

RECORD MATCHING--THEORY AND PRACTICE (Abstract)

Philip M. Hauser, University of Chicago and
Patience Lauriat, U. S. Bureau of the Census

After discussion of some of the principles and concepts behind matching of independent records, the paper presents some discussion of the experience of the U. S. Census Bureau in a study sponsored by the University of Chicago. The purpose of the study was to obtain data on the socio-economic characteristics of recent decedents by matching a sample of death certificates, representing deaths occurring during the period May through August 1960, to the Stage I and Stage II 1960 Census records. An overall match rate of about 80% was obtained. The matching was done clerically, using the address of usual residence, primarily, as well as name,

color, sex, and age of the decedent. If a person could not be matched at the usual address given on the death certificate, or the usual address could not be allocated to a census enumeration district, the enumeration book for the institution where the death actually took place, if the death occurred in an institution, was searched for a possible name match. In order to obtain data on socio-economic data of decedents under one year of age, a possible match was searched using the mother's address and name as reported on the death certificate. The paper also includes some information about basic demographic characteristics of matched and unmatched decedents.

AN EVALUATION OF SOME MECHANIZED MATCHING OPERATIONS (Abstract)

Joseph F. Daly, Herman Fasteau and Peter Ohs,
U. S. Bureau of the Census

The Bureau is now preparing to conduct Economic Censuses covering the calendar year 1963. This paper relates some experience bearing on our ability to mechanize three of the matching operations associated with these censuses. All three operations seem simple enough at first sight, but turn out to require a surprising amount of planning and control. One involves the matching of addresses (nearly 4 million in number) on a mailing list against a reference file to assign geographic codes needed

for the tabulations. The other two match returned reports against master lists to determine which firms have reported and which firms need further follow-up. Semiautomated techniques for checking in the reports have been tested on the 1962 Annual Survey of Manufactures and on the prec canvass of multi-unit firms with encouraging results. And the computer operations as associated with the advance location coding of the mailing list are now far enough along to be evaluated. (Related documents are available upon request to the authors.)