Measuring the uncertainty of population forecasts: a comparison of two approaches: JOEL E. COHEN (*Rockefeller University*)

A population projection (or, synonymously, forecast) is an estimate of the numbers, composition, or distribution of a future population. The uncertainty of past population forecasts has recently become the subject of systematic analysis (e.g. Stoto (1983)), which indicates that past forecasts have generally failed to recognize realistically the extent of their own uncertainty.

This paper reports and evaluates the first use for human populations of two methods of generating confidence intervals for forecasts of future population size. One method, developed by Heyde and Cohen (1985) and Heyde (1985), applies martingale limit theorems (Hall and Heyde (1980)) to a demographic model for age-structured populations with stochastic fluctuations in vital rates. This model is based on products of random matrices (Furstenberg and Kesten (1960)). The method of Heyde and Cohen gives both projections and confidence intervals for those projections. It does not give confidence intervals for projections derived from other models.

A second method adapts an empirical approach, proposed by Williams and Goodman (1971), to constructing confidence limits for economic forecasts. This method requires that population projections be derived from another source. The method estimates confidence intervals for the forecasts by empirical analysis of the errors in the forecasts.

In comparing the confidence intervals generated by the methods of Heyde and Cohen with those generated by the methods of Williams and Goodman, we forecast according to Heyde and Cohen in both cases. The forecasting procedure of Heyde and Cohen boils down to fitting a simple exponential curve exactly through the first and last data points of a time series of population sizes. This is a familiar and appealing procedure

(e.g. Stoto (1983)). The main interest of the analysis is clearly not the projection procedure itself, but the approaches offered to measuring the uncertainty of forecasts.

The data analyzed are the estimated total population sizes of Sweden at five-year intervals from 1780 to 1980 inclusive. Four checks of the model and of the estimators of Heyde and Cohen are carried out. First, the asymptotic long-term exponential growth predicted by products of random matrices is compared with the observed trend using all the data from 1780 to 1980. Second, some historical pseudo-experiments are performed using early data to project confidence intervals and later data to test them. Third, a 'confidence funnel' for the projected population in 1980 is constructed using data available up to 1930, 1935, $\cdots$, 1975. Fourth, the 95% confidence interval for 1970 projected from 1965 is examined to see if the range of variation is demographically plausible.

Comparison shows that the confidence intervals estimated by the method of Heyde and Cohen are consistently larger than those estimated, following Williams and Goodman, from the empirical distribution of forecasting errors. A 68% confidence interval for the Swedish population 5 years in the future based on 15 quinquennial observations varies by approximately $\pm 3{\cdot}1\%$ of the projected value according to the estimator of Heyde and Cohen, and varies by approximately $\pm 1{\cdot}5\%$ of the projected value according to the approach of Williams and Goodman.

Because so little analysis and prior experience are available concerning the behavior of the Heyde–Cohen and Williams–Goodman estimators of confidence intervals when used with demographic data, we can conclude only that the Heyde–Cohen confidence intervals may be too wide, or the Williams–Goodman intervals may be too narrow, or both methods may miss the mark. Further analysis and development of both methods would have practical value.

This work was supported in part by U.S. National Science Foundation grant BSR84-07461.

FURSTENBERG, H. AND KESTEN, H. (1960) Products of random matrices. *Ann. Math. Statist.* **31,** 457–469.

HALL, P. AND HEYDE, C. C. (1980) *Martingale Limit Theory and Its Applications.* Academic Press, New York.

HEYDE, C. C. (1985) On inference for demographic projection of small populations. In *Proceedings of the Neyman–Kiefer Memorial Conference,* ed. R. Olshen.

HEYDE, C. C. AND COHEN, J. E. (1985) Confidence intervals for demographic projections based on products of random matrices. *Theoret. Popn Biol.*

STOTO, M. A. (1983) The accuracy of population projections. *J. Amer. Statist. Assoc.* **78,** 13–20.

WILLIAMS, W. H. AND GOODMAN, M. L. (1971) A simple method for the construction of empirical confidence limits for economic forecasts. *J. Amer. Statist. Assoc.* **66,** 752–754.