

DISCUSSION

Jack Silver, Bureau of the Census

In comparing a set of study records with an independent set, the number of successful matches depends not only on the specifications for a match and the "true" proportion of cases in the independent file which actually match, but also on the reliability of the mechanism, whether computer or clerical. Many of us have experience with clerical matching operations in which the number of successful matches increases each time the operation is repeated. It would seem to me that a useful part of any clerical matching operation would be estimates of the "failure to match" rate (when match is possible) and the "successful match rate" when match is not possible under the specifications.

Computers must follow specifications. Clerks performing visual matching, however, have been known to modify specifications to increase the number of matches by taking advantage of information in ways not permitted by the specifications. On the other hand, they may fail to make

permissible matches for any number of reasons. Random samples of the study records could be used to determine the extent to which specifications are not followed strictly. (It is quite conceivable that one may want to modify the original specifications to take advantage of clerical ingenuity used in increasing the number of matched cases.)

While listening to the Simpson and Van Arsdol experience in matching juvenile delinquency records against the 1960 Census, I wondered whether part of the failure to match rate could be attributed to some undercoverage of a population which may be inherently more difficult to enumerate. I think it would be interesting to know whether the failure to match rates would have been statistically different had the same specifications been employed in matching the records for a more representative group of persons in this age group against the 1960 Census.