

September 6, 2002

**DRAFT**

## **Computation Technique for Intercensal Estimates**

This is a presentation of the general method used to produce intercensal estimates. It is produced in response to inquiries resulting from the recent release of county total intercensal estimates for 1990 through 2000, but is written in a general fashion since it is applicable to many situations.

### **Introduction**

There are several postcensal estimates products made annually by the Population Division. These are based on the latest decennial census and brought forward to the current estimate date using a variety of methods and data sources. Since they are computed *after* a decennial census, they are called *postcensal* estimates. When the results of a new decennial census are available, we typically generate a new series of estimates which is based on both the previous and new censuses. These are called *intercensal* estimates, since they are based on censuses at each end of the time series and are therefore *between* two censuses.\* The major factor to bear in mind is that the intercensal estimates that we produce are not only based on the two sets of data from the censuses, but also on the original set of postcensal estimates developed during the decade. The basic goal in making intercensal estimates is to take the existing time series of postcensal estimates and revise it so that the value for April 1 of the newest census year will have the same value found in the census.

### **Procedure**

*Preparation.* As mentioned above, even before actual new decennial census data are available, we can produce postcensal estimates through to April 1, 2000 ("new" census year). Since we typically produce estimates for July 1 of each year, the April 1 figures are usually obtained through interpolation or extrapolation, depending on what is available.

*Computation.* When we have both the April 1, census year postcensal estimates figures and the decennial results, we can produce intercensal estimates. At the center of the question of how to make these estimates so that the time series arrives at the decennial census figure rather than the postcensal estimates figure is the difference between these two numbers. It would indeed be miraculous if not impossible for all the numbers to be equal. The choice of method for making estimates for each year involves how to distribute this difference to the various annual estimates in the time series. The technique we currently use is known as Das Gupta 6, since it comes from a memorandum written in the 1980s by now-retired Prithwis Das Gupta, formerly of the Population Analysis Staff of the Population Division. This technique is given by the formula

$$P'_y = (C_c/E_c)^{y/n} * P_y$$

where

$P'_y$  is the intercensal estimate for year  $y$

$C_c$  is the decennial census number, year  $c$

$E_c$  is the postcensal estimate for the estimate year

and  $n$  is the number of years in the time period between censuses (10)

$P_y$  is the original postcensal estimate for year  $y$ .

Examining this formula, we see that it is basically adjusting the postcensal estimate for each year using the ratio of the census value to the postcensal estimate for the census date raised to an exponent which increases over time. An interesting result which can be used to check the computation is the fact that when  $y=10$ , i.e. the census date, the whole formula decomposes to the intercensal estimate equaling the decennial census value, which is the desired goal.

## Discussion

This technique for computing intercensal estimates, as the name implies, is the sixth method presented by Das Gupta. There are some desirable properties of this method which led to its selection. The very simplest method is simple linear interpolation. The problem with this is that it completely ignores the results of the postcensal estimates, and assumes linear change between decennial censuses. Both of these factors have deficiencies, since we believe the use of a postcensal estimates method contributes to ascertaining a pattern of population change, and we doubt the total linear change of an area's population during the intercensal period. This doubt comes from both evidence suggested by various data as well as less formal sources.

Even dismissing simple linear interpolation, there are still a variety of techniques for computing the intercensal estimates. The two properties of Das Gupta 6 which were attractive to those choosing it were that it preserved the overall "shape" of the time series over the decade, and it tended to distribute the difference between the postcensal estimate and the decennial census value toward the end of the decade. This second feature is consistent with our belief that postcensal estimates tend to vary from the truth to a greater extent for years further away from beginning decennial census.

\* There is at times some confusion in the use of the terms intercensal and postcensal. In demographic usage, the terms are used as they are in this document. In the parlance of management and budget persons, however, estimates that are properly called postcensal are referred to as intercensal. This relates to the time for which the estimates are made, i.e. years falling between censuses, rather than the data used to make the estimates. The reader may encounter this usage in the future in other documents, and should not confuse the intercensals discussed here with the intercensals referred to in other documents where what is actually intended is postcensal.