Abstracts for EIA's Spring 2006 Meeting
with the
American Statistical Association
Committee on Energy Statistics

1. How Can Modeling Suggest Data Needs? Open discussion between the Committee and EIA. This session is prompted by Committee remarks in the fall 2005 meeting. Nancy Kirkendall, Chair, Margot Anderson, Director, EMEU, John Conti, OIAF and likely other EIA senior management.

2. Measuring Perceptions of Applying Alternative Disclosure Limitation Methods**,** Jake Bournazian, SMG Suppression is the most common method that federal agencies use to protect the confidentiality of reported data when releasing an information product. During the past 15 years, alternative disclosure limitation methodologies have been developed for protecting tabular and microdata. These methodologies offer new options in releasing data products for statistical agencies to protect the confidentiality of the reported data. Although these alternative methods offer an improvement to the information loss caused by suppression, these methods impact the utility of the information product to the data user. Research is needed to measure the perceptions of the data user community and the survey respondents of applying alternative disclosure limitation methods to confidential EIA data.

3. A New Oil Production Representation for the SAGE Model: Methodology and Producer Behavior Assumptions Justine Barden, John Staub, Glen Sweetnam, OIAF, EIA

The United States Energy Information Administration (EIA) uses the System for the Analysis of Global Energy Markets (SAGE) model for its annual International Energy Outlook. The model has 16 regions; each has well defined demand, supply, technology, and import/export representations. The SAGE oil production representation has received more scrutiny recently as world oil prices are approaching $70 per barrel. Model results show that the current approach of using supply steps to represent oil production can over state the supply response of oil in a world where demand grows at an unexpected higher rate.

EIA plans to test a new oil production representation for the non-OPEC regions and examine the performance of the model in response to various demand and price scenarios. For each time period, the new model will incorporate key factors such as exploratory drilling, reserve additions, development drilling, production per well, oil

prices, optimal drilling activities, and resource constraints, in its dynamic production analysis.

This paper describes a linear/non-linear modeling methodology that EIA is considering and raises the issue of price expectation in the determination of optimal drilling and production activities. The handling of price expectation may have profound effect on the projection of oil production and will be tested extensively when the model is fully operational.

4. Improving the SAGE Petroleum Refinery Model, John Staub, OIAF, Phillip Tseng, SMG, EIA

The United States Energy Information Administration (EIA) completed initial development of the System for the Analysis of Global Energy markets (SAGE) in early 2003. The model is built on a liner programming platform and is solved for the least cost of meeting a predetermined set of energy service demand, period by period over a five year time interval from 2005 through 2030. SAGE includes sixteen world demand and supply regions. For each region, there are four end-use demand sectors (commercial, industrial, residential, transportation), petroleum refining, power generation, and supply of both fossil fuels and renewable energy. The structure of the SAGE is generic. A modeler can relatively easily increase the number of demands for energy services and introduce new technologies into the system.

Recent developments in the world market for crude oil and petroleum products prompt the need to enhance the refinery representation of the SAGE model. The U.S. refinery acquisition cost (RAC) of crude oil rose from less than $26 per barrel in January 2000 to more than $55 in October 2005. Price differential between light sweet and heavy sour also widens in the same period; in January 2000, U.S. FOB cost of crude gravity 20 per cent or less was $20.78 per barrel and the same barrel for crude gravity 40.1 to 45% was $26.9 per barrel. In October 2005, the cost was $44.21 and $59.24 for the heavy and light crude oils; differences in prices increased from about $6 per barrel in 2000 to more than $15 per barrel in 2005. The increased price differentials reflect several important market interactions: demand share for light products increased more than heavier products, supply of heavier crude oils was relatively more abundant than light crude oils, and refineries lack down stream capacities to process heavy crude oils.

The world demand for petroleum products is projected to increase by almost 50 percent between 2005 and 2030. Most of the increase will be in gasoline, diesel, and jet fuel. The improved refinery representation will help EIA capture several important features of future petroleum market. They include investment requirements, product pricing,

product trade flow, price differential between light and heavy crude oils, and more reasonable forecast of long term supply of petroleum products.

The key to developing a manageable refinery model is adding only the essential elements of refinery operations and minimize unnecessary details. The refinery representation needs to mimic only the essential physical structure of refineries. For example, an abstract representation may remove sulfur before rather than after distillation.

5. 2006 Manufacturing Energy Consumption Survey (MECS): Looking at Past Performance Statistics to Motivate New Methods of Collection, Robert Adler and Tom Lorenz, EMEU, EIA

EIA's 2002 Manufacturing Energy Consumption Survey (MECS) data file has a wealth of "metadata" available in the StEPS database in which it is housed. For each data item collected or derived, a flag indicates whether the data is reported, corrected by analyst intervention, or the result of using an alternate data source. The number of analyst corrections for collected data items gives an indication of data quality.

This paper will present statistics for some key items that may indicate a problem in the wording of the question, conceptual understanding, or other problems in respondent reporting. This metadata examination will yield potential changes to the 2006 MECS.

The 2002 MECS also had an electronic option for reporting for certain classes of respondents. For those classes of respondents, we will compare the analyst intervention flags and other performance statistics between the electronic and non-electronic reporting groups. The results will be used to justify the use and clearance of an Internet Data Collection.

Performance statistics, especially non response statistics for industry and size classes, will be used to demonstrate the desirability of:

- Form specialization based on type of industry;
- Shifting the sample away from smaller respondents and allowing their weights to rise.

If available in time for the presentation, we will provide an update of frame and sample changes anticipated for 2006.

6. EIA 914: Data Expansion Challenges to Include Crude Oil Production, John Wood, OOG. (John Wood is at 214-720-6160)

(Abstract outstanding)

7. Making Adjustments to Survey Data When the Collected Data Do Not Meet Expectations. Stan Kaplan, CNEAF, EIA. Paper is to be on the EIA-920 data and information challenges. Statisticians will be interested because the form was changed before on the basis of cognitive testing, but still has some challenges. It may be that the Committee modelers and energy members will have useful ideas about the concept we are trying to collect and model.

8. Preliminary Research Results on Respondent Cut-off Dates for EIA Electricity Data Collections Howard Bradsher-Fredrick and Alethea Jennings, SMG, EIA In order to achieve high response rates on establishment surveys, EIA expends significant resources in administering non-response follow-up to those surveys. For example, our analysis of the submission dates related to the EIA-860, *Annual Electric Generator Report,* shows that over 95% of the volume has been reported within two months of the deadline for submission while EIA continues to conduct non-response follow-up for over four months following the final deadline. Considering tightening budgets, the issue can be raised as to whether EIA will be able to continue to expend significant resources to achieve near 100% coverage by volume.

In order to make rational decisions on this issue, it is advisable to study past data collections to first assess when data had been submitted and to then determine the character of the respondents and lost respondents associated with an array of alternative cut-off dates. This paper summarizes these preliminary analyses on 2004 submissions of EIA-861, *Annual Electric Power Industry Report,* and EIA-860, *Annual Electric Generator Report,* data. In addition to overall summaries some analyses were also conducted on various strata important to data users.

In addition, we would like to discuss with the Committee some of the challenges, such as that both the EIA-861 and EIA-860 surveys are used as frames for sample surveys. We would also like to discuss plans for future analyses, such as the use of imputation to obtain data for the missing respondents. We would like to get the Committee's comments on the work we have done so far and on our plans. Because this is work in progress, the paper may differ slightly from the abstract.

9. Functional Requirements for EIA's Internet Data Collection System**,** Stanley R. Freedman, SMG, EIA. An EIA team has been working to develop functional requirements for an EIA-wide Internet Data Collection (IDC) system. These requirements will serve as a basis for developing an IDC that will meet the needs of EIA's respondents, and survey managers. The work of the team is nearing completion as

reflected in the accompanying PowerPoint presentation given to the Goal 4 subcommittee for EIA's Strategic Plan. The team would like input from the ASA Committee on the requirements we have developed to this point.

10. An Empirical Evaluation of the Relationship Between Crude Oil and Natural Gas Prices, Jose A. Villar, OOG, EIA.(202) 586-9613

This paper seeks to develop an understanding the salient characteristics of the economic and statistical relationship between oil and gas prices. This analysis identifies the economic factors suggesting how crude oil and natural gas prices are related, and assesses the statistical significance of the relationship between the two over time. A vector error correction model is estimated to distinguish between long-run and short run effects of changes in natural gas prices on oil prices, and vice-versa. A significant stable relationship between the two price series is identified. Oil prices are found to influence the long run development of natural gas prices, but are not influenced by them.