STATISTICAL AND POLICY ISSUES IN THE DEVELOPMENT OF A MANUFACTURING ENERGY CONSUMPTION SURVEY

Dwight K. French, Energy Information Administration

The Energy Information Administration (EIA), as the official source of information relating to energy in the U.S. economy, collects, analyzes, and disseminates data related to all aspects of energy-source throughput. The EIA has data systems in place to measure mining, drilling, importation, storage, transmission/delivery, refining, generation, sales, and end-use consumption of energy sources.

The EIA has divided the population of end-use consumers into five major sectors: residential households, residential transportation, commercial, industry, and nonresidential transportation. For each of these sectors, EIA has taken, is taking, and/or will take the following approach to meeting end-use consumption data needs: identify uses and needs for energy consumption and related information, investigate existing sources of data to see whether they satisfy these uses and needs, and where necessary, design and implement data systems to fill unmet needs. Data systems are currently in place for the residential household, residential transportation, and commercial sectors. The two sectors not currently covered are industry and nonresidential transportation.

The EIA is now developing a consumption survey for the manufacturing industries, which account for the majority of industrial energy consumption. The survey is being designed as a probability sample of the population of manufacturing establishments in the 50 States and the District of Columbia, using as an immediate frame the establishment list for the Annual Survey of Manufactures (ASM) conducted by the Bureau of the Census. The population of establishments consists of all those with primary activity in Standard Industrial Classification (SIC) categories 20 through 39. For purposes of sampling efficiency, the ASM sample will be stratified by SIC, geography, and establishment size. The manufacturing survey subsample will be drawn with largest probabilities given to the most intensive energy-use groups.

The purpose of this type of sampling is twofold; to minimize the level of error in national statistics, and to assure sufficient observations in important energy consuming industries to produce reasonably precise estimates for them. Data from other survey sources indicate that national consumption estimates for major energy sources with error levels comparable to those for other consumption data systems (4 to 6 percent relative standard error) could be derived from a carefully stratified sample of 10,000 responding establishments. Oversampling to allow for out of scope, out of business, and nonresponse cases leads to the current planned sample size of 12,000 establishments. The survey will be conducted by the Census Bureau as a followup to the ASM, and energy data from the manufacturing survey will be linked with economic data from the ASM to eliminate duplicative data collection and minimize respondent burden. A pilot study to test the MECS questionnaire and instructions is planned for later this year. Fieldwork for the national survey is now scheduled to begin in late spring of 1986.

The Manufacturing Energy Consumption Survey (MECS) has been the object of some controversy, with attention being brought to both its statistical and its policy implications. The purpose of this paper is to discuss the development of the MECS and address some of the statistical and policy issues that have been raised in assessing its value to the Federal statistical system.

STATISTICAL ISSUES: HOW THE SURVEY WAS DEVELOPED

Underlying Principles

The first thing to remember about the MECS survey is that it is intended to be a baseline survey, and is being designed to combine with other data systems into a coordinated end-use consumption data base. As a baseline statistical survey, the MECS is being developed according to a few standard planning assumptions that are common to most Federal survey efforts of this type:

(1) The survey will cover data issues of long-term, continuing interest to both the Government and the general public. That is not to say that other specific issues cannot be addressed in any particular cycle(s), but there will be a core of energy consumption and related information that forms the "backbone" of the data system.

(2) The survey will cover a comprehensive, well-defined population, with as few exceptions as possible for convenience of the survey operation.

(3) The survey will be conducted according to sound, proven statistical data-collection techniques, with proper attention given to estimation; variance

estimation; minimizing of, and adjusting for, nonresponse; data collection and data-processing techniques; and proper documentation of survey methods and sources of error.

Identifying Data Issues

In developing the data issues to be covered in the MECS, EIA surveyed data users both inside and outside Government and identified a wide variety of topics that were of varying degrees of interest. These include:

o Consumption and expenditures for energy sources.

o Breakdown of consumption by fuel use versus feedstock/-raw-material use.

o Consumption of fuels, broken out by end use.

o Conservation measures, including future plans for conservation investment.

o Energy saved through conservation.

o Inventories and storage capacity for energy sources.

o Descriptive information about individual pieces of energy-consuming equipment.

o Economic characteristics related to consumption, such as employment, value of production/value added, hours of operation, and capacity utilization.

o Onsite generation/-cogeneration of electricity, and the input fuels used.

o Onsite generation of steam and the input fuels used.

o Sources of energy inputs--purchases, transfers, onsite production.

o Energy outputs--sales, other transfers.

o Fuel substitution capability.

o Fuel substitution behavior.

o Building counts and square footage, and description of activities.

The EIA then went to representatives of individual industries and trade associations, meeting with them to talk about the way industry does its energy accounting, what kind of information industry can and cannot provide, and who fills out question- naires, when, and how. In addition, EIA staff visited the sites of more than 20 industrial plants in a variety of SIC groupings, employment sizes, and geographic locations, in an attempt to understand how various types of manufacturing processes work, how

companies operate their plants, and how the data issues listed above apply to these real-life situations.

The results were eye-opening, to say the least. The discussions and site visits showed that several of the issues mentioned above would be onerously burdensome to address for the purposes of a baseline statistical survey, or that meaningful data could not even be provided. Issues in these categories included conservation measures, equipment-specific information, consumption by end use, capacity utilization, decisionmaking about fuel substitution, and even building information. (One site we visited had so many buildings that the establishment energy manager could only estimate the number within about 300! And the number of buildings on the site literally changed from week to week.)

After EIA had finished these discussions and site visits, and had gone through its own internal issue discussions, the prospective questionnaire content had been reduced to the following core items, all of which had been stated by energy analysts both inside and outside Government as being vital to presenting a comprehensive overview of manufacturing energy use, and all of which looked to be feasible to collect:

o Consumption and expenditures for purchased fuels.

o Consumption of energy sources as feedstocks and raw materials.

o Inventories and storage capacities for appropriate energy sources.

o Transfers and onsite production of energy sources as waste or byproduct.

o Onsite generation/cogeneration of electricity, and sales back to the utility grid.

o Fuel switching capability.

The first major problem in developing a questionnaire from this list of issues was to merge these somewhat related topics into a single well-organized data collection. The method that has been chosen is to request data on specific consumption-related quantities as required to satisfy data needs, and to provide in the instructions two separate methods of energy flow accounting. Either or both of these methods can be used by respondents as appropriate to derive the quantities for which direct measurements or estimates are not

caused by omitting or misallocating inputs and/or dispositions.

Measuring Fuel Switching

The second major problem was defining and measuring fuel switching. In discussing the issue with data users and data providers, EIA never seemed to encounter two people who thought of fuel switching in exactly the same way. Discussions covered short-term switching capability, long-term capability, future switching strategy/intentions, investment in equipment with multiple fuel capabilities, prevalence of multiple-fuel equipment, prevalence and utility of redundant equipment (different pieces of equipment that can serve the same purposes with different fuels), emergency switching capability, business-as-usual switching capability, the relative costs of energy sources needed to trigger business-as-usual switching, switching by changing product sites for particular products, switching by changing product mix, practical limitations to switching capability, time limitations on switching due to maintenance requirements on equipment, temporary versus permanent (one-way) fuel-switching, installed versus potential switching capability, mandated versus voluntary fuel switching, historical switching capability, present switching capability, and whether or not the respondent has switched during a reference period. Out of this myriad of topics, EIA needed to decide upon a way to collect data on fuel switching that would fit in with the MECS questionnaire.

The approach that has been selected for pilot testing concentrates on estimating short-term switching capability, in no more than a 30-day period and with no significant capital expenditures. This approach asks the respondent to estimate switching based on the previous year's fuel consumption and production schedule, and narrows the capability being measured to swtiching to and from oil, the issue conceived to be greatest political concern and economic impact. Also, the approach chosen attempts to distinguish between maximum potential capability to switch and practical capability, taking into account legal, environmental and regulatory constraints.

Integrating MECS With Other End-Use Data

One final statistical problem that EIA is now working on is the

readily available from company records. These methods can help respondents to understand exactly what data are being requested, and how these quantities fit into energy throughput at an establishment. All accounting and data reporting for the MECS will use calendar year 1985 as a reference period.

The first method is direct "equipment/process" accounting, which requires that the respondent have records or estimates of 1985 consumption in individual manufacturing processes and pieces of energy-consuming equipment. This method of accounting, when feasible, would allow a respondent to estimate fuel and/or feedstock consumption for an energy source as a sum of its consumption at various stages of the manufacturing process. Much equipment and many processes are not individually metered by manufacturers. Therefore, this method could be expected to be applicable only for certain energy sources in certain closely-monitored establishments.

The second method is an "input/output" approach. Starting with beginning-of-year inventory, the accounting requires data on purchases, transfers in, onsite production and generation. Then feedstock use, sales, transfers out and end-of-year inventory are subtracted out. Any quantity remaining represents fuel consumption during 1985.

Both methods attempt to lead the respondent into answering questions about energy consumption in the more general context of energy throughput in the establishment. The approach is similar to the components approach commonly used to measure concepts such as income, and presumably generates a more accurate estimate than a direct request for a consumption figure, even if the concept of consumption is rigorously defined.

This approach not only unifies all the consumption-related data issues, but also provides a format for estimating consumption by accounting for all the components that contribute to it. This approach to estimation is similar to the components approach commonly used to estimate income, and presumably generates a more accurate estimate than a direct request for consumption. With the components approach, the respondent is forced to quantify energy flow and derive consumption, rather than give a single direct answer to a consumption question. The approach minimizes bias

integration of MECS data with data for other end-use consumption sectors. the fact that EIA has defined five end-use sectors does not mean that there is clean distinction between them. Far from it. Commercial activities, and even occasionally some industrial activities, take place in otherwise residential housing units. Household vehicles are used for both personal and business purposes. Buildings commonly house a combination of residential, commercial, and/or industrial activity. Most large industrial establishments contain offices for administrative staff, warehouses, and/or research laboratories. Such buildings, if they stood alone and were the primary activity of an establishment, would be classified as commercial. Large industrial establishments often have onsite fuel storage for shipping vehicles used off site, which can be considered part of nonresidential transportation. Trying to establish meaningful sectoral delineations using suppliers' data on customers is also a frustrating exercise, because suppliers (understandably) classify customers for their own convenience. Such classifications do not lead to consistent definitions of what constitutes a residential, commercial, or industrial customer. In fact, that inconsistency is a major drawback to sectoral "consumption" data in supply-based publications such as EIA's own "Monthly Energy Review."

The MECS design must thus be considered in the context of designs for other sectoral consumption surveys, most notably the Nonresidential Buildings Energy Consumption Survey (NBECS), also conducted by EIA. The NBECS covers the nonresidential buildings stock, and uses the building as its unit of sampling, data collection, and analysis. However, it is designed and conducted primarily as a survey of the commercial sector. Even though industrial buildings have been considered in scope for the two surveys conducted thus far (in 1979-80 and in 1983), estimates of industrial building counts and energy consumption have been poor to worthless because of the inherent neglect of industry in the sample design process. The EIA originally had included some questions about buildings in the MECS questionnaire, with the idea of completing the nonresidential buildings stock estimate via the MECS, and eliminating industrial buildings from future

cycles of the NBECS. However, there was never any real thought of collecting energy-consumption data allocated to buildings in MECS, and the building-characteristics data have now been dropped from the MECS. The EIA is, therefore, left with establishment-level data collection in MECS, which will cover energy consumption in manufacturing, but not building information. Establishment-level data can suffice as summary data for the manufacturing industries, and restructuring NBECS to get reliable industrial building data would be very expensive. Therefore, EIA's present plan is to keep MECS and NBECS separate, and perhaps eliminate any duplication between the two systems by screening industrial buildings out of the NBECS building interview process. That, of course, will lead to definitional problems, like what to do with mixed-use buildings that are partially industrial and how to screen out commercial buildings that are part of industrial complexes.

On the other hand, it is possible that the MECS could make a significant contribution to providing information about nonresidential transportation, a sector not currently addressed through end-use consumption data collection. The present MECS questionnaire identifies offsite use of fuel for vehicles, primarily as an output to be removed in the process of measuring onsite energy use. However, this particular segment of nonresidential transportation energy use may not be easily collectible by any other means, so that the MECS may be useful as a data source for this sector as well.

POLICY ISSUES: THE WHY BEHIND THE HOW
Industry Concerns

The EIA has gone through a process of explanation and justification of the MECS survey program to industry representatives that has been every bit as exhaustive and time consuming as its statistical development. This process has focused on three questions repeatedly raised by industry, namely:

(1) How can confidentiality of the data be assured?

(2) Can the data be put to any useful purposes that justify the burden of data collection?

(3) Do other data collections already provide the data EIA needs?

There are several reasons why manufacturers, as potential data providers, press for answers to these questions:

o Competitive
concerns--Manufacturers are in
direct, often high-stakes
competition, which makes them
leery of releasing any informa-
tion that might do them competi-
tive harm.

o Unease about data protection--
Manufacturers have been subjected
to past data collections,
especially regulatory
collections, that have forced
them to reveal information that
they considered sensitive and
have it put into the public
domain. These concerns are
heightened by the fact that EIA
has no confidentiality
legislation.

o Visibility as a sampling
unit--Even though the population
of manufacturers is rather large
(in excess of 300,000
establishments), list of parts or
all of the establishment
population are readily available
in the public domain. Thus, they
are a highly visible target of
data collectors.

o Quantity and variety of data
requests--Manufacturers are
approached repeatedly for
information about many of the
aspects of their operations.
This is particularly true of
large establishments. The top 10
percent of manufacturing .
establishments account for a
large majority of employment,
value of production, and energy
consumption, and are naturally
high-probability units in any
sample selection.

o Corporate burden--In
manufacturing, an additional
layer of individual burden is
created because large
corporations generally want
knowledge of and review rights
for data provided by their
establishments. Individual site
burden is thus multiplied in
corporate eyes.

o Previous bad experiences--In
1980, EIA obtained OMB approval
for, and subsequently fielded,
Form EIA-463, a survey of large
combustors in the manufacturing
industries. The survey, designed
for regulatory as well as statis-
tical purposes, collected both
detailed data on establishments
and (for establishments with
large combustors) detailed data
on individual pieces of equip-
ment. As a result of industry

protests, OMB Director David
Stockman withdrew data-collection
authority for the survey in
February of 1981, during the
fieldwork period, citing exces-
sive burden to collect unneces-
sary information.

In responding to these three main
issues, EIA has had an opportunity to
consider some searching questions
about the makeup of the Federal
statistical system, EIA's place in
that system, and the responsibility of
the data-collection community to its
data providers and vice versa.

Confidentiality

The issue of confidentiality has
been the easiest of the three issues
to deal with directly. The EIA does
not have its own confidentiality
legislation; in fact, some of its
supply data systems are specifically
designed to provide publicly available
information about individual
establishments in federally licensed
and/or regulated activities, such as
uranium processing and electricity
generation. End-use consumption
surveys do not ordinarily serve as
vehicles for regulation, and thus,
there is no need to identify data for
individual respondents. However, to
protect respondents against access to
their data by interested parties, EIA
has had a policy of collecting survey
data using only reputable survey
research firms with their own
confidentiality policies, and never
taking possession of any data that
could identify individual respondents.
In order to assure confidentiality,
contractors not only must remove names
and addresses from data files but must
also take the familiar steps of data
suppression, error inoculation or
other random adjustments, truncation,
recoding to categories, or replacement
of individual items with category
means, to avoid deductive disclosure.
The EIA has assumed responsibility for
overseeing data tabulations, to avoid
disclosure of information in cells
with small numbers of observations.
Table disclosure is usually not a
large problem, because sampling errors
usually force data suppression before
disclosure becomes a problem.

However, all the disclosure
protection in the world would not
prevent interested parties from
forcing disclosure of sensitive data
by legal means. There is one obvious
solution to that problem: remove any
legal means. One direct way of doing
that is to have the data collected by
the Bureau of the Census under its

Title 13 confidentiality provisions. In fact, the Census Bureau will serve as the data-collection agent for the MECS next year, collecting the data as an adjunct to its ASM. There are other good reasons for using the Bureau (the presence of the ASM sample as a frame for MECS subsampling, the Bureau's experience in industrial data collection, and the linking of ASM economic data to MECS energy data). However, the confidentiality issue is probably the single most compelling reason.

Worth Versus Burden

The second issue, the worth of the data being collected versus the burden it places on respondents, is much more difficult for the EIA to address, because it implies assessment of costs and associated benefits of data collection. A formal cost/benefit analysis is ordinarily impossible to perform with data-collection efforts, because the usual benefits are impossible to quantify. Questions that are asked (by OMB, among others) as a surrogate for benefit are:

o   Is the survey mandated by law or regulation?

o   Is the survey of direct benefit in formulating or carrying out Government policy?

o   Does the survey provide data needed for analysis or monitoring, or other indirect benefits that cannot be obtained from another source?

The MECS is being proposed as a necessary baseline statistical data set for the manufacturing industries. It will complement EIA's consumption data information in other end-use sectors, and will serve as input for EIA's industrial modeling and forecasting programs. The statement of the congressional conferees on the 1985 EIA appropriation identified $1.1 million specifically for this effort. Thus, attention focused not on whether to collect data, but on what data were required to satisfy the congressional mandate and had uses that were significant enough to justify the burden.

The MECS will provide important input to policymaking and program analysis in such areas as natural-gas deregulation, the effects of industrial conservation, and electric-utility construction and licensing. It will provide answers to questions asked within DOE during the past few years about tertiary fuel-storage

capacity and capability to switch to alternate fuels in times of crisis, two issues on which speculative and possibly misleading data are currently a part of official Government policy documents. It will provide data to assess the emergence of onsite generation and cogeneration of electricity, and electrification of key industrial processes. These events will have a profound effect on the future industrial fuel use mix, and thus will affect DOE's future research and project-support activities.

In the area of analytic benefits, the MECS will be used by a wide variety of analysts inside and outside the Government, as evidenced by comments received on users' needs. The list of potential users includes EIA modelers, of course; the Bureau of Economic Analysis; and the Commerce Department staff charged with overseeing the national energy accounts; along with utility groups, market analysts, State and local governments, and even industry analysts.

Of course, these general, inside benefits of data are no consolation to data providers, who see their time spent with no corresponding benefit. The problem is compounded in the case of manufacturers, because the wide range in scope of activities among establishments causes wide variability in burden. The MECS form is anticipated to take no more than 2 or 3 hours to complete for establishments with few fuels and no switching capability, but possibly 10 times that long for large establishments in industries such as chemical processing. In collecting energy data, there is simply no way to avoid the accounting needed for establishments engaged in a wide variety of processing activities. Even if the MECS form were to simply ask for consumption rather than derive it, a respondent would have to mentally go through the accounting anyway, in order to properly answer the question. In the same way, a respondent can only provide the type of fuel-switching data requested by following the logic of the fuel-switching section.

As mentioned earlier, large corporations see reporting burden not just in terms of individual establishments, but in terms of the corporate whole. One of the side benefits of the MECS sampling procedure is that it will reduce the number of reports required of large corporations. This

in turn will reduce the corporate oversight burden, which is important, according to some corporations, because of the relatively small staff devoted to energy management at the corporate level.

Duplication, Duplication, Duplication

The issue of duplication of existing data is the most interesting of the three issues from a statistical point of view, because it raises serious questions of data-system integration and a phenomenon that might best be described as adaptive estimation--that is, making the available data apply to the issue one wants to study. This issue arises directly from the concerns about burden, and the fact that manufacturers engage in so many facets of economic activity that everyone seems to need some type of information about them. Manufacturing establishments have received a great deal of attention from energy data collectors, especially since the 1973 oil embargo.

The first complete energy data system for the manufacturing industries was a supplementary collection of data on the consumption of purchased fuels and electric energy attached to the Census Bureau's ASM. The collection, which began in 1974, was funded by EIA and was a part of the data collection for all 50,000 to 60,000 establishments in the ASM. Data were collected annually through 1981, when EIA funding of the supplement was discontinued because of budget cuts and the data collection was terminated.*

At about the same time, two other data collections were under way. The first was a one-time inventory of major fuel-burning installations (MFBI), defined as establishments that had one or more combustors with a design firing rate of at least 100 Mm Btu/hour. Conducted by the Federal Energy Administration in 1975, the survey included fuel-consumption data on these large pieces of equipment and the establishments that operated them.

The primary purpose of this survey was to identify large consumers of oil and natural gas that could be switched to coal. The second data collection was a voluntary energy-consumption submission by major corporations within several energy-intensive industry groups, worked out as a result of discussions between industry and Government representatives. This data collection was expanded to cover large corporations in all 20 two-digit manufacturing industry groups as a

result of the Energy Policy and Conservation Act of 1978, which authorized mandatory collection of energy consumption and production information from all establishments in corporations that consumed more than a trillion Btu of energy in manufacturing within a single two-digit SIC code. This data system (commonly referred to by its form number, CE-189) is not designed as a benchmark statistical data collection, but is supposed to be used to indicate the progress that large energy consumers have made in reaching energy-conservation targets set for 1980. Data are collected from individual establishments, then aggregated to the corporate and (usually) trade-association level, before being submitted to the Department of Energy. The statistics that are produced from this annual data collection are national consumption estimates by energy sources for the corporations reporting for each two-digit SIC category and a single national conservation index for each two-digit category. Certain industry groups voluntarily compile annual data in more detail, but not ordinarily geographic detail. DOE is currently proposing elimination of this data collection in favor of MECS.

In 1980, EIA undertook a new survey of all manufacturing establishments having combustors with design firing rate of at least 50 Mm Btu. The Manufacturing Industries Energy Consumption Study and Survey of Large Combustors (Form EIA-463) was designed to collect consumption of both purchased and nonpurchased energy sources at the establishment and combustor level, together with descriptive data on the establishment and its large combustors. In February of 1981, the director of OMB withdrew OMB's approval of the survey during the fieldwork period, stating that the data collection was unnecessary.

Other EIA forms are currently being used to survey parts of the manufacturing industries. Forms EIA-3 and EIA-5, the Quarterly Coal Consumption and Coke Plant reports, respectively, provide quarterly data on coal receipts, stocks, prices, and consumption of coal for all uses. Form EIA-5 covers coal used for coke production; Form EIA-3 covers all other uses of coal. The surveys are designed to collect data from all manufacturing establishments involved in coal consumption. Forms EIA-810 and EIA-820 are monthly and annual

surveys, respectively, sent to all petroleum refiners and blenders in the United States. These forms collect data on actual and projected refining of crude oil and other unfinished petroleum inputs, along with product inventories, storage capacities, and onsite consumption of energy sources for refining.

In the meantime, other survey efforts are going on outside the Department of Energy. The Environmental Protection Agency maintains an ongoing inventory of emissions sources called the National Emission Data System (NEDS), which includes industrial sources, to provide data down to the emissions-source level (an individual smoke-stack, or even an individual process or piece of emitting equipment). This system is referred to as NEDS. Descriptive information about emissions sources customarily includes consumption by fuel, at least annually and often quarterly. The consumption values and other descriptive data are adjusted by EPA-supplied conversion factors to derive expected emissions values. The NEDS is supposed to include all emissions sources that emit at least 100 tons per year of all measured substances combined (sulfur, nitrogen oxides, particulates, and volatile organic chemicals). Another survey, of approximately 21,000 large manufacturing establishments, was started in 1983 by the Technical Economic Services Division of the Dun & Bradstreet Corporation. This system collects annual data on energy consumption, fuel switching, and equipment and process information, by a rather informal telephone-interview procedure supplemented by model-based imputations for missing data. Finally, there have been a number of survey efforts in recent years aimed at measuring energy consumption and conservation activities, conducted by State and local governments, utili-ties, industry groups, and academic groups.

Each of the data systems mentioned above--at least each of the Government systems--was developed for its own legal purposes, covering a subset of manufacturers specific to the purposes of the survey, and collecting a specialized set of data. When a specialized approach to data collec-tion is used, there is a natural inclination to design new systems to meet new data needs.

In speaking with the EIA, industry representatives have been very concerned about the proliferation of surveys. The concern is greatest in large, diversified corporations that have activity in almost every area of inquiry. The question we have been most frequently asked is, "Isn't there some way you can produce the data you need from the data available?" As anyone who has worked with statistical data knows, it is rarely possible to piece together various data collec-tions to produce a meaningful composite. Usually in such cases the available data are incomplete, there are large gaps in coverage, and the data available are plagued by such drawbacks as inconsistent definitions, different reference periods, and varying degrees of analytic capa-bility. These and other logistical problems cause the composite data to be overwhelmed by caveats, to the point that it is not at all certain what the composite data represent or how they could be used. Such is the case with manufacturing-energy-use data. In fact, the other existing energy data systems are so far removed from the purposes of the MECS that it does not appear that any great benefits could be derived by linking MECS with any of them. The one vehicle that appears to be suitable for integrating the MECS is the ASM, which is not in itself an energy data collection but which adapts itself well to the purpose. Using the ASM data system to field the MECS would continue the tradition of system integration that has been characteristic of EIA end-use consumption surveys.

When the MECS is fielded, a pertinent question then becomes, "Can other related information needs be met by the MECS system?" Current EIA policy is to restrict the MECS data collection to a minimal set of baseline data. However, one of the advantages of a baseline data system is its flexibility. The EIA needs to be sensitive to the possibility of using the MECS as a system for incorporating other related data-collection requirements and thus helping to minimize total burden from Government data collections.

SUMMARY AND CONCLUSIONS

The MECS has undergone an extended, laborious shakedown period in its development. Serious questions have been asked both from within and outside EIA about what constitutes appropriate baseline data for manufacturing, the largest single end-use consumption sector in the U.S. economy. The fact that the MECS is designed to satisfy benchmarking

purposes and to provide answers to specific policy questions sets it apart from other data collections developed since 1970. That flexibility, which is character-istic of EIA end-use consumption surveys, can allow the MECS to be a dynamic, appropriate tool for surveying manufacturing in coming years.

There have been several important issues raised, and lessons learned, in the development of the MECS that are of importance not only to EIA but to other data-gathering organizations throughout the Federal Government and outside. Some of these issues have been made particularly clear because of the current economic and political environment in which Government statistics organizations operate. They include:

(1) Know the subject matter--This issue is especially important for a complex subject area such as industrial energy throughput. Subject-matter specialists can offer valuable input about their area of expertise, but they cannot take the place of a statistician in designing efficient, thorough data-collection procedures or appropriate estimators. On the other hand, a statisti-cian without proper subject-matter understanding (not just advice) is in danger of designing collection and estimation procedures that are statistically sound but cannot be implemented in the real world.

(2) Get the big picture--With the current emphasis on efficiency and burden reduction in Federal survey activities, there is nothing wrong with looking for opportunities to coordinate and integrate one's data system with others, even in other agencies. There may be instances where loosely related or even unrelated topics may appear on the same questionnaire, but if such procedures produce the most efficient distribution of burden and do not significantly compromise the collection of any important characteristics, they should be encouraged. One point should be made here, however; the most efficient distribution of burden does not necessarily mean minimum burden. Individual and total burden must be weighed against each other. For populations such as households, where surveys ordinarily are carried out for a small proportion of the population and no individual units are high-probability selections, constructing separate surveys often makes sense, to avoid compounding burden on particular respondents. On the other hand, for a population such as manufacturers, which is dominated by a small cohort of major units that are highly likely to be selected for any survey, combining surveys makes more sense. If individual burden is a concern for smaller units, one possible approach is to break the comprehensive survey into sections, administer the entire survey to large units, and select panels of smaller units to repond to individual sections. To help data users get the big picture, it might be worthwhile to include references to related or complementary information in published reports.

(3) Build adaptability into surveys--One of the great advantages of justifying a survey system by enabling legislation as opposed to specific, restrictive law is that the survey system can adapt itself to changing data needs. The EIA, as a relatively young statistical organization, needs to be prepared to take advantage of such opportunity as fully as possible, although needs for trend data demand that certain key measures be kept constant if possible. Adaptabi-lity need not be limited to questionnaire content but can extend to population definition, sample size and population subgroups of analytic interest, data-collection method, and speed of analysis and data presentation.

(4) Consider a "life-cycle" approach--Government data-collection systems seem generally to fall into two categories, those specifically planned as one-time efforts, and those that are proposed for continuing collection, with no specific plan or means to declare that a program is complete or outdated and should be revised or discontinued. The easiest survey system in the world to run is one that is firmly entrenched and requires no more than minor modifications from time to time. However, in a field such as energy, in which changing

economic patterns, emerging technologies, and creative thinking are constantly changing policy emphasis and data needs, entrenched systems may not be practical.

We may have to design survey systems with the full intention of making major modifications or even discontinuing/starting over after a certain period. This procedure is currently taking place for EIA's NBECS, where major revisions in sample design, population coverage, and questionnaire design are being investigated/developed as a result of the first two cycles of data collection.

(5) Take review responsibilities seriously--One problem with the life-cycle approach to systems design is that the Government would find it difficult to render an impartial judgment on an existing system. The organization sponsoring the existing system has a vested interest in supporting its continuation. On the other hand, an organization that sees a responsibility to gather related but not duplicative data has no reason to support an existing system if it does not meet its needs. In fact, because of the pressures of the Paperwork Reduction Act and current attitudes toward data collection, the organization may even be pushed toward withholding support of an existing system or support-ing its own system at the expense of the existing system. This type of relationship is highly competitive and dangerously close to being adversarial, and cannot be tolerated.

The most obvious approach to the problem is through independent peer review. I realize that commenting on survey proposals that cross your desk from other organizations is not a glamorous job, nor is it easy to do well. Many times a reviewer feels totally unable to provide authoritative commentary. However, when a reviewer finds that (s)he has subject-matter expertise and interest, (s)he can do the Government a large favor by providing careful review, to the point of addressing other related data systems, if possible. As a really wild suggestion, the ASA itself might want to take an increasing role in assessing data needs, and the relative merits and drawbacks of related data systems.

The ideas of maximizing efficiency, avoiding duplicate data collection, and minimizing respondent burden are as old as survey statistics itself. However, in the conditions of tight money and public concern under which we now operate, we need to think of aggressive and innovative ways to work toward these goals.

---

*These data were supplemented for 1977 by feedstock information gathered for that year's Census of Manufactures, which was conducted for all establish-ments on the Bureau's standard list of 300,000-plus establishments.