DISCUSSION

Peter A. Morrison, The Rand Corporation

Population estimates usually are prepared with practical rather than experimental objectives foremost in mind. The contexts of many comparative studies so far have arisen fortuitously, leaving little room for experimental design in advance. Even when such forethought has been possible, adherence to a rigid scientific design usually proves impractical. The following points illustrate complications that typically arise:

- The quality of the data used for estimation generally varies from one population to another, with uneven effect on individual methods. The accuracy of Component Method II, for example, is vulnerable to poor school-enrollment data since these form the basis for its migration estimate. The quality of data and the precision of a method, therefore, cannot be separately distinguished.

- In small-area applications, modifications often are necessary to adapt an estimating method to the local data environment. Instances where one method has been applied to every population in a computationally consistent manner are the exception rather than the rule. These variations reduce strict comparability to an unknown degree.

- Comparisons among methods that have been applied to separate universes are especially hazardous, since the accuracy of an estimate varies systematically with a population's absolute size and its relative rate of growth. One method may appear more precise than another simply because the former was applied to a disproportionate number of heavily populated or slowly growing areas, both of which lend themselves to more precise estimation.

- Statistical measures used to gauge the relative precision of estimating techniques are inadequate. Effective comparisons are difficult, and statistical appraisals of differences are rarely conducted. The conventional measure adopted in most studies is the mean of percentage deviations, neglecting signs, between estimated and enumerated populations (symbolized hereafter as $\overline{D}$).[1] This measure indicates relative error independent of an area's absolute population size, thereby weighting large and small study populations equally. As a result, a few numerically small populations--for which relative error can be large--may swamp the measure, overstating the actual degree of imprecision.

- The short estimating periods used in some comparative studies occasionally favor simple extrapolative procedures over more analytical techniques. These results must be viewed with caution. A minimum imprecision is inherent in several of the latter methods; and although simple extrapolation will sometimes outperform them in the short run, it should not be concluded that this advantage will hold for longer intervals.

All of these complications discount the value of most comparative evidence now available. Exhaustive performance tests based on 1970 census data are planned by the Bureau of the Census. For now, the current evidence lends itself to only a few general conclusions.

CONCLUSIONS SUPPORTED BY COMPARATIVE EVIDENCE

1. *No single method of estimating local population shows consistently greater accuracy, although the Regression Method continues to look most promising.*
2. *Evidence consistently shows that lower average error can be attained by averaging together estimates made by different methods.*
3. *Average error tends to be lower for counties whose populations are large or metropolitan.*
4. *Average error varies with rates of population growth.* $\overline{D}$ is lowest among slowly growing counties, followed by rapidly growing counties, followed by counties losing population.

$$\overline{D} = \frac{\sum_{i=1}^{N} \left| \frac{\text{Estimated} - \text{Enumerated}}{\text{Enumerated}} \times 100 \right|}{N},$$

where N is the number of populations for which estimates are prepared.

Reference:

Morrison, Peter A., *Demographic Information for Cities: A Manual for Estimating and Projecting Local Population Characteristics*, R-618-HUD, The Rand Corporation (June 1971).