

DISCUSSION

Daniel H. Garnick, Bureau of Economic Analysis

To discuss six papers, all weighty, all carefully constructed, all on different applications of different administrative record systems, and all deserving of similarly carefully constructed comment is not a trivial assignment. Fortunately, the constraints of time permitted in this session (and not having received two of the papers) limits the vastness of the abyss of ignorance from which I must ascend to address these papers with the seriousness they each deserve.

There is no unifying theme running through the papers except by hearkening back to the Report on the Statistical Uses of Administrative Records, (SUAR), Statistical Policy Working Paper No. 6. In that report, general recommendations were made with respect to potential uses and to overcoming significant problems inhibiting major new uses of administrative records.

The Cartwright/Levine/Buckler paper is most closely identified with the SUAR report and it is fitting to begin with their paper. They deal with the fundamental issue of developing common identifiers for establishments among the various administrative record systems. It surpasses the imagination to suppose that inconsistencies of results among the data sets originating in the different administrative records systems can be statistically resolved without resolution of the question of common identifiers for establishments. The proposal to tie the Social Security Administration (SSA) establishment reporting plan to the Standard Statistical Establishment List (SSEL) appears to be a practicable way to circumvent some of the well-known and widely regretted problems of access to the SSEL. How can one explain the continued resistance to rescue the SSA establishment reporting plan and thereby save a deteriorating, albeit demonstrably useful, Continuous Work History Sample? It does not appear that the SSA establishment reporting system can be saved without closer links to the SSEL or the Unemployment Insurance systems. The resistance of the statistical agencies to the interagency cooperation required to effect this rescue speaks louder than words about the needs for a strong coordinating function in the fragmented, and, often irrational, Federal statistical system. Only decisive action from the Office of Management and Budget can adequately coordinate these efforts. The community of users of government statistics must be heard in order to focus OMB's attention on this vital issue.

Michael Colledge concludes that "many of the problems in the provision of frame data for economic surveys are ubiquitous." There are distinct parallels between the Canadian and U.S. situations, but by virtue of its centralized structure, Statistics Canada has avoided many of the problems created in the U.S. by lack of access to the SSEL outside the Census Bureau. Canada may very well have

inadequate coordination among surveys in terms of maintaining and improving a central list. But all Canadian survey programs have access to the main list so there is no problem of complete independence and inconsistency as exists in the U.S. between the Bureau of Labor Statistics and Census, not to mention the problem described above with respect to the establishment reporting plan of SSA. Moreover, "detailed data exchange protocols" have recently been developed in Canada to standardize and coordinate use of the list. How long will the U.S. statistical system remain behind? But if Canada is ahead of the U.S. in access and coordination, it seems to be somewhat behind in terms of defining and delineating establishments. For example, the definition of establishments on page 7 appears to ignore the geographic dimension. The practical reporting unit is only vaguely defined in terms of units from which "principal production account statistics can actually be obtained." Nor, from my perspective, is the multi-establishment company problem sufficiently developed in the context of geographic data.

Susan Hostetter's "verification method" for updating SIC codes for previously coded businesses involves sending each coded employer a 4-digit industry coded description with which the employer can agree or disagree. This procedure is designed to reduce the reporting burden. The major potential problems identified in the paper are 1) employers may answer "yes" when that answer is not appropriate (in order to avoid the burden of redescribing their industry), and 2) the precoded industry description may be inadequate and lead to incorrect or needlessly detailed responses. BLS has devised a test to check the "yes" responses. And unless it is found that biased responses are unacceptably high, it appears that the verification method makes sense. Some questions are not treated in the paper: 1) While the telephone followup study is probably called for, has BLS considered that a first approximation on the potential magnitude of the "yes" bias problem might be gotten by comparing the (90%) "yes" response rate with the proportion of codes which have remained unchanged under the old, "full classification" approach to reclassifying, and 2) on the assumption that there is some chance of misclassifying anytime a description of activities is requested, why did BLS decide not to have a telephone followup on any of the "no" respondents? Possible inconsistencies in the group of "no" respondents might demonstrate that even if the verification approach has a bias toward "yes," that bias would not necessarily imply that the "inferiority" of the verification vs. the full-classification approach is as great as would be guessed solely on the basis of the number of respondents giving inconsistent "yes" activity descriptions when pressed in

the telephone followup.

The Klein/Kasprzyk paper is concerned with integrating data files providing information on applicants and beneficiaries under the Titles II and XVI disability programs administered by SSA. The authors state that their efforts are only a first step in design for paring and synthesizing the data files towards creation of an integrated data base that would serve administrative needs and provide a "rich source of information for

disability research." This is a useful exercise, as far as they have gotten, one that might eventually help meet real data needs cost-effectively. It is to be hoped that the authors will go on to estimate the resources required for the integrated data system and to estimate the research needs, not currently being met, that would be met with this system. They should also consider, against this, how much savings might accrue by eliminating some of the underlying data systems.