Title

The Accuracy of Hamilton–Perry Population Projections for Census Tracts in the United States

Journal

Population Research and Policy Review, 40(6)

ISSN

0167-5923

Authors

Baker, Jack
Swanson, David
Tayman, Jeff

Publication Date

2021-12-01

DOI

10.1007/s11113-020-09601-y

Peer reviewed

The Accuracy of Hamilton–Perry Population Projections for Census Tracts in the United States

Jack Baker, David Swanson & Jeff Tayman

Population Research and Policy Review
in cooperation with the Southern Demographic Association (SDA)

ISSN 0167-5923

Popul Res Policy Rev
DOI 10.1007/s11113-020-09601-y

ONLINE FIRST

Springer

Population Research and Policy Review
https://doi.org/10.1007/s11113-020-09601-y

ORIGINAL RESEARCH

The Accuracy of Hamilton–Perry Population Projections for Census Tracts in the United States

Jack Baker[1] · **David Swanson**[2,3] · **Jeff Tayman**[4]

Received: 3 March 2020 / Accepted: 21 July 2020

Abstract

In a first-ever nation-wide census tract evaluation, we assess the accuracy of the Hamilton–Perry population projection method for 65,221 census tracts. We started with 73,607 census tracts but eliminated those for which zeros appeared in age/sex groups. The test uses 1990 and 2000 census tract data by age and gender to construct cohort change ratios, which are then applied to 2000 census tract data to generate 2010 Hamilton–Perry projections that are evaluated in an ex post facto test against the reported 2010 census tract data by age and gender. The projections include: (1) uncontrolled age and gender projections; and (2) age and gender projections controlled to a projection of the population total by census tract. Mean Absolute Percent Error (MAPE) is used to evaluate precision and Mean Algebraic Percent Error (MALPE) is used to evaluate bias. We find that controlling the Hamilton–Perry projections by age for each tract to the linearly projected total population of each tract reduces both MAPE and MALPE within age groups by gender and for total females and total males. As this result suggests, simple linear extrapolation provides more accurate projections of the total population than does the Hamilton–Perry Method. However, even with controlling we find the Hamilton–Perry projections by age to be biased upward. Finally, we use MAPE-R (MAPE- Revised) to evaluate the effect of extreme outliers and find that high MAPEs in the uncontrolled projections are largely driven by extreme errors (outliers) found in less than 1 percent of the 65,221 census tracts used in the study.

Keywords Small area · Projection · Extrapolation · Cohort change ratios · National test

✉ David Swanson
dswanson@ucr.edu

1 Transamerica Life, Cedar Rapids, IA 52404, USA

2 University of California Riverside, Riverside, CA 92521, USA

3 Center for Studies in Demography and Ecology, University of Washington, Seattle, WA 98195, USA

4 University of California San Diego, La Jolla, CA 92093, USA

Published online: 20 September 2020

Introduction

Among many other uses, sub-county population projections are important for public safety assessments (Tayman et al. 1997), land use, environmental, and transportation planning (Waddell 2002), siting business establishments (Ambrose and Pol 1997; Siegel 2002, pp. 220–287), church membership analysis (Cline 2008), determining K-12 staffing and facility needs (Swanson et al. 1998), political re-districting (Morrison and Bryan 2019), and healthcare planning (Clark and Rees 2017; Pol and Thomas 2013). However, sub-county population projections present problems not found in projections for larger areas (Baker et al. 2017; Smith and Shahidullah 1995; Swanson et al. 2010). This is especially the case if age and gender data are desired because some or all of the "components" (fertility, mortality, and migration) required by the standard approach to population projection, the cohort-component method (CCM) (Smith et al. 2001), are often not available and components that are available at higher levels of geography often do not fit well at lower levels (Baker et al. 2017; Swanson et al. 2010). Without these components, it is not possible to move "cohorts" through time using the CCM.

Three alternatives are usually considered when fertility, mortality, and migration data are not available to fully implement the CCM: (1) structural modeling (2) Time Series modeling; and (3) simple and ratio extrapolative methods (Smith and Shahidullah 1995; Smith et al. 2001). Unlike the CCM, none of these methods requires age-specific data on births, deaths, and migration. However, structural modeling requires a wide range of social, economic, and demographic data while time series modeling requires data for many points in time (Tayman 1996). As Smith and Shahidullah (1995) observe, such data are seldom available for small sub-county areas. Thus, sub-county population projections are usually implemented with the third alternative, simple and ratio extrapolation methods, which have minimal data requirements. However, using them typically limits the range of demographic characteristics that can be projected (Smith et al. 2001, pp. 161–183).

However, all is not lost if one desires a range of demographic characteristics in a set of sub-county projections. While direct measures of the components of population change are difficult to implement at the sub-county level, these same components are indirectly measured in cohort change ratios (CCRs), a less-data intensive variation of the cohort-component method. One strong advantage of the CCR method is, that as is the case for the CCM, it also is algebraically equivalent to the fundamental population equation (Baker et al. 2017, pp. 100–101).

The construction of CCRs requires only data by age (and by race, ethnicity, gender, and other ascribed characteristics, as desired) from two successive population counts to implement (Baker et al. 2017), which fortuitously are usually available for sub-county areas in the form of census counts. Once constructed, CCRs provide the input for what is known as the "Hamilton–Perry Method" (HP) for projecting population (Hamilton and Perry 1962; Smith et al. 2001). It consists of two steps. In the first, CCRs are constructed and in the second, they are applied to age groups, which generate population projections by age. The HP method not only can be easily extended to age-based projections that include gender, race, ethnicity and other ascribed characteristics, such as nativity, but it is highly useful for doing small-area projections because of its minimal data requirements. With some adjustments, it also can be extended to include achieved characteristics such as marital status, educational attainment, and tobacco use (Baker et al. 2017, pp. 119–141).

A major strength of the HP method is it requires only census (or survey or population registry) data collected by age at two points in time. In the United States, for example, which conducts a census every 10 years, the first step in using the HP method would typically be used to capture population change over a 10-year period using CCRs constructed from the two most recent census counts. As the term CCR implies, one finds the ratio for a given cohort by dividing it into the number of people in this same cohort 10 years later. For example, the cohort of those aged 0–4 in 2000 ($({}_5P_0)$2000) becomes those aged 10–14 in 2010 ($({}_5P_{10})_{2010}$), the 10-year CCR between 2000 and 2010 for this cohort is: ${}_{10}\mathrm{CCR}_{0\text{–}4} = ({}_5P_{10})_{2010}/({}_5P_0)_{2000}$.

Because CCRs for those aged less than the number of years between the time of the most recent census (e.g., 2010) and the preceding census (e.g., 2000) cannot be computed (due to the fact that these individuals are not in the preceding census, their dynamics are modeled using child/adult ratios (Swanson and Tayman 2012; Smith et al. 2013; Baker et al. 2017). For a census that takes place every 10 years, the child/adult ratio for those aged 0–4 in 2010, for example, would be formed by using those aged 0–4 in the denominator and the adults likely to be their parents in the denominator, e.g., those aged 20–34: $({}_5P_0)_{2010}/({}_{15}P_{20})_{2010}$. The ratio for those aged 5–9 in 2010 would then be $({}_5P_5)_{2010}/({}_{15}P_{25})_{2010}$. A commonly used child/adult ratio is the child-woman ratio, which divides the population ages 0–4 and 5–9 by females ages 15–44 and 20–49, respectively. If a census takes place every 5 years, then one only needs the child adult ratio for those children age 0–4 years.

Also, there is typically a terminal, open-ended age group in the data used to construct CCRs. This is dealt with in the same manner as finding a survivorship ratio for a terminal, open-ended age group in a life table. For example, the CCR for those 75 years and over in the 2000 census is formed by taking ratio in which those aged 75 years and over become the denominator and those 85 years and over in the 2010 census become the numerator, ${}_{10}\mathrm{CCR}_{75+} = ({}_\infty P_{85})_{2010}/({}_\infty P_{75})_{2000}$. For details see Baker et al. (2017, p. 3) and Swanson et al. (2010).

In order to proceed to the second step and fully implement the HP method, the age data need to be organized in age groups such that the length of the input data cycle (e. g., 10 years) is exactly divisible by the width of the age groups (e. g., 5 years). For example, with decennial census data (e.g., the U.S., Mexico, Argentina), the age groups can be 10 years, 5 years, or 1 year in width. With a quinquennial census cycle (e.g., Australia, Canada, and the United Kingdom), 1-year and 5-year age groups meet this requirement. Because CCRs are applied against a base population, the projection cycle will be equal to the number of years between the points in time that the CCRs were constructed. For decennial data, the projection cycle is 10 years, for quinquennial data the projection cycle is 5 years.

In a country with a population registry system (e.g., Estonia, Finland, and Sweden), it is possible to implement the HP method using single-year of age data taken from the register in two successive years, which would be used to produce projections running in a single-year cycle. This can also be done where annual population estimates by age are available, as is the case in the United States where these estimates are done by the Census Bureau for the nation as a whole, states, and counties. If 5 year age groups were desired in the projections from data collected on an annual basis, then one would use register or estimated data that were 5 years apart (e.g., 2014 and 2019) to construct the cohort change ratios, which would maintain the rule of exact divisibility and produce projections running in a 5-year cycle. Further details of the method, as well as a wide variety of potential applications across many relevant areas of interest in the demographic sciences, are covered in Swanson and Tayman (2012), Smith et al. (2013), and Baker et al. (2017).

In this paper, we use the HP method to develop 10-year population projections (2000–2010) at the level of U.S. census tracts and then evaluate them in terms of *ex post facto* accuracy against the 2010 decennial census counts. While previous research has comprehensively-characterized the accuracy of the HP method and its variants for states and counties in Florida (Smith and Tayman 2003) and U.S. counties (Sprague 2012; Hauer 2019), this paper is the first to report a comprehensive, nation-wide evaluation of projection error for *census tracts* within the United States. As such, it fills an important gap in the literature around the method and a benchmark against which improvements in the method for small-area demographic analysis may be compared. We evaluate these HP projections from the standpoint of accuracy using two dimensions, precision and bias, measured, respectively, by Mean Absolute Percent Error (MAPE) and Mean Algebraic Percent Error (MALPE) (Smith et al. 2013; Rayer and Smith 2014), which we define later. Before turning to the evaluation, we provide some background in the form of a review of relevant literature.

Background

Even where the CCM can be properly implemented, there is no evidence that it is more accurate than the HP method (Smith 2017, Smith and Tayman 2003, Smith et al. 2001, pp. 307–316).[1] Thus, its simplicity, minimal data requirements, flexibility, and algebraic equivalency with the fundamental demographic equation suggest that the HP method should be given more attention, (Baker et al. 2017, pp. 101, 102, 247–250). Adding to this point, the HP method has been shown to provide a comprehensive and tractable approach to the stable population dynamics, which lends it some theoretical utility (Swanson et al. 2016).

At the level of states, the HP method appears to perform as well as the CCM (Smith and Tayman 2003; Smith et al. 2013; Swanson and Tayman 2017). At this

[1] Even more broadly, Green and Armstrong (2015) argue that for all types of forecasting, there is no evidence that complex methods increase accuracy.

level, HP errors ranging between less than 4% and ~ 12% for ages 10–85 + and between 20 and 25% for the 0–4 and 5–9 age intervals are observed (Smith and Tayman 2003; Swanson and Tayman 2017). At the level of U.S. counties or equivalents, the HP method has generally performed well, with errors ranging from 6% to 16% reported in previous comprehensive, nation-wide evaluations in both the U.S. (Sprague 2013; Hauer 2019) and Australia (Wilson 2016).

At the sub-county level, the HP method has been argued to be particularly attractive because it has significant potential to avoid pitfalls associated with the use of incompletely-geocoded administrative data sources (Baker et al. 2013, 2014, 2017). As such, it really has little competition from the CCM. Thus, the question is how well does the HP method perform at the sub-county level?

At the level of U.S. census tracts, the HP method has been evaluated in a number of studies; however, in each case these evaluations were limited in geographic scope, usually to specific states including Florida (Smith 1987; Smith and Shahidullah 1995) and New Mexico (Baker et al. 2013, 2014, 2017, pp. 59–82). In each of these cases, where age/gender breakouts have been considered specifically, the method has displayed average errors ranging between as low as 10% and as high as 70% or more, falling consistently in the 20–60% range (Smith and Shahidullah 1995; Baker et al. 2013, 2014, 2017, pp. 59–82). Moreover, authors have consistently pointed to the presence of an uncomfortable number of significant, large-scale outliers that in practice would require some form of case-by-case resolution to produce acceptable projections for use in policy analysis or other forms of decision-making that they are typically designed to support (Smith et al. 2013; Swanson et al. 2010; Baker et al. 2014).

As inferred by the increasing reported errors as previous studies move from state to county to census tract population projections, it is well-known that projection errors increase as population size decreases (Smith et al. 2013; Swanson and Tayman 2012). Population projections are also known to be associated with rates of intercensal population change, but not always in straightforward ways (Tayman et al. 2011). Projection error might be expected to be u-shaped with respect to the rate of change: those tracts which grow or shrink most during the intercensal period should be expected to tend to show the greater errors (Smith et al. 2013; Swanson and Tayman 2012; Baker et al. 2017, pp. 58–82; Smith and Shahidullah 1995). This pattern makes sense for a method based on historical trends that may shift in complex ways between intercensal periods (Smith et al. 2013; Smith and Shahidullah 1995; Wilson 2016; Baker et al. 2014). As with other projection methods, we also anticipate that projections made using the HP method will become less accurate as the length of the horizon is increased (Smith et al. 2013).

Though we are unaware of any study that examines the relationship between base population size and projection errors within age/gender intervals at the tract level, one study exists that examines the relationship between rate of intercensal population change and errors within age/gender breakouts at the tract level (Baker et al. 2013). Although limited to tracts within four counties in New Mexico ($n = 182$), the results of this analysis do not suggest a relationship between rate of intercensal population change and errors within age/gender groups. A more recent look at 2010 New Mexico census tracts ($n = 499$) overall (Baker et al. 2017, pp. 66, 67) showed that Hamilton–Perry methods that modified CCRs and child/adult ratios provided lower average errors within age/gender groupings. This suggests that the same patterns may be observed in a fuller-sample of census tracts and, indirectly, raises the point that shifts in intercensal dynamics can increase anticipated projection errors.

Taken together, both the paucity of studies aimed at quantifying projection errors for small geographies such as census tracts as well as the limited geographic scope of existing evaluations challenge the generalizability of previous evaluations. Given the importance of small-area projections to applied demographic practice, a further and more comprehensive evaluation of expected errors is a high-priority for further research.

Data and Methods

This study was made possible by the use of the National Historic GIS Database (NHGIS), housed at the University of Minnesota and made freely available to researchers at https://www.nhgis.org (Ruggles et al. 2010). This source provides population counts in 5-year age/gender-specific groupings for all 2010 census tracts with data for 1990 and 2000 normalized to these geographies using a simple method of block-level re-aggregations and interpolations. The method is described in detail at: https://www.nhgis.org/documentation/time-series.

Starting with these normalized tracts, we excluded census tracts in Puerto Rico and "island areas" (American Samoa, Guam, Commonwealth of the Northern Mariana Islands and U.S. Virgin Islands), which yielded 73,607 tracts with age/gender data in 1990, 2000, and 2010. Using these data, we then eliminated any tract for which zero was reported in one or more age sex groups. This means that if any zeros were found in age/sex groups for a given tract in any of the three decennial years, 1990, 2000, and 2010, it was eliminated from this study, a point we consider in the discussion section of this paper. This reduced the number of tracts from 73,607 to 65,221.

We then created CCRs for each of these 65,221 tracts based on 1990 to 2000 populations by age and gender cohort and then used them to project their respective 2000 census tract populations forward to 2010 for ages 10–85 +. To project populations for intervals including 0 to 4 and 5 to 9 years, we utilized the 2000 child/adult ratios for each census tract. This is to say that we used the standard algorithm for the HP method (Baker et al. 2017; Smith et al. 2013), which was constructed without population controlling methods (Judson and Popoff 2004) or special adjustments (Swanson and Tayman 2012; Smith et al. 2013) of any kind. It is acknowledged that this reflects the simplifying assumption that rates of change for the 1990–2000 period, as well as the 2000 child/adult ratios, reflect the dynamics of the 2000–2010 period (Baker et al. 2017, pp. 59–82; Tayman and Swanson 2017).

Projected values for 2010 were compared to observed 2010 census counts using mean absolute (MAPE) and algebraic (MALPE) percentage errors that are comparable to previous studies of ex post facto error in the applied demography literature (Tayman et al. 2007; Swanson and Tayman 2012; Tayman and Swanson 2016). Percentage errors allow us to summarize errors while negating the known effects of population size on ex post facto error calculations (Tayman et al. 1997; Smith 1987; Swanson and Tayman 2012).

MAPE and MALPE respectively capture two dimensions of projection error—precision and bias (Swanson 2015; Swanson et al. 2011). Error is defined as the difference between the projection and a census count: Projection–Census. MAPE, consequently, is defined as the mean of the summed absolute relative errors expressed in percentage terms, [|Projection–Census|/Census] × 100. The mean absolute percent error (MAPE) measures projection precision in which positive and negative errors do not offset each other. It shows the average percentage difference between the projected and observed population, ignoring the sign of the error.

The mean algebraic percent error (MALPE) is a measure of bias in which positive and negative values offset each other. It is defined as the mean of the summed relative errors expressed in percentage terms, [(Projection–Census)/Census] × 100. A positive MALPE reflects the tendency for the projections to be too high on average and a negative MALPE reflect the tendency for the projections to be too low on average.

Because we anticipated in advance that outliers may be a factor, we also report MAPE-R, a measure designed explicitly to minimize the effect of outliers on MAPE (Coleman and Swanson 2007; Swanson et al. 2018, 2011; Tayman and Swanson 1999). MAPE-R is a rescaled mean that uses a Box-Cox power transformation to adjust the original absolute percent error distribution. Rescaling is designed to address the impact of outlying observations, while still preserving the valuable statistical properties of the mean. Comparing the difference between MAPE and MAPE-R allows us to evaluate the impact of outliers on MAPE. As found in Coleman and Swanson (2007), the definition of MAPE-R is

$$\text{MAPE–R} = \left(\frac{1}{n}\sum\nolimits_{i=1}^{n}\text{APE}_i^{\lambda}\right)^{1/\lambda} = M_n^{[\lambda]}(\mathbf{APE})$$

where $M_n^{[p]}(\mathbf{a}) = \left(\frac{1}{n}\sum\nolimits_{i=1}^{n} a_i^p\right)^{1/p}$ is the pth power mean of the vector $a = (a_1, \ldots a_n)'$ and **APE** is the vector of the APE, where APE = [|Projection–Census|/Census] × 100.

Finally, we use the HP method to generate projections under both "uncontrolled" and "controlled" conditions. In the former, the HP projections by age are summed to obtain the total population; in the latter, we control the age/gender projections to projections of the total population in each census tract generated using linear extrapolation. We used this approach because the literature suggests that "controls" have the potential to improve accuracy (Smith et al. 2013, pp. 259–272) and simple extrapolative methods have been found to yield reasonably accurate results at the county (Swanson and Hough 2012) and sub-county levels (Smith and Shahidullah 1995).

In terms of the controls, we apply the numeric change in the total population between 1990 and 2000 of each census tract to the total population in 2000 to generate a 2010 projection of the total population. That is, $P_{2010} = P_{2000} + \Delta$, where $\Delta = P_{2000} - P_{1990}$. We then multiply the total population generated by the extrapolation by the proportion in each age/gender group generated by the HP method, which

Table 1 Mean percentage errors by age and gender 10-year Hamilton–Perry forecasts (uncontrolled)

Age	Males			Females		
	MALPE	MAPE	MAPE-R	MALPE	MAPE	MAPE-R
0 to 4	23.7	37.9	19.8	23.4	37.9	19.8
5 to 9	22.6	35.8	20.2	21.9	35.5	20.0
10 to 14	18.0	30.9	17.0	17.4	30.5	17.1
15 to 19	19.9	29.8	15.4	17.4	30.2	15.9
20 to 24	18.1	35.4	18.8	18.1	34.6	19.5
25 to 29	26.0	42.4	21.7	21.2	38.4	21.0
30 to 34	24.5	43.4	20.8	17.1	36.2	19.0
35 to 39	19.7	35.9	17.5	13.4	28.9	16.3
40 to 44	13.9	29.1	14.8	9.9	23.9	13.5
45 to 49	10.4	24.2	12.7	8.7	21.1	12.1
50 to 54	8.6	21.9	11.9	7.8	20.1	11.5
55 to 59	6.6	21.3	12.3	6.8	20.2	11.7
60 to 64	7.4	22.0	12.7	8.2	21.4	12.1
65 to 69	7.7	24.6	13.5	9.6	23.9	12.6
70 to 74	7.2	26.2	14.2	7.4	22.7	13.0
75 to 79	2.0	25.6	16.2	9.2	27.2	14.1
80 to 84	9.8	38.6	19.8	13.4	34.6	17.4
85 Plus	12.5	50.7	26.9	15.1	44.1	23.0
Total	13.7	20.8	8.4	12.0	19.1	8.0

2010 Census Tracts ($n = 65{,}221$)

"controls" the HP age/gender projections to the total population generated by the extrapolation method. In 42 census tracts, the extrapolation method was taking the total population below zero so we fixed these at zero.

Results

Age/gender-specific MALPE's and MAPEs are presented in Tables 1 and 2. In Table 1, the HP projections are not "controlled" while in Table 2, they are controlled to projected population totals generated by linear extrapolation.

As seen in Table 1, the overall MALPE across all age groups for both males and females by age is positive, which indicates that the HP method is biased upward. It is 13.7% for males and 12.0% for females. In terms of precision, MAPE shows an average absolute percent error across all age groups of 20.8% for males and 19.1% for females. The highest MAPE for males and for females is found for those age 85+, 50.7% and 44.1%, respectively. A key characteristic is that the differences between MAPE and MAPE-R are large. This indicates the presence and influence of outlying errors in these data. For example, the overall MAPE for males is 20.8 while the corresponding overall MAPE-R is 60% lower at 8.4%. This and the other differences between MAPEs and their respective MAPE-Rs, both overall and by age groups for

Table 2 Mean percentage errors by age and gender, 10-year Hamilton–Perry forecasts (controlled)

Age	Males			Females		
	MALPE	MAPE	MAPE-R	MALPE	MAPE	MAPE-R
0 to 4	14.6	30.5	20.6	14.3	30.4	20.5
5 to 9	14.9	29.6	20.1	14.3	29.5	19.9
10 to 14	11.3	26.7	17.9	10.9	26.6	18.0
15 to 19	11.7	23.2	15.7	10.5	25.4	16.3
20 to 24	10.1	29.0	18.8	10.5	28.4	19.4
25 to 29	15.1	32.4	21.8	12.2	30.9	21.1
30 to 34	13.4	34.1	21.2	8.2	29.1	19.8
35 to 39	11.0	29.9	18.8	7.1	25.3	17.4
40 to 44	6.4	24.0	15.6	3.8	20.3	14.1
45 to 49	3.9	20.1	13.3	3.0	17.8	12.4
50 to 54	2.6	18.1	12.4	2.1	16.4	11.5
55 to 59	1.0	17.8	12.4	1.1	16.9	11.8
60 to 64	1.7	18.7	13.0	2.3	18.0	12.5
65 to 69	0.5	20.0	14.0	2.0	19.0	13.1
70 to 74	−0.4	21.5	15.0	1.9	20.1	13.9
75 to 79	−2.5	24.2	16.9	0.8	21.9	15.1
80 to 84	−1.3	30.5	21.0	3.4	27.6	18.6
85 Plus	1.9	42.7	27.8	15.1	44.1	23.0
Total	−6.1	15.9	10.3	−5.0	29.7	18.8

2010 Census Tracts ($n = 65{,}221$)

both males and females, clearly indicate that arithmetic averages as summary measures of HP projection error are driven by outliers. Specifically, we find that these errors are driven by extreme errors representing less than 1% of the observations.

In spite of the presence of extreme outlying errors, from the point of view provided by past studies, the uncontrolled HP projections are surprisingly accurate at the census tract level with few discernible differences between projections of males and females by age. These errors are, in fact, in the range of projections errors reported at state and county levels in previous evaluations (Swanson and Tayman 2012; Hogan and Mulry 2015; Smith 1987, 2013).

Still considering the uncontrolled HP projections, in general, MAPEs are higher in the age groups between 20 and 40. This indicates, not surprisingly, that (domestic) migration is a factor in the errors found for those in these age groups because they represent the ages of adults who are most likely to migrate within a given country (Bernard et al. 2014, pp. 214–217). Looking at the MALPEs, we see that for both sexes there is a distinct tendency for the highest average over-projection errors to be found in the younger age groups, 0 to 49, with lower over-projection errors found in the age groups over 50.

Turning to Table 2, we see that controlling the HP projections creates noticeable reductions in the MAPEs in every age group for both males and females. Generally, in both the uncontrolled and controlled projections, the errors for females are higher than those for males. This is likely due to the higher propensity of females to migrate (Hernández-Murillo et al. 2011, pp. 174, 175). However, the influence on the MAPEs of extreme errors still persists. For example, the MAPE for males aged 25–29 is reduced from 42.4% in the uncontrolled projections to 32.4% in the controlled projections while the corresponding MAPE-Rs are virtually unchanged at 21.7% and 21.8%, respectively. The overall MAPE for males declines from 20.8 to 15.9% in going from the uncontrolled to the controlled projections. Note, however, that for females the overall MAPE across all age groups increases from the uncontrolled to controlled projections, 19.1% versus 27.9%. Having reductions in each age group and an increase in the overall MAPE for females is an example of the Simpson-Yule paradox (Simpson 1951; Yule 1903) and is due to the differences in the size of the age groups. The same paradox occurs in looking at the overall MALPEs for both males and females in the controlled projections. For males, only three of the eighteen age-specific MALPEs are negative while the overall MALPE is negative (−6.1%); for females the MALPEs in all of the eighteen age groups are positive, but the overall MALPE is negative (−5.0%). We suspect this is drive by outliers, but along with other issues noted in the subsequent section, we believe that this is a future research project.

Discussion

The first point we discuss is the decision to eliminate any tract with one or more zeros reported in its age/sex groups in any of the three study years, 1990, 2000, and 2010. We did this largely because unlike Cohort Change Differences (CCDs) CCRs are not well-suited to deal with zeros because they are a ratio, a measure that is undefined when the denominator is zero. Even with the restriction imposed by having a zero-valued denominator, we prefer to use CCRs over CCDs the reasons provided by Baker et al. (2017, p. 6): Unlike a CCR, a CCD can become negative, not bounded on the lower end by zero. In terms of comparisons, when a CCR is greater than 1.00, its corresponding CCD will be positive and when a given CCR is less than 1.00, its corresponding CCD will be negative (less than zero). Given this and other properties (e.g., neither division nor subtraction is associative and commutative in terms of their mathematical properties), there is no major advantage in using a CCD compared to its corresponding CCR. In addition, dating back to its introduction by Hamilton and Perry (1962), ratios have been used by analysts rather than differences (Baker et al. 2017, p. 6). Finally, we note that the CCR method is not alone in regard to dealing with zeros in that neither does the Cohort-Component Method, especially in terms of age-specific rates representing the components of change, fertility, mortality, and migration.

Table 3 shows CCRs and Child/Adult ratios for both genders combined for the US at two points in time. It shows two related points: (1) that the 1990–2000 CCRs for age groups 10–14 to 50–54 are larger than the 2000–2010 CCRs for these same age groups; and (2) that the 1990–2000 CCRs for age groups 55–59 to 85+ are smaller than the 2000–2010 CCRs for these same age groups. These comparisons are consistent with the population aging experienced by the U. S.

Table 3 CCRs and child/adult ratios, both genders combined, United States, 1990–2000 and 2000–2010

Age	Population			Cohort Change and Child/ Adult Ratios[a]		1990–2000 ratio larger
	1990	2000	2010	1990–2000	2000–2010	
0 to 4	18,354,443	19,175,798	20,201,362	0.29511	0.32581	0
5 to 9	18,099,179	20,549,505	20,348,657	0.28666	0.32828	0
10 to 14	17,114,249	20,528,072	20,677,194	1.11843	1.07830	1
15 to 19	17,754,015	20,219,890	22,040,343	1.11717	1.07255	1
20 to 24	19,020,312	18,964,001	21,585,999	1.10808	1.05154	1
25 to 29	21,313,045	19,381,336	21,101,849	1.09166	1.04362	1
30 to 34	21,862,887	20,510,388	19,962,099	1.07834	1.05263	1
35 to 39	19,963,117	22,706,664	20,179,642	1.06539	1.04119	1
40 to 44	17,615,786	22,441,863	20,890,964	1.02648	1.01856	1
45 to 49	13,872,573	20,092,404	22,708,591	1.00648	1.00008	1
50 to 54	11,350,513	17,585,548	22,298,125	0.99828	0.99360	1
55 to 59	10,531,756	13,469,237	19,664,805	0.97093	0.97872	0
60 to 64	10,616,167	10,805,447	16,817,924	0.95198	0.95635	0
65 to 69	10,111,735	9,533,545	12,435,263	0.90522	0.92323	0
70 to 74	7,994,823	8,857,441	9,278,166	0.83434	0.85866	0
75 to 79	6,121,369	7,415,813	7,317,795	0.73339	0.76758	0
80 to 84	3,933,739	4,945,367	5,743,327	0.61857	0.64842	0
85 Plus	3,080,165	4,239,587	5,493,433	0.32276	0.33091	0
						9

[a] ${}_4P_{0,t-10}/{}_{34}P_{20,t-10}$ ages 0–4 (child/adult ratio), ${}_9P_{5,t-10}/{}_{39}P_{25,t-10}$ ages 5–9 (child/adult ratio), $P_{x,t}/P_{x-10,t-10}$ ages 10–84, $P_{85+,t}/P_{75+,t-10}$ ages 85+

from 1990 to 2010. They also are consistent with the MALPEs found in Tables 1 and 2 in that for both males and females, they are larger (and all positive in the younger age groups and smaller (with some being negative) in the older age groups. That is, the 1990-CCRs tend to over-project the younger age groups in 2010, but do less so with the older age groups.

In terms of the projections by age, the upshot of a nation-wide evaluation of the simplest possible 10-year HP projections against the 2010 census is that: (1) the uncontrolled HP projections are surprisingly accurate at the census tract level; (2) outliers drive summary measures of error; (3) errors are reduced when the projections by age are controlled to the total populations generated by linear extrapolation; and (4) extreme values play a major role in summary measures of error even when the projections are controlled. Given the significant presence of outliers in our nation-wide evaluation, we believe that MAPE-R provides a more accurate picture of precision than does MAPE.

The fact that summary measures of error are driven by outliers is not surprising. Smith et al. (2001, p. 159) note that the HP method can lead to unreasonably high projections in places where the CCRs were generated during a period of rapid growth and unreasonably low projections in places where the CCRs were generated during a period of rapid decline The cause of this is due to rapid growth that resulted in a tract being "filled in" with housing and population whereby it was no longer capable of rapid growth. Similarly, rapid loss can lead to a tract that is "emptied out" and no longer capable of rapid decline. Swanson et al. (2010) found that developing "ceilings" and "floors" mitigated these problems by placing limits on the projections. Our finding that controls reduce errors also is consistent with findings by Swanson et al. (2010). In part this is due to the fact that controls themselves place ceilings and floors on tracts that experienced rapid growth or rapid decline during the period when CCRs were constructed.

Applied demographers using this method should anticipate greater than previously-believed levels of accuracy and the ongoing challenge of upward bias and outliers for both producing publishable datasets and understanding the potential and pitfalls of the method. As noted earlier, we are encouraged by the fact that the levels of error of the uncontrolled tract projections are in the range of projections errors reported at state and county levels in previous evaluations (Swanson and Tayman 2012; Hogan and Mulry 2015; Smith et al. 1987, 2013). It is, indeed, a bonus that the controlled projections are even more accurate.

One caveat to our results is the issue of zero-valued observations. If tract level projections are being done for a limited area (e.g., a city or a county) where an analyst is very familiar with the demographic landscape, then it is likely the case that manual interventions can be used such as suggested by Swanson et al. (2010). With more resources, such interventions can be applied tracts representing a larger area (e.g., a state, region, or the US as a whole).

In regard to future research, this study indicates that initial population size and rate of intercensal population change should be considered. Future research may benefit from focusing on understanding the population characteristics of tracts and block-groups within the outlier sets defined here (e.g., size, growth rate, and special populations). Here, a comparison with other simple projection methods such as "shift-share" and geometric extrapolation would likely be useful. The former would be interesting because the tract level projections are, in effect, controlled to county level projections. Finally, the results shown here for census tracts also suggest that an evaluation of accuracy for block group HP projections is feasible. Hopefully, the "Differential Privacy" initiative for the 2020 census will not preclude the availability of sufficiently accurate data needed to implement these suggestions (IPUMS, no date).

References

Ambrose, D., & Pol, L. (1997). Motel 48: Evaluating the profitability of a proposed business. In H. Kintner, T. Merrick, P. Morrison, & P. Voss (Eds.), *Demographics: A casebook for business and government* (pp. 144–154). Santa Monica, CA: Rand Corporation.

Baker, J., Alcantara, D., Ruan, X., Watkins, K., & Vasan, S. (2013). A comparative evaluation of error and bias in census tract-level age/sex specific population estimates: Component I (net migration) v. Component III (Hamilton–Perry). *Population Research and Policy Review, 32*, 919–942.

Baker, J., Alcantara, D., Ruan, X., Watkins, K., & Vasan, S. (2014). Spatial weighting improves accuracy in small area demographic forecasts of urban census tract populations. *Journal of Population Research, 31*(4), 345–359.

Baker, J., Swanson, D., Tayman, J., & Tedrow, L. (2017). *Cohort change ratios and their applications*. New York: Springer.

Bernard, A., Bell, M., & Charles-Edwards, E. (2014). Life course transitions and the age profile of internal migration. *Population and Development Review, 40*(2), 213–239.

Clark, S., & Rees, P. (2017). The drivers of heath trends: A decomposition of projected health for local areas in England. In D. Swanson (Ed.), *The frontiers of applied demography* (pp. 23–40). New York: Springer.

Cline, M. (2008). A comparison of potential demographic influences on church membership of two United Methodist congregations in San Antonio, Texas. In S. Murdock & D. A. Swanson (Eds.), *Applied demography in the 21st Century* (pp. 237–249). New York: Springer.

Coleman, C., & Swanson, D. (2007). On MAPE-R as a measure of cross-sectional estimation and forecast accuracy. *Journal of Economic and Social Measurement, 32*(4), 219–233.

Green, K., & Armstrong, J. S. (2015). Simple versus complex forecasting: The evidence. *Journal of Business Research, 68*, 1678–1685.

Hamilton, C. H., & Perry, J. (1962). A short method for projecting population by age from one decennial census to another. *Social Forces, 41*, 163–170.

Hauer M. (2019). Population projections for U.S. counties by age, sex, and race controlled to shared socioeconomic pathways. *Scientific Data*. https://www.nature.com/articles/sdata20195.pdf.

Hernández-Murillo, R. L., Ott, M. O., & Whalen, D. (2011). Patterns of interstate migration in the United States from the survey of income and program participation. *Federal Reserve Bank of St. Louis Review, 93*(3), 169–185.

Hogan, H., & Mulry, M. (2015). Hogan: Statistical properties of different measures. In N. Hoque (Ed.), *Emerging techniques in applied demography* (pp. 119–136). New York: Springer.

IPUMS. No date. Changes to Census Bureau data products. https://ipums.org/changes-to-census-bureau-data-products.

Judson, D., & Popoff, C. (2004). Selected general methods. In D. Swanson & J. S. Siegel (Eds.), *The methods and materials of demography* (2nd ed., pp. 667–732). San Diego: Elsevier.

Morrison, P., & Bryan, T. (2019). *Redistricting: A manual for analysts, practitioners and citizens*. New York: Springer.

Pol, L., & Thomas, R. (2013). *The demography of health and healthcare* (3rd ed.). New York: Springer.

Rayer, S., & Smith, S. K. (2014). Population projections by age for Florida and its counties: Assessing accuracy and the impact of adjustments. *Population Research and Policy Review, 33*(5), 747–770.

Ruggles, S., et al. (2010). *Integrated public use Series: Version 5.0 [Machine-readable database]*. Minneapolis, MN: University of Minnesota.

Siegel, J. (2002). *Applied demography: Applications to business, government, law, and public policy*. San Diego, CA: Academic Press.

Simpson, E. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society Series B, 13*, 238–241.

Smith, S. (1987). Tests of forecast accuracy and bias for county population projections. *Journal of the American Statistical Association, 82*, 991–1003.

Smith, S. (2017). Foreword. In J. Baker, D. Swanson, J. Tayman, & L. Tedrow (Eds.), *Cohort change ratios and their applications* (pp. v–vi). New York: Springer.

Smith, S., & Shahidullah, M. (1995). An evaluation of population projection errors for census tracts. *Journal of the American Statistical Association, 90*(429), 64–71.

Smith, S., & Tayman, J. (2003). An evaluation of population projections by age. *Demography, 40*, 741–757.

Smith, S., Tayman, J., & Swanson, D. (2001). *State and local population projections: Methodology and analysis*. New York: Kluwer Academic/Plenum Press.

Smith, S., Tayman, J., & Swanson, D. (2013). *A practitioner's guide to state and local population projections*. New York: Springer.

Sprague, W. (2013). *Wood's Method: A method for fitting Leslie Matrices from age-sex population data, with some practical applications*, Unpublished Ph.D. Dissertation, University of California Berkeley.

Swanson, D. (2015). On the relationship among values of the same summary measure of error when it is used across multiple characteristics at the same point in time: An examination of MALPE and MAPE. *Review of Economics and Finance*, *5*(1), 1–14.

Swanson, D., & Hough, G. (2012). An evaluation of persons per household (PPH) data generated by the American Community Survey: A demographic perspective. *Population Research and Policy Review, 31*, 235–266.

Swanson, D., Hough, G., Clemans, C., & Rodriguez, J. (1998). Merging methods and judgment for K-12 enrollment forecasting. *Educational Research Service Spectrum, 16*(Fall), 24–31.

Swanson, D., Schlottmann, A., & Schmidt, R. (2010). Forecasting the population of census tracts by age and sex: An example of the Hamilton–Perry method in action. *Population Research and Policy Review, 29*, 47–63.

Swanson, D., & Tayman, J. (2012). *Subnational population estimates*. New York: Springer.

Swanson, D., & Tayman, J. (2017). A long-term test of the accuracy of the Hamilton–Perry method for forecasting state populations by age. In D. Swanson (Ed.), *The frontiers of applied demography* (pp. 491–513). New York: Springer.

Swanson, D., Tayman, J., & Bryan, T. (2011). MAPE-R: A rescaled measure of accuracy for cross-sectional, sub-national forecasts. *Journal of Population Research, 28*, 225–243.

Swanson, D., Tayman, J., & Bryan, T. (2018). A note on rescaling the arithmetic mean for right skewed positive distributions. *Review of Economics & Finance, 14*(4), 17–24.

Swanson, D., Tedrow, L., & Baker, J. (2016). Exploring stable population concepts from the perspective of cohort change ratios: Estimating the time to stability and intrinsic r from initial information and components of change. In R. Schoen (Ed.), *Dynamic demographic analysis* (pp. 227–258). New York: Springer.

Tayman, J. (1996). The accuracy of small area population forecasts based on a spatial interaction land-use modeling system. *Journal of the American Planning Association, 62*, 85–98.

Tayman, J., Parrott, B., & Carnevale, S. (1997). Locating fire stations: The response time component. In H. Kintner, T. Merrick, P. Morrison, & P. Voss (Eds.), *Demographics: A casebook for business and government* (pp. 203–217). Santa Monica, CA: Rand Corporation.

Tayman, J., & Swanson, D. (1999). On the validity of MAPE as a measure of population forecast accuracy. *Population Research and Policy Review, 18*, 299–322.

Tayman, J., Smith, S., & Lin, J. (2007). Precision, bias, and uncertainty for state population forecasts: An exploratory analysis of time series models. *Population Research and Policy Review*, *26*(6), 347–369.

Tayman, J., & Swanson, D. (2016). New insights on the impact of coefficient instability on ratio-correlation population estimates. *Journal of Economic and Social Measurement*, *41*(2), 121–143.

Tayman, J., & Swanson, D. (2017). Using modified cohort change and child-woman ratios in the Hamilton–Perry forecasting method. *Journal of Population Research, 34*(3), 209–231.

Tayman, J., Smith, S., & Rayer, S. (2011). Evaluating population forecast accuracy: A regression approach using county data. *Population Research and Policy Review, 30*(2), 235–262.

Waddell, P. (2002). UrbanSim: Modeling urban development for land use, transportation and environmental planning. *Journal of the American Planning Association, 68*(3), 297–314.

Wilson, T. (2016). Evaluation of alternative cohort-component models for local area population forecasts. *Population Research and Policy Review, 35*, 241–261.

Yule, E. (1903). Notes on the theory of association of attributes in statistics. *Biometrika, 2*(2), 134.