"Evaluating County Population Projections Using a Century of Census Data"

By Stefan Rayer

University of Florida

stefanr@bebr.ufl.edu

Paper submitted for presentation at the annual meeting of the Population Association of America, March 31 to April 2, 2005, Philadelphia, PA.

Extended Abstract

This paper evaluates summary measures of population projection accuracy and bias for a large sample of counties and county equivalents in the continental United States over the period 1900–2000. Many of the studies that have evaluated projection accuracy and bias did so using a limited set of data both with regard to space and time. The ex post facto projections analyzed in the present study utilize decennial census data for the entire 20th century based upon a sample of close to 2,500 counties for which comparable data were available. A total of 10 trend extrapolation techniques are examined: seven primary techniques – linear, modified linear, share-of-growth, shift-share, exponential, constant-share, and constant – in addition to three averages. For each of these techniques, 125 projections are calculated, covering a range of 10–50 year projection horizons and 10–50 year base periods for each census target year between 1920 and 2000.

The analysis has three primary purposes. First, it examines the appropriateness of using the mean absolute percent error (MAPE) as a summary measure of projection accuracy, and the mean algebraic percent error (MALPE) for determining bias. Although widely used, mean-based measures such as the MAPE have been criticized for their tendency to overstate forecast error. To evaluate this claim, two robust summary measures of error that have been advanced as alternatives – the median absolute percent error (MedianAPE) and an M-estimator (Tukey's Biweight) –

are compared to results obtained using MAPE. The study finds that MAPE can be unduly influenced by outliers and that in some instances more robust measures should be considered when evaluating the accuracy and bias of population projections.

Second, the study reexamines the relationship between the length of the projection horizon and the length of the base period. Using a rich dataset that covers the entire 20th century allows for a wide range of projection horizon/base period combinations to be analyzed. In particular, the study examines whether the length of the projection horizon should correspond to the length of the base period or whether a 10–20 year base period is sufficient, even for long range population projections. The study finds that for most projection techniques extending the base period from 10 to 20 years improves forecast accuracy, but that beyond 20 years little additional improvement is achieved.

Third, the paper investigates whether particular trend extrapolation techniques are more appropriate than others for projecting the population of counties exhibiting specific size and growth characteristics. While averaging has been advanced as a response to the inherent uncertainty related to deciding among competing projection techniques, some researchers speculate that specific methods could be more suitable than others for certain types of counties, although little comparative research has actually been conducted on this issue. The present study finds that some trend

extrapolation techniques can be more appropriate than others, but also concludes that averaging remains a powerful, and sometimes preferable, alternative.

In addition to these specific objectives, a main feature of this study is the attempt to understand how past trends of population change can inform decisions about the future. Simple trend extrapolation models are often not held in high regard among population forecasters. Yet numerous studies have found that more complex and sophisticated techniques are generally no more accurate. While the past may not always repeat itself, this study concludes that using historical data for a wide range of years, a broad sample of geographical units, and a rich cross-section of projection techniques can provide valuable guidance that aids the projection of future population.