S3. Intercensal cohort method

One method for diagnosing age exaggeration that does not require a validated external data source is what Palloni, Pinto, and Beltrán-Sánchez (2015) describes it as the "Preston Index," which seems to be an extrapolation of what Elo & Preston (1994) term the "intercensal cohort method." In the absence of net migration, the population for cohort z at the 2nd census (P_{z,t_2}) should be equivalent to the population of that cohort at the 1st census (P_{z,t_1}) minus the deaths (D_z) that occurred to that cohort between t_1 and t_2 (Elo and Preston 1994, pp. 430-431). Thus, the ratio of the population cohort z counted by the 2nd census to the expected population should be 1.0:

$$\frac{P_{z,t_2}}{\hat{P}_{z,t_2}} = \frac{P_{z,t_2}}{P_{z,t_1} - D_z} \sim 1.0 \ .$$

We compute this ratio for each intercensal interval separately by sex for cohorts aged 65-69, 70-74, and 75-79 at the end of each interval. Palloni , Pinto, and Beltrán-Sánchez (2015, p. 28-29) notes that a ratio less than 1.0 could indicate age exaggeration more so in the population counts than in the deaths, whereas values greater than 1.0 could indicate age exaggeration more so in the deaths than in the population counts OR age exaggeration in both sources. Unfortunately, it can be difficult to ascertain the reasons for the observed inconsistencies in the ratios (Elo and Preston 1994). Many factors can affect this ratio, including net migration, under-registration of deaths, incomplete enumeration in one or both censuses, and age misreporting (in deaths and/or census counts). Consequently, the only condition under which we can clearly attribute a deviation in the ratio from 1.0 to age misreporting is if we have complete death registration, complete population enumeration, and accurate data on net migration (or a closed population). Any country in which age misreporting is a serious problem is unlikely to be free from problems in these other areas. Furthermore, even in the absence of problems other than age misreporting, this method cannot diagnosis the absolute level of age misreporting in deaths and population counts; rather, it indicates only whether there is age misreporting in population counts *relative to* deaths. As Palloni, Pinto, and Beltrán-Sánchez (2015) acknowledge, age exaggeration in both sources could offset one another producing a ratio close to 1.0.

Analysis of the consistency between death counts and population estimates at older ages suggest some inconsistencies, particularly for the 1950-63 inter-censal interval (Table S1). For example, among the cohorts aged 75-79 in 1963, the census counted 6.2% fewer women and 4.3% fewer men than expected based on the 1950 census count and the recorded deaths during 1950-63 for those cohorts. As noted by Palloni , Pinto, and Beltrán-Sánchez (2015, p. 28-29), a ratio less than 1.0 could indicate age exaggeration more so in the population counts than in the deaths, but it could also reflect under-registration of deaths, population coverage that is *less* complete in the 2nd census compared with the 1st census, or net out-migration. Estimates from INEC-CCP indicate that the census coverage *improved* between 1950 (6.5% omission) and 1963 (3.2% omission), so that would not appear to explain the subpar ratios (Gómez Barrantes 1972; Rosero-Bixby 2004). Neither does net out-migration appear to be the source of the inconsistency because levels of migration were very low during this period. Nonetheless, it is unclear whether the inconsistency stems from age exaggeration or from under-registration of deaths.

By the 1973-84 interval, ratios were close to 1.0, but in the latest interval (2000-11) the ratios are somewhat lower for women (3-4% fewer than expected). Census coverage deteriorated between the 2000 (2.9% omission, Instituto Nacional de Estadistica y Censos and Centro Centroamericano de Población 2002) and 2011 (6.2% omission, Centro Centroamericano de Población 2013) censuses, which could explain ratios less than 1.0 during that interval. Yet, the fact that ratios were below 1.0 only for women, but not men seems unlikely if changes in census coverage is the reason. Given that death registration is thought to be virtually complete in these recent years (Solano 2009) and net migration is very low at these ages, the results may indicate age exaggeration. But, why there would be more exaggeration among women than men is unclear given that rates of literacy (a factor typically driving the quality of age reporting) are close to 100% for both sexes (World Economic Forum 2017).

Table S1. Ratios of Counted to Expected Population Based on the Intercensal Cohort Method, Costa Rica

Age at the end of the interval[a]	1950-63	1963-73	1973-84	1984-2000	2000-11
			Females		
65-69	0.976	0.978	0.989	1.015	0.969
70-74	0.958	0.995	0.981	1.013	0.965
75-79	0.938	0.979	0.961	1.017	0.961
			Males		
65-69	0.982	0.982	0.989	1.019	1.014
70-74	0.947	0.977	0.984	1.016	1.014
75-79	0.957	0.991	0.970	1.012	1.009

[a] Age range refers to the age of the cohorts at the end of the inter-census interval. At the beginning of the interval, those cohorts would be *n* years younger, where *n* is the length of the interval. For example, the cohorts aged 75-79 in 2000 would have been aged 59-63 in 1984.

Note: Ratios between 95% and 105% are highlighted in green, values of 90-94.9% or 105.1-109.9% are highlighted in yellow, and values below 90% or above 110% are highlighted in red.

References

Centro Centroamericano de Población. (2013). *Evaluación demográfica del X Censo Nacional de Población de Costa Rica 2011 y de otras fuentes de información*. San José, Costa Rica: Centro Centroamericano de Población, Universidad de Costa Rica.

Elo, I.T. and Preston, S.H. (1994). Estimating African-American mortality from inaccurate data. *Demography* 31(3): 427-458.

Gómez Barrantes, M. (1972). *El descenso de la fecundidad en Costa Rica (Primera parte)*. San José, Costa Rica: Universidad de Costa Rica, Escuela de Ciencias Económicas, Cuidad Universitaria Rodrigo Facio.

Instituto Nacional de Estadistica y Censos and Centro Centroamericano de Población. (2002). *Costa Rica: Estimaciones y proyecciones de poblacion 1970-2050 actualizados al año 2000 y evaluacion del censo del 2000 y otras fuentes de informacion*. San José, Costa Rica: INEC.

Palloni, A., Pinto, G., and Beltrán-Sánchez, H. (2015). Chapter 2: Estimation of Life Tables 1850-2010: Adjustments for Relative Completeness and Age Misreporting. In: *Two Centuries of Mortality Decline in Latin America: From Hunger to Longevity*. Madison: University of Wisconsin: 3-75.

Rosero-Bixby, L. (2004). La explosión demográfica. In: Rodríguez-Vega, E. (ed.). *Costa Rica en el Siglo XX Tomo II*. San José, Costa Rica: Editorial Universidad Estatal a Distancia: 241-255.

Solano, E. (2009). *Evaluacíon Censal con Métodos Indirectos. Experiencia de Costa Rica.* Paper presented at Conferencía Estadistica de las Américas – CEA CEPAL, Santiago, Chile, June 3-5 June 2009.

World Economic Forum. (2017). *The Global Gender Gap Report 2017*. Geneva, Switzerland: World Economic Forum.