DISCUSSION

Daniel Kasprzyk
National Center for Education Statistics

The Division of Science Resources at The National Science Foundation is actively and aggressively involved in working through a complete redesign of their Scientific and Technical Personnel Data System (STPDS). This is complicated, not just because of the complex sampling aspects of the study, but because the NSF is dealing with multiple surveys, multiple contractors, and multiple users, many with specific needs. The operation is also somewhat more complicated and difficult because of the fact that many members of the research community clearly felt very strongly that the system of the 1980's needed substantial improvement.

The managers at the NSF during the 1980's should be commended for funding a Committee on National Statistics (CNSTAT) Panel to review the data system design of the Scientific and Technical Personnel Data System. This review identified a number of probing questions and issues, and provided recommendations designed to alter the status quo of the system. The papers presented here today represent, I believe, the first discussion at these meetings (ASA) of how NSF has followed up on the recommendations of the CNSTAT panel.

In that sense this session is a watershed. This type of activity ought to occur at every annual ASA meeting. These meetings provide an excellent opportunity to bring together and report on recent activities and initiatives. The NSF ought to take pleasure at today's reporting of their recent contributions, but should not be complacent about their work in the future. That is, they need to continue to take advantage of the opportunities provided at these meetings.

The data system includes three surveys;
1. National Survey of College Graduates (NSCG);
2. Survey of Doctorate Recipients(SDR); and
3. Survey of Recent Science and Engineering Graduates (New Entrants).

The general goal is to provide information on the number of U.S. scientists and engineers and their demographic, educational, and employment characteristics.

The four papers presented today describe the activity during the last three years in addressing the concerns of the CNSTAT panel:

1. Dr. Shettle's paper characterizes how NSF views their work in relation to the major recommendations from the CNSTAT panel. There are good reasons to be pleased with her report. The NSF appears to be using the CNSTAT report for all major directions.

2. The Tsapogas and Gannon paper describes the constraints they worked with and the process they went through to identify the content of their questionnaires.

3. The Wilkinson and Hines paper provides background information on the data collection problems occurring in the 1980's and the NSF response to these problems in the next data collection cycle.

4. The Hardy paper describes the scope of the surveys, the sampling frames, sample design goals, and a few selected issues.

I have some comments on each individual paper, and then I will follow with a few general remarks.

Shettle

The paper by Shettle provides the broad view of the background to the current redesign effort and a status report on the major CNSTAT recommendations. Clearly much progress has been made and that is good, but much more remains to be done. Several specific points about the paper:

1. The author describes the reason for commissioning CNSTAT is to help ensure the design problems of the 80's do not reoccur in the 90's. While there were certain design problems that needed to be resolved, a number of decisions made with regard to the data system in the 80's had to do with budget cuts and ways the data system had to be altered to accommodate these budget cuts. The data system suffered for this.

Data collections are expensive and a small statistical unit in an agency known for its grant giving function is bound to suffer. I am not sure any panel can deal with the vagaries of an agency's budget reductions. It's not at all obvious how budget decisions in the 90's will be made and what impact the CNSTAT panel report will have on an agency's decision to cut funds for data collection. In the absence of increased funding for the system, the NSF staff should realize and be prepared to make difficult decisions concerning the data system as federal budgets continue to be reduced. I think the current staff will make decisions with an eye to the ultimate consequences on the quality of the data in their data system.

2. Two points about the coverage of the scientist and engineer population follow below:

a. Issues of population coverage are important, but we are never given any magnitude of the problem as exhibited in the 80's surveys. Assurances are given that the coverage will improve, but I wish the amount of coverage improvement with the New Entrants Survey had been quantified.

b. I also wish there were more discussion of missing populations -- those who flow into scientist and engineer occupations without a scientist or engineer's degree as well as the immigrant population. In particular, I had thought that some effort was being put into how best to deal with the immigrants. NSF's work in this area will be most interesting -- I only wish more had been discussed in the paper, including some of the past work.

3. The issue of data dissemination has been delayed in the plans of the NSF staff. This is understandable given limited time and staff. However, it can not be delayed much longer. NSF will need to develop a process of obtaining user input to identify the data products of the data system. This is particularly true for the issue of how microdata products will be made available to the public. Balancing the need for data for research purposes with confidentiality issues should be addressed soon. Reaching closure will be painstaking and difficult; the process could be drawn out. Don't delay the start of this activity. Some ideas which I am sure NSF will consider include: 1) restricted access to microdata files, perhaps through a licensing agreement; 2) establishing a system to provide tabulations and analyses for data users; 3) establishing a "fellow" or post-doctoral research program; and 4) establishing data analysis research centers.

Tsapogas and Gannon

This paper describes the history of content decisions for the Scientific and Technical Personnel Data System and the more recent initiatives to revise the content of the data system.

As described by the authors, the staff over the last 5-7 years have taken the opportunity to obtain suggestions on broad topical areas as well as specific items. They have used many forums -- data users conferences, workshop on data needs, the CNSTAT panel report, technical working groups on survey design and content. This is really quite good. What is not all that clear in the paper are the areas of content common to these groups and those areas that were not common. The authors then describe how decisions

on content were made as well as the questionnaire testing program. NSF is clearly working very hard at improving their questionnaires. Some specific points:

1. The authors suggest in the paper that historical comparability is a consideration in keeping questionnaire items. I certainly can understand the concern, but it seems to me that a considerable amount of bias is apparent in the estimates from the 80's. NSF has said this themselves. So it is not clear to me why one would want to encourage such comparisons. As it is, change estimates will be difficult to interpret. NSF will need to write a technical note on this for their users.

2. I am struck by the authors' statement that more recent forums (research users in 1989) provided information similar to data generated from earlier efforts. Apparently, the issues have not changed. Nevertheless, it is important for NSF to continue to communicate with the participants of these forums. Working groups are a good idea to maintain communication, but bringing together a large group of interested parties at a conference will maximize outreach.

3. The discussion on how content decisions were made seems slightly misdirected. Questions were dismissed because they were not easily formulated, too burdensome (income by source), and too intrusive (Social Security Number), for example. Shouldn't the criterion be more along the lines of the need for these data? If items are important, but difficult to formulate, shouldn't some effort be expended to simplify the formulation and collect the data that are needed? The criterion ought to be the item and its importance.

4. A number of lab sessions, focus groups, and retrospective think aloud interviews were held culminating in a pretest of the National Survey of College Graduates questionnaire. This seems reasonable and thorough. To improve response, don't forget the look of the form is important. You may need to consider significant formatting changes to help elicit a response.

With regard to the pretest, it seems to me that conducting a professional review by the staff of a subset of the returned questionnaires will provide important insight into how the respondent interprets the questions. In addition, the continuation of small-scale tests is advisable.

Wilkinson and Hines

Wilkinson and Hines describe the earlier data collections indicating the considerable response problems these surveys had. They describe their considerable efforts at improving the data collection during this cycle of the program. Some specific comments:

1. Locating respondents is a serious problem for the STPDS data system. As proposed, using the change of address file and look up operations will help. Because of costs, the personal visit should be used as a last step. If a personal visit is conducted, I suggest that a separate cost study also be conducted to get better idea of the cost and effort of this task. Certainly to reduce costs, adopting the personal visit program could be done on subsample of the unable to locate part of the sample. Results of such a study could feed into decisions on how best to handle the unable to locate in the estimation phase.

2. Concerning the use of non-monetary incentives to motivate response - while I know the use of an informational brochure was tested, I say, why bother. I personally think such information about the survey should be an integral part of any correspondence with the respondent.

3. Concerning the use of the monetary incentive to motivate response, the authors seem to feel this feature improves response rates, but are you willing to do this for all respondents or just the difficult to convert respondents?

4. Details of the test on the effect of questionnaire length on response were not provided in the paper. This test seems potentially important to the statistical community; results ought to be reported.

5. The authors briefly discuss improving longitudinal response rates. Maintaining contact between waves of the panel will improve the longitudinal response rates. The authors suggest the use of informational materials at this contact. That is a good idea, but I would also add including a thank you letter and possibly propose an incentive ($5 or more) to report a change of address.

6. The authors suggest additional work on nonresponse bias, mode effects, and improving the ability to locate respondents. Nonresponse studies should address the issue of how best to adjust for unit nonresponse.

If a reasonable proportion of interviews are taken over the telephone, mode effect studies can help indicate data quality problems.

According to the authors, respondents seem to have some difficulty with the contact person concept as a way to maintain contact rates. I have not heard or seen this a major problem in other surveys. Perhaps, it's the population of interest. I would be very reluctant to drop this item until I was absolutely certain that other methods worked.

The authors suggest a number of possible studies. How will priorities be assigned? Can they really do all of these studies before the 1993 data collection?

Hardy Paper

The Hardy paper discusses the framework for the sample redesign, identifying the wide diversity of uses and users as well as the need to allow for subpopulation of interest -- minorities and women. It provides some general guidelines for the sample allocation the most important of which is to limit the range of sampling rates. In general, this paper provides the background for future work concerning sampling in the STPDS data system.

The point of view it takes seems directly related to the CNSTAT panel. Some specific points:

1. While some design goals are identified, the paper does not show data on the effects of alternative designs nor does it provide explicit information on how decisions will be reached.

2. The author argues that using a variable which is set to a fixed proportion across strata, thus fixing the variances, is a reasonable way to look at the relative effect of various designs. It is, but articulating this effect in concrete terms by identifying the variable, the number in the cell, and the expected coefficient of variation will help users to understand the implications.

3. Estimating the out-of-scope by detailed cell is difficult, but can we not use several years of survey data to help--for example, using the CPS?

4. I am not sure when data collection begins, but we must be getting very close to the deadline for supplying sample allocation numbers. So the lack of detail in this report on this topic is surprising.

General Remarks

1. Several of the papers, particularly the Tsapogas and Gannon and the Wilkinson and Gannon papers, refer to focus groups, cognitive lab work, tests and experiments implemented and analyzed by the NSF. There were no citations, however. I think it is very useful to cite the source documents and/or provide more information about this work at future meetings of the ASA. Several years from now, after staff have

changed jobs, those involved in the program will want to refer to these documents. This is particularly true if a Quality Profile is to be developed.

2. It is unclear as to the future balance of funds between "research and evaluation" and data collection. The program will have invested wisely if some funds are allocated in the future to a continuing research and evaluation program. This includes establishing a small research sample in the data collection years.

3. The emphasis on nonrespondents, knowing who they are and how they differ from the respondents, etc., should provide big dividends to the program, particularly in its longitudinal aspects. Use of this information for nonresponse adjustment should be investigated.

4. I wish the papers had more discussion of the difficulty of comparing items across data systems in the U. S. government.

5. There is no mention of how NSF plans to monitor the quality of the collected data; nor is there a discussion of plans for evaluating the data. Developing a plan that includes a reinterview component to measure response bias and response variance is advisable.

6. I suspect the use of multiple agencies for data collection complicates the analyses from this data system. Procedures and definitions need to be as consistent as possible, but even with this in place each collection agency has its own idiosyncracies. Steve Cohen (AHCPR) did some work on this a few years ago. Is there a data collection agency effect?

7. Develop an articulated program to evaluate data collected in these surveys. Make the commitment.

Much has occurred over the last several years; much more has to be done. I hope the staff continues to show the energy and enthusiasm for the program and that the commitment to a first class data system does not waver in the future. Thank you for the opportunity to discuss these papers.