

On the Likelihood of Improving the Accuracy of the Census Through Statistical Adjustment

David A. Freedman and Kenneth W. Wachter

Abstract

In this article, we sketch procedures for taking the census, making adjustments, and evaluating the results. Despite what you read in the newspapers, the census is remarkably accurate. Statistical adjustment is unlikely to improve on the census, because adjustment can easily put in more error than it takes out. Indeed, error rates in the adjustment turn out to be comparable to—if not larger than—errors in the census. The data suggest a strong geographical pattern to these errors even after controlling for demography, which contradicts a basic premise of adjustment. Complex demographic controls built into the adjustment mechanism turn out to be counter-productive.

Proponents of adjustment have cited “loss function analysis” to compare the accuracy of the census and adjustment, generally to the advantage of the latter. However, these analyses make assumptions that are highly stylized and quite favorable to adjustment. With more realistic assumptions, loss function analysis is neutral or favors the census. At the heart of the adjustment mechanism, there is a large sample survey—the post enumeration survey. The size of the survey cannot be justified. The adjustment process now consumes too large a share of the Census Bureau’s scarce resources, which should be reallocated to other Bureau programs.

Keywords: Census; adjustment; heterogeneity; correlation bias; demographic analysis; dual-system estimation; non-sampling error; loss function analysis

1 Introduction

The census has been taken every ten years since 1790. Counts are used to apportion Congress and redistrict states. Furthermore, census data are the basis for allocating federal tax money to cities and other local governments. For such purposes, the geographical distribution of the population matters rather than counts for the nation as a whole. Data from 1990 and previous censuses suggested there would be a net undercount in 2000; the undercount would depend on age, race, ethnicity, gender, and—most importantly—geography. This differential undercount, with its implications for sharing power and money, attracted considerable attention in the media and the court-house.

There were proposals to adjust the census by statistical methods, but this is advisable only if the adjustment gives a truer picture of the population and its geographical