intensities and fuel-switching capacities vary considerably between industry classes, so would their response capacity. Features such as the lag time between a market signal and demand response, whether it is temporary or permanent, and the relative impact by industry sector are not measureable or are missed by infrequent data collection. Enormous changes have occurred in the structure of American industry, where labor and the supply and cost of energy are key factors in global competitiveness. There is a significant value to energy policymakers,

the Bureau of Economic Analysis, the Federal Reserve Board, and the Bureau of Labor Statistics in having a clearer understanding of the flow of energy, capital and labor. MECS is conducted primarily via the Internet, so the marginal burden to respondents and cost to EIA of updating this series biennially would be minimized. A more frequent MECS could integrate policy-relevant topics into the survey in a more timely manner. *Initiative 3.11. Cost per 2-year Cycle: $3,987,700.*

— *Conduct CBECS, RECS, and RTECS triennially.* Adhering to a 3-year cycle would ensure that EIA could best account for factors that are important in forecasts of national consumption, and in identifying trends and structural shifts caused by changes in policy, technology, and behavior. EIA would realize some efficiencies in survey management with more frequent data collection, more staff in more specialized roles, and new resources to improve data, management and release processes. *Initiative 3.12 (CBECS). Cost per 3-year Cycle: $13,763,000. Initiative 3.13 (RECS): Cost per 3-year Cycle: $9,551,000. Initiative 3.14 (RTECS): Cost per 3-year Cycle: $4,740,000.*

Increase Timeliness of End-Use Data Releases

- *Improve Timeliness of Data Releases.* To increase the availability of end-use data, EIA would need to reduce the time lag between data collection and data release in part by making better use of the Internet for data release. In the past 10 years, three major computing changes occurred while budget (in real terms) and staff resources declined: EIA migrated from a centrally controlled mainframe to a distributed local-area network processing environment, data collection moved from paper-administered forms to computer-assisted interviews or Internet data collection instruments, and reports moved from physical publications to electronic media on the Web. The processing environment requires more skilled coordination, management, and documentation. The data collection technology requires more pre-survey preparation, which backs into the previous survey cycle. Web publications require ongoing, specialized support functions to meet user needs and expectations. Although EIA has begun work to improve performance on all these fronts, additional resources could help attain consistency and efficiency across the consumption surveys. *Initiative 3.15. Start-up Cost: $500,000; Cost per 2, 3 or 4-year Cycle: $800,000 (increment over EIA's 2009 budget request).*

Conduct Feasibility Studies for Alternative Sources of End-Use Data

- *Feasibility Studies for Alternative Sources of Data for End-Use Estimates.* Evaluate alternative Federal, State, and commercial data sources for meeting the statutory requirements of EIA's end-use program under budgetary constraints. Identify other surveys and administrative records that may prove valuable for benchmarking, modeling, or filling data gaps in the program. Evaluate the methods used and quality produced relative to EIA's needs. Identify the potential for interagency data and cost sharing, collaborative data collections, and value-added analyses to meet the challenges of scarce Federal resources, economic disruptions from the global recession, climate change, and health or national security events. *Initiative 3.16. Annual Operating Cost: up to $500,000 (increment over EIA's 2009 budget request).*

Additional Challenges to Improving Energy Consumption Data

Where there are gaps in the scope, frequency, timing, and/or publication of EIA consumption data, users increasingly resort to *ad hoc* means to address them. For example, EPA is assisting businesses, user

groups, and trade associations in collecting their own consumption data to augment the CBECS program. In other situations, States are independently collecting data to produce more localized benchmarks. These efforts are vulnerable to funding cuts, which would likely cause quality challenges similar to or worse than EIA's. Furthermore, EIA is subject to OMB's statistical standards, whereas State data collections are not. Although resources vary considerably, States are relying more on smaller data collection firms and less expensive, less accurate modes of data collection.

In another example, the Department of Housing and Urban Development produces a Utility Schedule Model using RECS, the basis for calculating utility allowances for various Federal programs. Because RECS is conducted infrequently, users are left to make their own idiosyncratic or no adjustments to the model for non-RECS years. Adjusting a model in rapidly changing energy markets is beyond the scope, resources and expertise of most of these users. As a result, EIA analysts receive direct requests from metropolitan housing authorities (or their consultants) for inter-survey estimates for geographic areas smaller than in the current RECS. Trade associations for commercial buildings and for manufacturing firms that are trying to work around Federal data gaps also request *ad hoc* advice and support. Such requests suggest that the methods used to define and evaluate efficiency and other energy program targets are idiosyncratic and will diverge until EIA can provide more and better consumption data for smaller geographical divisions and analytic subclasses.

Stakeholders and other data users report that the demand for high-quality State-level data will continue to grow. EIA terms, definitions, and survey methods then become the *de facto* basis for benchmarking and assessing energy policies led or enacted by States, as well as by other Federal statistical agencies. Improving the consumption survey program would add considerable value and coherence to data that are central to policy and other decision makers.

Chapter 4. Supplier Surveys: Data Quality and Scope

EIA's supplier surveys cover all stages of the energy supply picture. For crude oil and petroleum products, the program covers reserves, production, transportation, refining, blending, imports, exports, and storage. Much of the data collected is reported at the national, regional, and State levels. For natural gas, EIA collects and publishes data on the operation of natural gas markets at the national and State levels to provide information about natural gas supplies (from domestic and foreign sources), natural gas movement and storage, natural gas consumption by major end-use sectors, and prices along the path from wellhead to end user. EIA surveys cover the electric power industry by collecting data at the individual facility level and made available at the national, State, sector, and facility level (subject to confidentiality requirements) and are designed to capture data from traditional utilities as well as emerging participants (e.g., power marketers, and other entities engaged in the production, sales, or distribution of electricity). EIA also collects extensive data on the coal industry and on the nuclear power and the uranium mining and milling industries. EIA collects and disseminates a limited amount of data from the manufacturers of solar thermal collectors, solar photovoltaic cells and modules, and geothermal heat pump equipment. These data along with information collected by EIA on its electric power surveys and the MECS provide statistics on the growth in renewable energy in the United States, assess the effectiveness of incentive programs, and provide a basis for projections of renewable energy consumption. Finally, EIA collects and disseminates data on the production of alternative-fueled vehicles and hybrid-electric vehicles by suppliers, and on the use and fuel consumption of alternative-fueled vehicles by Federal, State, and fuel- provider fleets, and local transit companies.

Many stakeholders and customers have said that EIA supplier survey data are important for State-level policymaking, as well as for responding to emergencies and understanding markets within and affecting their States. Stakeholders and customers emphasize the importance of EIA's ability to maintain quality in its data. They have pointed out room for improvement in EIA's data, mentioning issues such as occasional discrepancies between weekly and monthly data and gaps in what EIA is currently able to collect and publish. Taking the feedback received from stakeholders over the years, as well as from workshop participants and others during our 2008 outreach efforts, EIA has developed several data quality options aimed at reducing statistical error in EIA surveys. Statistical error is inherent in all survey data, regardless of their source and the care and competence of data collectors. There are various potential sources of statistical error, such as the following:

- **Sampling.** A data collection may be based on a sample of the population rather than on a complete enumeration. Variation occurs by chance because a particular sample is surveyed. This variation decreases with increased sample size. A preferred sample size is determined in order to keep key statistics under a specified threshold of sampling error, but sometimes the actual sample size is solely determined by cost.

- **Coverage.** The frame (the list of those surveyed) may not be complete or may contain companies that are out of scope or no longer in business. Complete and up-to-date frames are important in reducing statistical error due to lack of coverage in EIA surveys. Frame improvements are particularly important in the deregulated electricity sector, which has undergone a tremendous number of mergers, acquisitions, divestitures, and company name changes over the past several years.

- **Target population.** Reporting thresholds can be used to determine a survey frame and thereby exclude smaller companies from reporting. Exclusions result in a frame that does not completely

cover the target population of companies of all sizes. Also, a frame may be incomplete and not cover the parts of the target population that are more difficult or more expensive to reach.

- **Respondent.** Respondents may commit errors in reporting their data. Respondents frequently report statistics in units other than those requested by EIA. For example, respondents will report data in gallons rather than in barrels.

- **Nonresponse.** Not all of the units that are surveyed respond (unit non-response), and some respondents may not provide all of the information requested (item non-response).

- **Processing.** Errors may occur from transcribing data incorrectly.

- **Concept.** The data collection elements may not measure the items that they were intended to measure, or the estimation methodology may provide inaccurate results.

Most of the options discussed below focus on the first three sources of survey error identified above, either by increasing sample sizes to provide better statistics or by improving survey coverage by updating frames. Other quality initiatives include improving the scope of several EIA surveys by increasing the sample size to obtain State-level estimates for all 50 States and better estimates for current statistics, or by adding new surveys. Current statistical data edits, nonresponse follow-up methods, and other internal EIA quality checks are covered by strict adherence to EIA statistical standards and control the last four sources of errors; therefore, no initiatives are proposed for any of these sources of error at this time.

Upgrade Frames

A comprehensive frame is essential to conducting a statistically valid survey. Frames-related initiatives include:

- *Develop Frame for Weekly Retail On-Highway Diesel Price Survey (EIA-888).* Construct an outlet frame of truck stops and service stations to improve the quality of EIA's weekly diesel price data. EIA is currently unable to provide average prices for more detailed geographic areas than national, Petroleum Administration for Defense (PAD) Districts, sub-PAD regions, and California. *Initiative 4.1. Start-up Cost: $250,000; Annual Operating Cost: $50,000 (increment to EIA's 2009 budget request).*

- *Update Frame for Weekly Retail Gasoline Price Survey (EIA-878)*. The frame of motor gasoline outlets was last updated in 2001. Currently, EIA provides select State and city prices in addition to national, PADD, and sub-PADD average prices. The updated frame could be used to select a new, larger sample that could produce State-level data for all States. *Initiative 4.2. Start-up Cost: $260,000; Annual Operating Cost: $50,000 (increment to EIA's 2009 budget request).*

Expand Scope

Additional initiatives would expand the number of data elements on current EIA survey forms or would increase the sample size in order to obtain more accurate State-level data[7]. Others would increase the

[7] Expanding sample sizes to include some or all U.S. territories would require additional funding for each initiative.

sample size in order to reach additional States. Still others propose new surveys to collect information on new energy data elements, including renewable energy. This group of initiatives includes the following:

- *Expand Natural Gas Production Survey (EIA-914)* to collect data from all natural gas-producing States and to add crude oil and lease condensate production from all oil-producing States. Expanding the survey would promote improved and timelier data on production trends nationally and particularly for a number of States where oil and gas production is rapidly changing. *Initiative 4.3. Start-up Cost: $400,000; Annual Operating Cost: $325,000 (increment to EIA's 2009 budget request).*

- *Expand Scope of the Monthly Gas Marketer Survey (EIA-910)* to obtain data on sales by marketers for more States. The current scope of the survey includes only the 12 States that offer natural gas retail choice programs with the highest concentration of natural gas sold by marketers. Currently, there are 22 States with residential sales by marketers and 49 States with commercial and industrial sales by marketers. *Initiative 4.4. Start-up Cost: $600,000; Annual Operating Cost: $500,000 (incremental to EIA's 2009 budget request).*

- *Expand Power Plant Operations Survey (EIA-923)* to collect fossil-fuel receipts from smaller electric plants (those with capacities between 1 megawatt and 50 megawatts). An expanded survey would complete the picture of the electric power industry and match the reporting threshold by which fuel consumption and electricity generation data are collected. *Initiative 4.5. Start-up Cost: $50,000; Annual Operating Cost: $65,000 (incremental to EIA's 2009 budget request).*

- *Expand Winter Heating Fuels Telephone Survey (EIA-877)* from 24 to all 50 States and from its current data collection period of October through mid-March to a year-round survey. This survey collects weekly data on retail prices of No. 2 heating oil and propane. These data are used to assess hardships experienced by heating oil and propane users during periods of critical short supplies. The current survey is a cooperative data collection effort between EIA and 24 States. (New initiatives for joint data collection are discussed in Chapter 5.) *Initiative 4.6 Start-up Cost: $750,000; Annual Operating Cost: $750,000 (incremental to EIA's 2009 budget request).*

- *Develop and Launch Weekly State Inventory Survey for Selected Petroleum Products* similar to EIA's current monthly collection of inventory data for selected petroleum products such as distillate fuel oil, motor gasoline, kerosene, residual fuel oil, and propane. EIA would use the weekly data to better monitor any pending supply issues. *Initiative 4.7. Start-up Cost: $1,000,000; Annual Operating Cost: $500,000 (incremental to EIA's 2009 budget request).*

- *Develop and Launch New Survey to Collect Biomass Energy Consumption* using a frame developed by EPA of all forest-product facilities. This survey would permit EIA to publish State-level biomass data. *Initiative 4.8. Start-up Cost: $190,000; Annual Operating Cost: $105,000 (incremental to EIA's 2009 budget request).*

- *Add Methane Data Elements to Existing Survey (EIA-895A), "Annual Quantity and Value of Natural Gas Production Report,"* which currently collects vented and flared natural gas as a single value, despite the fact that vented natural gas is a methane emission and flared natural gas results in carbon dioxide emissions. The distinction is important because methane is a more potent greenhouse gas than carbon dioxide. This initiative would require respondents to provide separate values for vented and flared gas and would lead to improved accuracy of State-level

greenhouse gas emissions. *Initiative 4.9 Start-up Cost: $70,000; Annual Operating Cost: $70,000 (incremental to EIA's 2009 budget request).*

Chapter 5. Communication, Collaboration, and Outreach

One of the requirements of Section 805 is to establish guidelines to "ensure the quality, comparability, and scope of State energy data, including data on energy production and consumption by product and sector and renewable and alternative sources, required to provide a comprehensive, accurate energy profile at the State level. . . ." EIA's stakeholders and customers note that it takes Federal initiatives to provide data that are comparable across States. Section 805 also instructs EIA to "assess any existing gaps" in EIA's data and to consult with State officials and others on a regular basis.

As a recent independent study noted, "EIA has a wide variety of customers and a long history of collecting feedback in order to better understand the needs of the public and to improve the quality of its products."[8] Current consulting and collaborating activities include the following:

- *Consultations with Other Government Agencies*. EIA consults with other State and Federal agencies. For example, EIA consults with the Federal Energy Regulatory Commission (FERC) on a regular basis concerning new and existing survey forms, particularly in conjunction with seeking the necessary periodic approval from the Office of Management and Budget for EIA survey forms.

- *Participation with State Organizations***.** One way EIA stays informed about changes in energy markets that affect States in particular is through participation in workshops and conferences with groups such as the National Governors Association, the National Conference of State Legislatures, the National Association of Regulatory Utility Commissioners, and others. The National Association of State Energy Officials (NASEO) is a good example. EIA staff regularly attend NASEO conferences and co-sponsor conferences and workshops with NASEO every year. EIA's communications with NASEO are intended to assess changes in interest for State energy data, to identify emerging energy issues and trends among the States, and to share solutions to State energy market or distribution problems. EIA staff often provide presentations or suggest and recruit energy industry conference speakers at NASEO's request.

- *Joint Data Collection.* One of EIA's best known collaborations with the States is the State Heating Oil and Propane Program (SHOPP), a cooperative data collection program between State energy officials and EIA that operates during the winter months of October through March. EIA provides funding to the States through grants to cover 50 percent of their costs for participating in the weekly heating oil and propane price survey. EIA selects the sample for each State, edits and verifies the incoming data, aggregates the data, and disseminates the data on the EIA Web site. The States make weekly phone calls to the companies in their sample and transmit the prices to EIA via an Internet data collection system. SHOPP is important in terms of outreach. It provides EIA with a contact person in each State Energy Office, enabling a communication network with the participating State Energy Offices. The cooperative program promotes the States' ability to monitor market conditions for heating oil and propane during the winter season.

[8] *Challenges, Choices, Changes: An External Study of the Energy Information Administration*, May 2006, p. 21.

In addition to continuing activities such as those described above, EIA could expand its interaction with the States through the following initiatives, which also could involve the U.S. territories, as feasible and appropriate:

- *Develop a Plan To Evaluate and Address Emerging State and U.S. Territory Data Needs*. New data needs related to emerging State and territorial energy issues are of particular concern to them. In order to collect and publish data that are responsive to those needs and also comparable across States and territories, EIA proposes to initiate an ongoing plan. In each cycle, the first step would be to identify emerging issues and to evaluate the data needs through iterative interactions with State and territorial officials and others. The second step would be to interview EIA analysts and to research current data collections. The third and final step would be to write recommendations about the ways in which EIA could address the needs by expanding the scope and coverage of existing surveys, by adding new surveys, or by other means. EIA's recommendations to fill the gaps could help ensure that current, comparable datasets are available for State- and territorial-level energy analysis. *Initiative 5.1. Start-up Cost: $130,000; Annual Operating Cost: $65,000 (increment to EIA's 2009 budget request).*

- *Conduct Biannual State and U.S. Territory Energy Data Workshops.* EIA could conduct workshops to educate users of State and territorial energy data about current and new EIA initiatives to enhance the relevancy, scope, quality, comparability, and timeliness of EIA's data, as well as to seek feedback. The workshops would be aimed at a broad audience. Each workshop would result in a record of written and oral comments from workshop participants and proposed follow-up actions which, if funded, would address the data needs identified in the workshop. *Initiative 5.2. Start-up Cost: $65,000; Annual Operating Cost: $65,000 (increment to EIA's 2009 budget request).*

- *Consult with Other Government Agencies on State and U.S. Territory Data Needs.* EIA could regularly consult with Federal and State agencies and organizations about the relevancy, scope, quality, comparability, and timeliness of EIA's data in order to gain information about State and territorial data needs. EIA would write annual reports describing the data needs and recommending changes in EIA's data collection activities which, if funded, would meet the needs. *Initiative 5.3. Start-up Cost: $130,000; Annual Operating Cost: $130,000 (increment to EIA's 2009 budget request).*

- *Investigate Options for Supporting Independent State and U.S. Territory Data Collection.* Currently, EIA publishes a set of statistical and methodological guidelines on EIA's Web site. One option for cooperative data collection between EIA and the States and territories would be for EIA to work with them to assess their data needs, plan an approach to collecting the data, and develop detailed, individualized guidance on how to collect the targeted data. With that approach, State and Territory data collection efforts could mesh with EIA's, yielding comparable datasets. *Initiative 5.4. Start-up Cost: $130,000; Annual Operating Cost: $130,000 (increment to EIA's 2009 budget request).*

- *Establish a "Wiki" to Allow for Collaboration in a Common Workspace.* EIA understands the mutual benefits of good communications with the States and other stakeholders who are interested in State- and territory-level information, and EIA communicates with stakeholders in a variety of ways. Some electronic media currently in use are "one-way," in that EIA posts content to its Web site or sends list-serve notices about new products. EIA also uses interactive approaches, such as those that allow users to create customized graphs and tables. There are also

options for two-way electronic communications, including the use of wikis (Web sites, or parts of a Web site, using software that allows for content to be edited and revised by all users.) The use of a wiki would allow EIA's data users to collaborate in a common workspace on data projects. The main benefit of a wiki would be quick and efficient communications during collaborations and a public record of data projects. *Initiative 5.5. Start-up Cost: $50,000; Annual Operating Cost: $45,000 (increment to EIA's 2009 budget request).*

Sharing of Company-Level Data

Section 805 directs the Administrator to "share company-level data collected at the State level with each State involved, in a manner consistent with the legal authorities, confidentiality protections, and stated uses in effect at the time the data were collected, subject to the condition that the State shall agree to reasonable requirements for use of the data, as the Administrator may require. . . ." Information collected by EIA falls into three groups with respect to how it is shared:

- *Confidential* information is collected under the pledge of confidentiality pursuant to the *Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA)*.

- *Protected* information is collected under a pledge that EIA will protect information to the extent it satisfies exemption 4 of section 522(b) of the Freedom of Information Act (FIOA). Title 15 Section 796(d) of the United States Code also contains a provision that requires EIA to protect certain types of information that, if made public, would reveal methods or processes that are entitled to protection as trade secrets. Broadly stated, information is considered *protected* by EIA if the release of respondent-identifiable information would be likely to cause harm to the respondent's competitive position or impair EIA's ability to obtain similar data in the future. EIA staff and contractors are subject to the penalties (imprisonment, fines and termination) cited in Title 18 Section 1905 of the United States Code for disclosing trade secret information or confidential statistical data in any manner not authorized by law.

- *Public Information* is information that is collected with the notice to respondents that the information may be publicly released in company- or individually-identifiable form.

Respondent-level confidential and protected data may be shared under certain conditions as outlined in CIPSEA and other authorizing legislation pertinent to EIA, including the DOE Organization Act, the Federal Energy Administration Act, and the Energy Supply and Environmental Coordination Act of 1974. Requests for data sharing are reviewed on a case-by-case basis. EIA formalizes the sharing of confidential and protected data with a Data Sharing Agreement signed by EIA's Administrator. Several factors are evaluated when a request is received, including: 1) the legal authority controlling the sharing of requested data; 2) the resources involved in providing the requested data, such as if any special file formatting or re-programming is required; 3) what direct identifiers, if any, need to be removed from the data file before its release; 4) the duration of the sharing activity, i.e., is it one-time or on-going; 5) can the agency requesting the data appropriately safeguard it; and 6) will any tabular data satisfy the requesting agency's data needs or is respondent-level data necessary for its use and purpose for accessing the information.

Confidential data, or data that is collected under CIPSEA, can only be shared for statistical purposes and CIPSEA attaches stringent data protection and eligibility requirements. CIPSEA allows EIA to designate other agents, besides contractors, who may have access to identifiable information collected under CIPSEA as long as the agent's access is properly controlled by EIA. These agents, such as researchers,

function under the agency umbrella, i.e., under the supervision of EIA employees through contracts or special agreements and subject to the same limitations and penalties that are described in CIPSEA. The fact that CIPSEA allows EIA to share information collected under CIPSEA with approved agents does not create a researcher's right to such information. If EIA chooses to designate an agent, it still remains responsible for all confidential information protected under CIPSEA. EIA must approve the researcher and the researcher must agree in writing to comply with all terms and conditions required by EIA prior to any access.

Protected data can be shared for official use by other DOE components, other Federal agencies, the General Accounting Office, and any Committee of the Congress. Official use of EIA data by other Federal agencies may include both statistical and non-statistical uses. A court of competent jurisdiction may obtain protected information in response to a court order. With respect to the States, Title 15 Section 779(a)(3) of the United States Code requires the Administrator to provide, upon request, to State governments all relevant information he/she possesses concerning the status and impact of energy shortages and the extent and location of available supplies of crude oil, petroleum products, natural gas, and coal, within the distribution area serving that particular State. EIA has, in fact, worked closely with the States over many years to share relevant information in its possession concerning the status and impact of energy shortages. However, because Section 779(a)(3) explicitly states that such information is to be provided in accordance with other provisions of this Code chapter, which includes Section 796(d) cited previously, the sharing of protected information remains subject to confidentiality requirements.

As noted, existing laws also permit the sharing of protected or confidential information for statistical purposes related to the EIA mission provided that the entity seeking access to the data has the ability to protect it from disclosure. EIA's implementation of data-sharing agreements provides State entities or agencies with an avenue to access company-level data for statistical purposes.

Appendix A. Legislation

ENERGY INDEPENDENCE AND SECURITY ACT OF 2007 (P.L. 110-140)
Sec. 805. ASSESSMENT OF RESOURCES

(a) 5-Year Plan-

(1) ESTABLISHMENT- The Administrator of the Energy Information Administration (referred to in this section as the `Administrator') shall establish a 5-year plan to enhance the quality and scope of the data collection necessary to ensure the scope, accuracy, and timeliness of the information needed for efficient functioning of energy markets and related financial operations.

(2) REQUIREMENT- In establishing the plan under paragraph (1), the Administrator shall pay particular attention to—

(A) data series terminated because of budget constraints;
(B) data on demand response;
(C) timely data series of State-level information;
(D) improvements in the area of oil and gas data;
(E) improvements in data on solid byproducts from coal-based energy-producing facilities; and
(F) the ability to meet applicable deadlines under Federal law (including regulations) to provide data required by Congress.

(b) Submission to Congress- The Administrator shall submit to Congress the plan established under subsection (a), including a description of any improvements needed to enhance the ability of the Administrator to collect and process energy information in a manner consistent with the needs of energy markets.

(c) Guidelines-

(1) IN GENERAL- The Administrator shall--

(A) establish guidelines to ensure the quality, comparability, and scope of State energy data, including data on energy production and consumption by product and sector and renewable and alternative sources, required to provide a comprehensive, accurate energy profile at the State level;
(B) share company-level data collected at the State level with each State involved, in a manner consistent with the legal authorities, confidentiality protections, and Stated uses in effect at the time the data were collected, subject to the condition that the State shall agree to reasonable requirements for use of the data, as the Administrator may require;
(C) assess any existing gaps in data obtained and compiled by the Energy Information Administration; and
(D) evaluate the most cost-effective ways to address any data quality and quantity issues in conjunction with State officials.

(2) CONSULTATION- The Administrator shall consult with State officials and the Federal Energy Regulatory Commission on a regular basis in--

(A) establishing guidelines and determining the scope of State-level data under paragraph (1); and
(B) exploring ways to address data needs and serve data uses.

(d) Assessment of State Data Needs- Not later than 1 year after the date of enactment of this Act, the Administrator shall submit to Congress an assessment of State-level data needs, including a plan to address the needs.

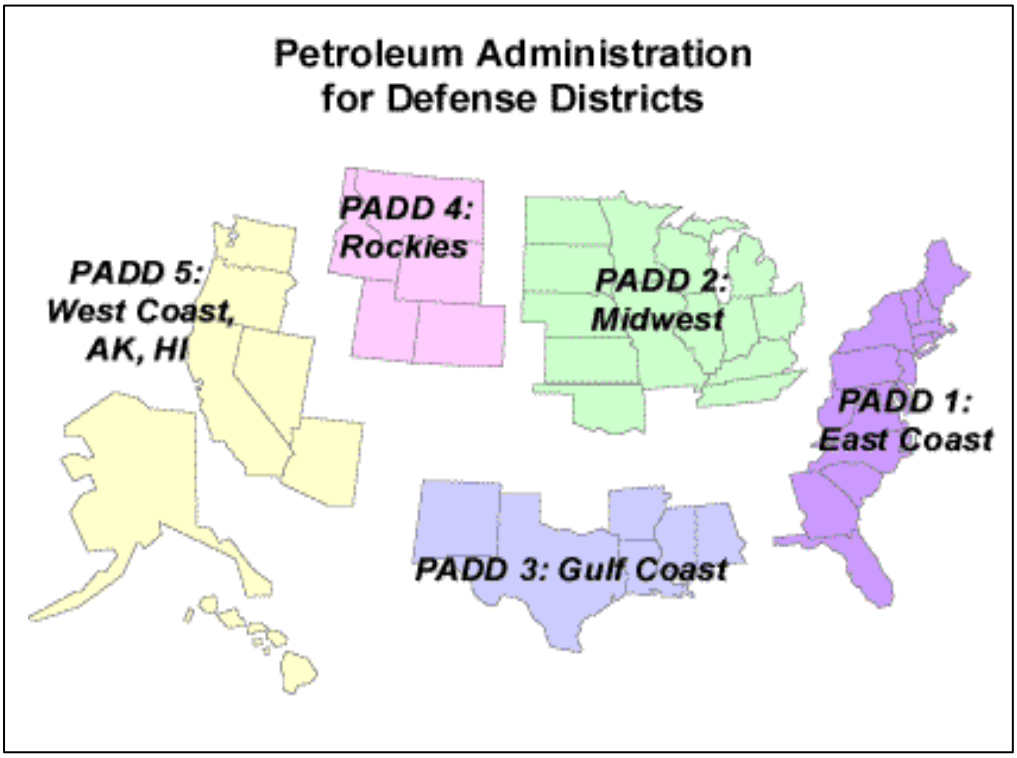
(e) Authorization of Appropriations- In addition to any other amounts made available to the Administrator, there are authorized to be appropriated to the Administrator to carry out this section--

(1) $10,000,000 for fiscal year 2008;
(2) $10,000,000 for fiscal year 2009;
(3) $10,000,000 for fiscal year 2010;
(4) $15,000,000 for fiscal year 2011;
(5) $20,000,000 for fiscal year 2012; and
(6) such sums as are necessary for subsequent fiscal years.

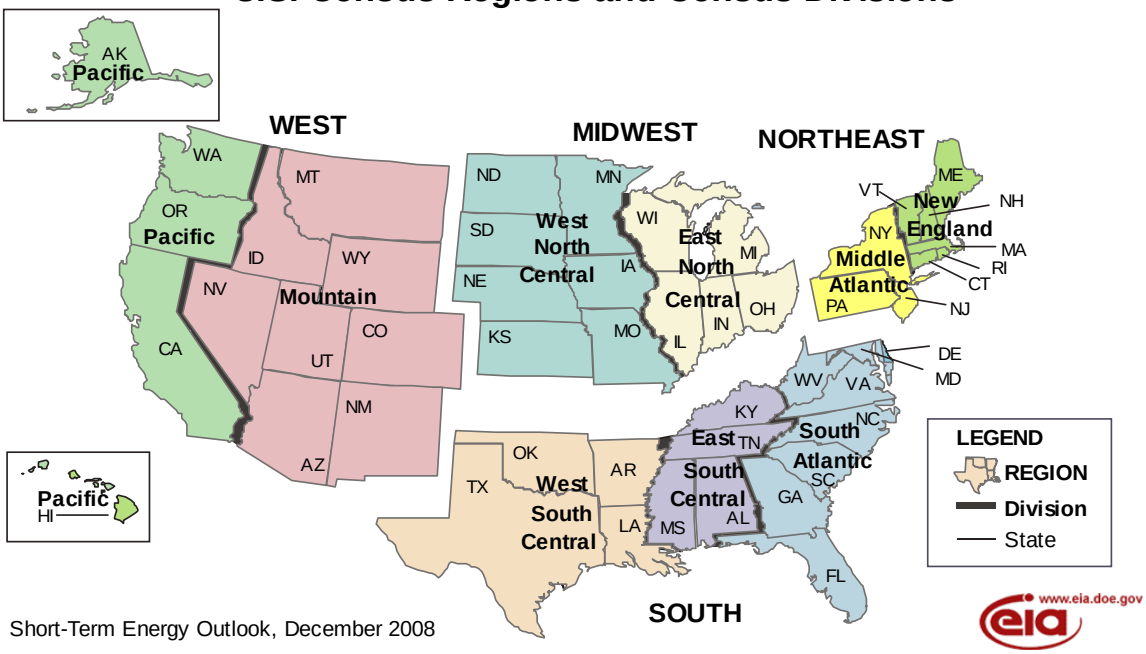
Appendix B. Geographic Reporting Levels for EIA Data

EIA currently collects and publishes data at different geographic levels depending on the type of fuel or energy source (e.g. petroleum, natural gas, coal, uranium, renewables, and electricity) and the type of data being reported, such as prices, reserves, supply, distribution, trade and marketing, consumption, sales, and end use. Only data relevant to a given State are collected and reported for that State because not all States have production plants, refineries, storage facilities, pipelines, ports, etc. Additionally, insufficient sample sizes can prohibit State-level aggregations, as discussed in Chapter 5, and some data are more easily interpreted if they are reported by regions or districts that are associated with established production, marketing, or transmission boundaries. Some data are available by both State and by region, other data are available by region alone. For example:

- Much of EIA's petroleum data are reported at the five Petroleum Administration for Defense Districts (PADD) in addition to U.S and State levels (see map below).
- EIA's Energy Consumption Surveys are reported primarily at the Census Divisions and Regions level (see map below).
- EIA's *Electric Power Annual* reports summer and winter data at the NERC (North American Electric Reliability Council) region level (see map below).

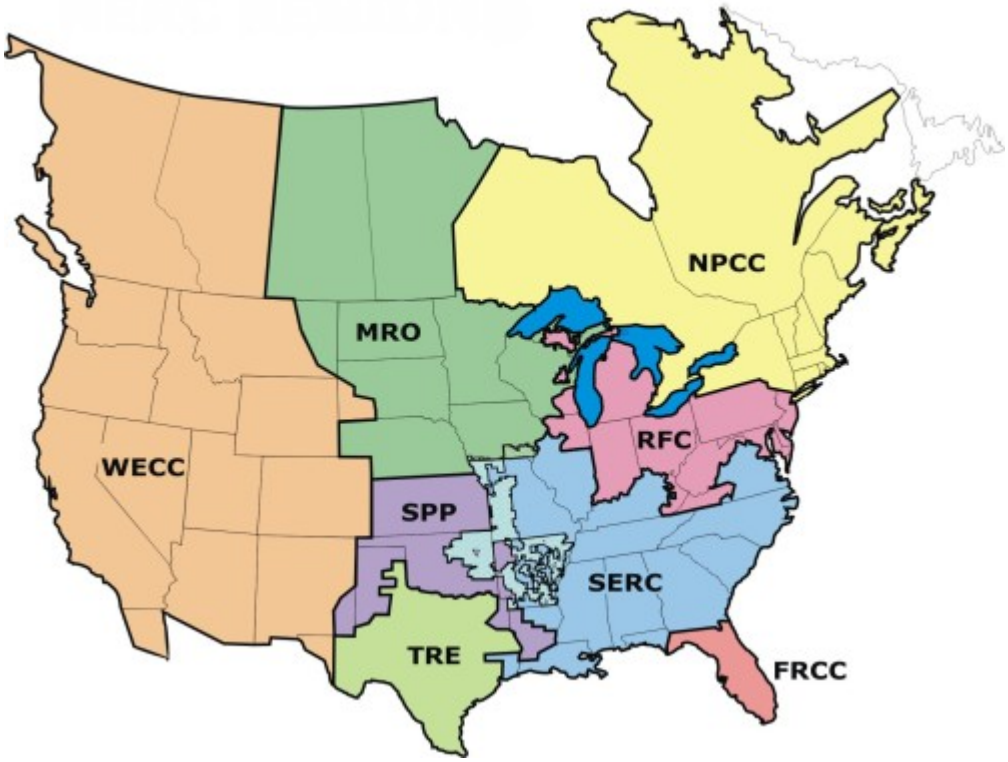


U.S. Census Regions and Census Divisions



Short-Term Energy Outlook, December 2008

North American Electric Reliability Corporation (NERC) Regions



FRCC - Florida Reliability Coordinating Council
MRO - Midwest Reliability Organization
NPCC - Northeast Power Coordinating Council
RFC - Reliability*First* Corporation
SERC - SERC Reliability Corporation
SPP - Southwest Power Pool, RE
TRE - Texas Regional Entity
WECC - Western Electricity Coordinating Council