A Simple Example of the Censal Ratio Method for Population Estimation

Mathew E. Hauer *

Department of Sociology, Florida State University

August 8, 2023

Abstract

Lorem Ipsum.

Keywords: keyword, keyword, keyword

---

*Thanks y'all!

1   Introduction

Population estimation is generally a straightforward process: any population must result from a past population number plus the births minus the deaths plus the net migration. This cohort-component method is often considered the 'gold standard' for population estimation (Gerland 2014). However, the components of change are often lagging indicators, sometimes preventing immediate estimation of a population until all births, deaths, and migrants are enumerated, some populations do not have components of change due to their small population size or geography, and sometimes more temporally granular population estimates are desired for which no components of change exist.

Alternative, indirect population estimation methods exist to produce population estimates in situations when a traditional cohort-component method cannot be readily deployed. One such useful method is the Censal Ratio Method (Bogue 1950) which has both a straight-forward interpretation and requires very little data. The method assumes that population change between two periods occurs in the same ratio as a given symptomatic indicator – ideally with a symptomatic indicator correlated with the population total. Traditionally, this is accomplished with vital registration data but symptomatic indicators are not limited to vital events. Mathematically,

$$P_t = \frac{S_t}{S_{t-1}/P_{t-1}}$$

Where S_t is a given symptomatic indicator at time t. S_{t-1}/P_{t-1} is simply the ratio (r) of the symptomatic indicator to the population. This equation could also be written as $P_t = S_t/r$.

Such symptomatic indicators could be vital event registrations, school enrollments, voter registration, drivers licenses, electrical usage, etc., making the method attractive for estimating populations where the more data-hungry Cohort-Component is not feasible. The more closely a symptomatic indicator correlates with the population to be estimated, the better the method performs. Further refinements of the Censal Ratio Method exist, including regression-based approaches (Bryan 2004) and the inclusion of uncertainty (Swanson et al. 2012).

Consider the following example using voter registration data for Alabama Counties[1]. Using voter registration data in 2010 and 2019, we could estimate the population in Alabama Counties using only population data for 2010, assuming that the relationship between voter registration and total population is the same in both 2010 and 2019. For example, Autauga County Alabama had 33,885 registered votes in 2010, a County population total of 54,773 at Census 2010, and had 39,986 registered voters in 2019. Assuming the relationship between registered voters and total population remained unchanged in 2019 compared to 2010, the Censal Ratio Method would estimate Autauga's 2019 population as 64,635.

$$r = \frac{S_{t-1}}{P_{t-1}} = 0.6186 = \frac{33,885}{54,773}$$

$$P_t = \frac{S_t}{r} = 64,635 = \frac{39,986}{0.6186}$$

Oftentimes a symptomatic indicator is used at face value to estimate a population when it might seem closely correlated with population totals. For example, announced attendance for a sporting event could be closely correlated with the daytime population during that sporting event, the net difference between in-bound and out-bound flight passengers might be closely correlated with net migration, or the number of social media users in a given area might be closely correlated with the population total of that area. But symptomatic indicators rarely perfectly correlate with population values and using the symptomatic indicator alone will lead to erroneous estimation.

Figure 1 shows the Censal Ratio Method for all Alabama Counties and shows such erroneous estimation. Voter registration is closely correlated with county populations in the United States but using only voter registration would lead to an underestimation of Alabama county populations (**Figure 1a**). However, by incorporating the ratio of voter registration to the total population, our population estimates using the Censal Ratio Method perform considerably better (**Figure 1b**).

Thus, the Censal Ratio Method is a simple, useful, and intuitive approach to estimating populations when a Cohort-Component approach is not feasible.

---

[1]Data accessed 8/08/2023 via `https://www.sos.alabama.gov/alabama-votes/voter/election-data`
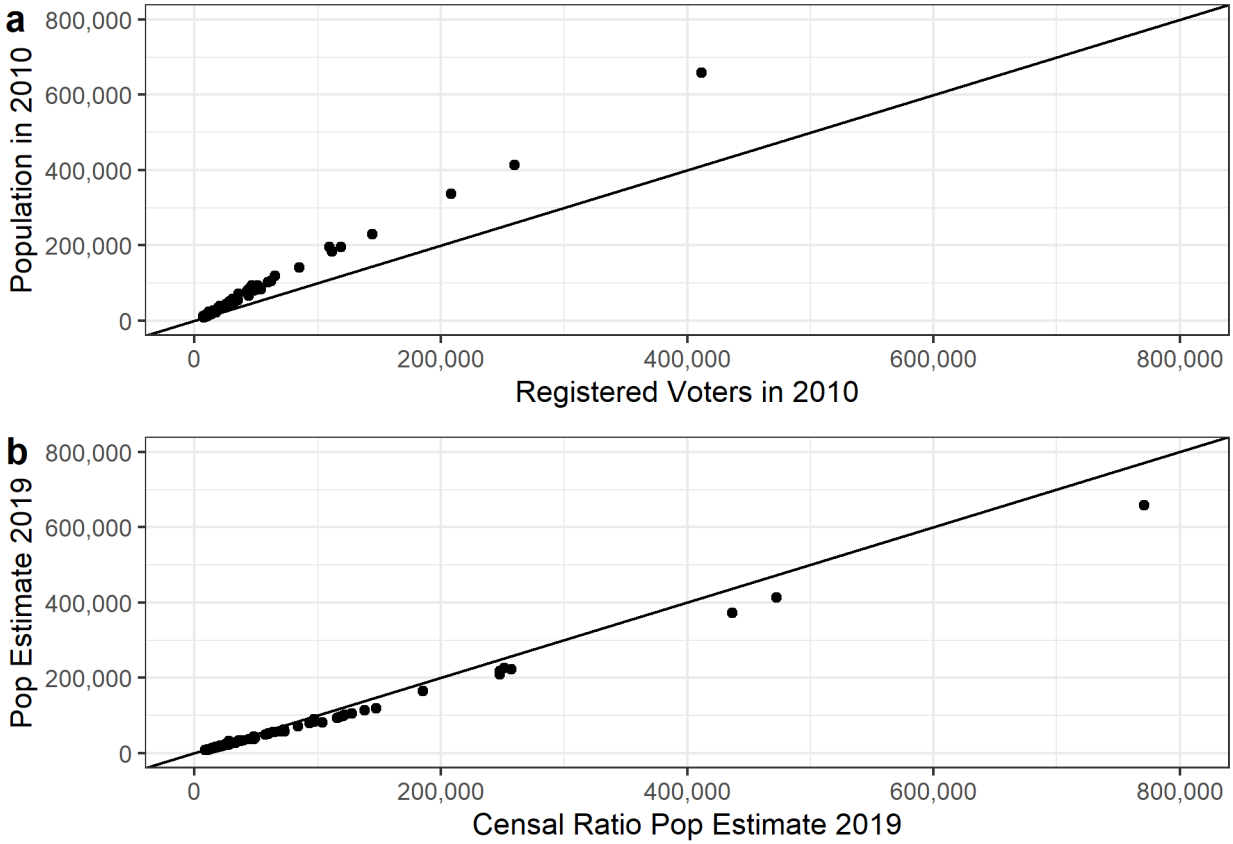
Figure 1: **Relationship between registered voters and population in Alabama Counties.** Solid diagonal lines are y=x. (a) shows the relationship between registered voters and population totals in Alabama Counties in 2010. Notice that there are always fewer registered voters in each county than total people. Strictly using a symptomatic indicator – voter registration in this case – would yield an underestimate of populations. (b) shows the relationship between county populations in 2019 and estimates produced using the Censal Ratio method. Notice how the Censal Ratio Method produces much more accurate estimates than the symptomatic indicator alone.

References

Bogue, D. J. (1950), 'A technique for making extensive population estimates', *Journal of the American Statistical Association* **45**(250), 149–163.

Bryan, T. (2004), 'Population estimates', *The methods and materials of demography* **2**, 523–560.

Gerland, P. (2014), 'Un population division's methodology in preparing base population for projections: case study for india', *Asian Population Studies* **10**(3), 274–303.

Swanson, D. A., Tayman, J., Swanson, D. A. & Tayman, J. (2012), 'Censal-ratio methods', *Subnational Population Estimates* pp. 187–194.