Title

On the Relationship among Values of the same Summary Measure of Error when used across Multiple Characteristics at the same point in time: An Examination of MALPE and MAPE

Journal
Review of Economics and Finance, 5(1)

Author
Swanson, David A

Publication Date
2015-04-01

Peer reviewed

Review of Economics & Finance
Submitted on 15/04/2015
Article ID: 1923-7529-2015-03-01-14 David A. Swanson

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[1]

Dr. ***David A. Swanson***
Department of Sociology, University of California Riverside
Riverside, California 92521, U.S.A.
Tel.: +1-951-827-4373 E-mail: dswanson@ucr.edu

Abstract: This paper deals with an issue that appears to be unexplored by demographers and others who conduct cross-sectional forecasts of populations with multiple characteristics. The issue is based on the question: "if one was conducting an ex post facto evaluation of a forecast or estimate that includes more than one characteristic (e.g., age, gender, race, and geography), how can one explain the fact that there are differences among the summary measures of error for all of the characteristics?" Using a hypothetical demographic forecast as an illustration, the paper examines this issue for two characteristics (race and geography) using a standard summary measure of forecast errors for each of two error dimensions, bias and precision. For the bias dimension, we examine the "Mean Algebraic Percent Error" (MALPE) and "Mean Absolute Percent Error (MAPE) for the precision dimension. The paper finds that hitherto unknown relationships across characteristics exist for each of the two summary error measures. For example, in evaluating a forecast that includes more than one characteristic, one can express MAPE taken for one characteristic (e.g., race) in terms of each of the other characteristics (e.g., age and geography). This finding allows one to explain the reason for differences in the summary error measures used to evaluate a forecast or estimate done for more than one characteristic. The findings are informative and suggestive and are likely to be generalizable. While the discussion is not formally rigorous, a formal general proof is presented of the relationship between values of MALPE across multiple characteristics as well as the values of MAPE.

Keywords: Forecasting; Errors; Summary measure reconciliation
JEL Classifications: B4, C8, and Z1

1. The Problem

Summary measures of error are part of the toolkit for those who evaluate forecasts and estimates. In the field of demography, the evaluations are often in the form of ex post facto assessments against census data (Chi and Voss, 2011; Chu, 1974; Coleman and Swanson, 2007; Long, 1995; Rayer, 2008; Shahidullah and Flotow, 2005; Smith and Sincich, 1992; Smith, Tayman, and Swanson, 2013: 324-328; Swanson, Tayman, and Barr, 2000; Tayman, 1996; Tayman and Swanson, 1999; Tayman, Swanson and Barr, 1999). The most commonly used summary measures of error are Mean Algebraic Percent Error – MALPE - and Mean Absolute Percent Error – MAPE

[1] This research was prompted by Dr. George Hough, the first person known by the author to ask how the differences among summary measures of forecast error taken across different characteristics of the same population could be explained.

(Smith, Tayman and Swanson, 2013: 324-328; Swanson and Tayman, 2011: 268-272). MALPE is used to measure bias – is the estimate or forecast high or low, on average; MAPE is used to measure precision – how close is the estimate to the benchmark numbers on average?

Their ubiquity testifies to the desirable properties of MALPE and MAPE. However, like all measures they have their shortcomings (Coleman and Swanson, 2007; Swanson and Tayman, 2011: 268-280; Swanson, Tayman and Barr, 2000; Swanson, Tayman, and Bryan, 2011; Tayman and Swanson, 1999; and Tayman, Swanson and Barr, 1999). One shortcoming that has not been addressed is one specific to forecasts and estimates done for a range of characteristics (e.g., age, gender, race, geographic area) at a given point in time. Perhaps, the best way to introduce this shortcoming is through an anecdotal description.

Suppose we have done forecasts back in 2002 for the year 2010 by race (using three race groups) for a state that has four counties. That is, we have done forecasts for two population characteristics, race and geography. The sum of the population by race is equal to the sum of the population by geography, both of which are equal to the grand total. We now want to do an ex post facto evaluation of the accuracy of our multi-characteristic forecast across the two major dimensions of error, precision and bias. To assess these two dimensions, we use, respectively, the two summary measures most commonly used for these purposes, Mean Algebraic Percent Error (MALPE), and Mean Absolute Percent Error (MAPE). As we will see in Sections III through V, when we evaluate our hypothetical multi-characteristic forecast, the MALPEs associated with these characteristic will not only be different from one another, but also different from the overall algebraic percent error. We also will find that this holds true for the MAPEs.

These differences in the same measure across the different characteristics may lead to questions that will be difficult to answer if one does not understand the nature of the relationship for a given summary error measure across a range of characteristics that roll up to a single grand total. This paper is aimed at providing this understanding. Before we turn to an explanation of these relationships, however, it is important to define the two summary error measures, MALPE and MAPE.

2. Summary Measures for the Two Major Dimensions of Error

Before turning to a definition of MALPE and MAPE, a definition of error is needed. One can define forecast error (E) as the difference between a given forecast (F) for a particular population and the 2010 census (CEN):

$$E = F - CEN \qquad (1)$$

The error will be positive when the forecast is larger than the census count and it will be negative when it "under-forecasts" the census. Errors are often expressed as percent differences rather than absolute differences. The use of percent errors is particularly helpful when making comparison across geographic areas. A forecast error of 2,000 has a very different meaning for a place with 20,000 residents than a place with 200,000 residents. The definition of Error given above can be broadened to include age, sex, race, and ethnicity. For example, the definition of error for a particular age group i would be

$$E_i = F_i - CEN_i \qquad (2)$$

Without adjustments for population size, errors for places with large populations (or age groups with large numbers of people) would dominate the effects of errors for places with small populations (or age groups with small numbers of people). Thus, the definition of error can be broadened to provide a "relative" perspective as follows:

$$ALPE = (E / CEN) * 100; \text{ and} \quad (3)$$

$$APE = |(E / CEN) *100 \quad (4)$$

ALPE (algebraic percent error) preserves the sign of the percent error; it has a theoretical minimum of -100% and no upper bound, while the APE (absolute percent error) has a minimum at zero and no upper bound. ALPE and APE represent individual forecast errors for the set of geographic areas under study and form a distribution of forecast errors. As before, the definition of relative error given above can be broadened to include age, sex, race, and ethnicity. For example, the definition of relative error for a particular age group i would be

$$ALPE_i = (E_i / CEN_i) * 100; \text{ and} \quad (5)$$

$$APE_i = |(E_i / CEN_i)| * 100 \quad (6)$$

Two common summary measures of relative error are MALPE and MAPE, both of which are arithmetic means. MALPE measures bias and MAPE measures precision. MALPE is defined as:

$$MALPE = \sum_i ALPE_i / n \quad (7)$$

where $ALPE_i$ is as defined above; n = number of groups.

MALPE is a measure in which positive and negative values offset each other (Smith, Tayman, and Swanson 2013: 324-328). Consequently, it is often used as an average measure of *bias*. A positive MALPE reflects the average tendency for forecasts to be too high and a negative MALPE reflects the average tendency for forecasts to be too low. A zero MALPE would indicate no bias in the set of forecasts, as the sum of the positive percentage errors would equal the sum of the negative percentage errors. A MALPE less than +5% and greater than -5% is considered as an indication that the forecast is not substantially biased. A MALPE greater than +10% but less than +25% or less than -10% but not less than -25% indicates considerable bias and one greater than +25% indicates substantial upward bias (the forecast is way too high) and one less than -25% indicates substantial downward bias (the forecast is way too low). MAPE is defined as:

$$MAPE = \sum_i APE_i / n \quad (8)$$

where APE_i is as defined above; n = the number of groups.

MAPE is a measure in which positive and negative values do not offset each other; it measures the precision of the forecasts by showing the average percent difference between forecasts and actual activities regardless of whether the individual forecasts were too high or too low. MAPE has several desirable properties including reliability; ease of use and interpretation. It also incorporates all of the information in its calculation, but MAPE has a major drawback. Like any average, MAPE is affected by extreme values, but in the case of MAPE, the extreme values most often occur at the high end of the distribution (Swanson, Tayman, and Bryan, 2011). Thus, the error distribution of the APEs is often asymmetrical and right-skewed because it is bounded on the left by zero and unbounded on the right. Therefore, MAPE is susceptible to being pulled upward and to overstating the error represented by most of the observations. Given this undesirable property, a zero MAPE would indicate perfect accuracy in the set of forecasts, as the sum of the positive percentage errors would equal the sum of the negative percentage errors. A MAPE less than 5% is considered as an indication that the forecast is acceptably accurate. A MAPE greater than 10% but less than 25% indicates low, but acceptable accuracy and MAPE greater than 25% very low accuracy, so low that the forecast is not acceptable in terms of its accuracy.

3. The Problem Revisited

Suppose we have done forecasts back in 2002 for the year 2010 by race (using three race groups) for a state that has four counties. That is, we have done forecasts for two population characteristics, race and geography. We now want to do an ex post facto evaluation of the accuracy of our forecasts across the two dimensions of error, bias and precision. To assess these two dimensions, we use, respectively, MALPE and MAPE.

Table 1 contains the results of our 2010 forecasts for the state by race and county, while Table 2 contains the actual 2010 population counts, Table 3, the numerical errors, Table 4, the percent errors, and Table 5, the absolute percent errors.

Note that in each table, 1 through 5, we have a 3 (race groups) by 4 (geography) table, with marginal totals for race in the fifth row and marginal totals for geography in the fourth column.

Table 1. 2010 Illustrative Forecast (F_{ij}) for the State by Race & Geography

Forecasted Population				
	race1	race2	race3	total
geog1	250	100	55	405
geog2	500	125	40	665
geog3	80	10	5	95
geog4	125	20	10	155
Total	955	255	110	1320

Table 2. 2010 Illustrative Census Counts (CEN_{ij}) for the State by Race & Geography

Census Population				
	race1	race2	race3	total
geog1	291	98	45	434
geog2	490	130	51	671
geog3	85	8	4	97
geog4	130	25	12	167
Total	996	261	112	1369

Table 3. Numerical Error (F_{ij}-CEN_{ij}) in the 2010 Forecast by Race & Geography

Numerical Error				
	race1	race2	race3	total
geog1	-41	2	10	-29
geog2	10	-5	-11	-6
geog3	-5	2	1	-2
geog4	-5	-5	-2	-12
Total	-41	-6	-2	-49

Table 4. Relative Error ((F_{ij}-CEN_{ij})/CEN_{ij}) in the 2010 Forecasts by Race & Geography

Relative Error				
	race1	race2	race3	total
geog1	-0.1409	0.0204	0.2222	-0.0668
geog2	0.0204	-0.0385	-0.2157	-0.0089
geog3	-0.0588	0.2500	0.2500	-0.0206
geog4	-0.0385	-0.2000	-0.1667	-0.0719
Total	-0.0412	-0.0230	-0.0179	-0.0358

Table 5. Absolute Relative Error |(F_{ij}-CEN_{ij})/CEN_{ij}| in the 2010 Forecasts by Race & Geography

Absolute Relative Error				
	race1	race2	race3	total
geog1	0.1409	0.0204	0.2222	0.0668
geog2	0.0204	0.0385	0.2157	0.0089
geog3	0.0588	0.2500	0.2500	0.0206
geog4	0.0385	0.2000	0.1667	0.0719
Total	0.0412	0.0230	0.0179	0.0358

Table 6. Summary Error by Characteristic

Characteristic	Algebraic Error	Absolute Error
RACE	-0.0274	.0274
GEOG	-0.0421	.0421
TOTAL ERROR	-0.0358	.0358

As was the case for Table 4, keep in mind that when at the "totals" for the rows and columns in Table 5 are not the sums of the errors across the rows and columns, respectively, but the result of finding the relative difference between the forecast for a given geographic area or race group and the census number. For example, the absolute relative error for race in geographic area 1 is 0.0668 = |-29/434|, while the absolute relative error for geographic within race group 1 is 0.1412 = |-41/996|.

Table 6 provides the MALPEs for Race (RMALPE) and Geography (GMALPE), and the TALPE, calculated from Table 4 as follows:

RACE MALPE (RMALPE) = -0.0274 = (-0.0412-.0230 -.0179)/3

Thus, RMALPE is found by summing the "total" relative errors for each of the three race groups across all four geographic areas and dividing the result by three.

GEOG MALPE (GMALPE) = -0.0421 = (-.0668-.0089-.0206-.0719)/4

GMALPE is found by summing the "total" relative errors for each of the four geographic areas across all three race groups and dividing the result by four.

TOTAL ALPE (TALPE) = -.03580 = [(1320-1369)/1369]

Finally, TALPE is found by simply subtracting the census population by the total forecasted population and dividing the difference by the census population.

In percent terms, we have Race MALPE = -2.74%, Geog MALPE = -4.21% and Total MALPE = -3.58%. Not only is it the case that none of the MALPEs is equal to another, but there is no clear relationship among them.

Table 6 provides the MAPEs for Race (RMAPE) and Geography (GMAPE), and the TAPE, calculated from Table 5 as follows:

RACE MAPE (RMAPE) = 0.0274 = (.0412+.0230+.0179)/3

Thus, RMAPE is found by summing the "total" absolute relative errors for each of the three race groups across all four geographic areas and dividing the result by three.

GEOG MAPE (GMAPE) = 0.0421 = (.0668+.0089+.0206+.0719)/4

GMAPE is found by summing the "total" absolute relative errors for each of the four geographic areas across all three race groups and dividing the result by four.

TOTAL APE (TAPE) = 0.0358 = [|(1320-1369)/1369|]

Finally, TAPE is found by simply subtracting the census population by the total forecasted population, dividing the difference by the census population and taking the absolute value of this difference. In percent terms, we have RMAPE = 2.74%, GMAPE = 4.21% and TAPE = 3.58%. As you can see, not only is it the case that neither of the two MAPEs is equal to one another or to the Total APE, but also, that there is no obvious relationship among these three measures.

Now imagine yourself in front of a group of "stakeholder" discussing these forecasts and you are describing these error measures (and how accurate your estimates are) when you are asked "why are the MAPEs different from one another" and "why are the MALPEs different from one another?" The answer is not apparent, likely because it appears that demographers have not yet addressed the issue of simultaneous measures of error on a multi-characteristic forecast or estimate. What follows is an initial attempt to answer the question, "Why are they different?" As you will see, the answer has to do with "weighting" in the case of MALPE and MAPE, and lack thereof. We start with MALPE

3.1 MALPE and Weighted MALPE

The calculations shown for the MALPEs in Table 6 reveal that the two MALPEs are "not weighted." That is, the algebraic proportionate errors (ALPEs) underlying each of them are calculated without regard to the proportion of the "actual" numbers they are associated with (the data shown in Table 2). Suppose, however, we do weight them. The two weighted MALPEs are calculated as follows:

WEIGHTED RACE MALPE (WRMALPE) = -0.0358 =

$$[-(.0412)*(996/1369)] + [-(.0230)*(261/1369)] + [-(.0179)*(112/1369)$$

Note that WRMALPE is found by summing the product of the "total" relative errors for each of the three race groups and the corresponding population weight across all four geographic areas. This corresponds to how RMALPE was calculated, given that we are weighting each of the three race group's "total" relative error instead of dividing the sum by three (which is equivalent to giving them all the same "weight.").

WEIGHTED GEOG MALPE (WGMALPE) = -0.0358 =

$$[-(.0668)*(434/1369)] + [-(.0089)*(671/1369)] + [-(.0262)*(97/1369)] + [-(.0719)*(167/1369)]$$

WGMALPE is found by summing the product of the "total" relative errors for each of the four geographic areas and the corresponding population weight across all three race groups.

Finally we have

WRMALPE = WGMALPE =

TOTAL ALPE (TALPE) = 0.0358 = $[\,|(1320-1369)/1369|\,]$

That is, once the two MALPEs have been "weighted," we see they are equal not only to one another, but also to the Total ALPE (Table 7).

Table 7. Summary Algebraic Error by Characteristic

Characteristic	Error
WEIGHTED RACE MALPE	-0.0358
WEIGHTED GEOG MALPE	-0.0358
TOTAL ALGEBRAIC ERROR	-0.0358

These weighted MALPEs provide a key to understanding the relationship between the two unweighted MALPES and the relationship between each of them and the Total ALPE.

Let
F_i = Forecast counts for population in a given characteristic (e.g., race);
CEN_i = Census counts in a the same characteristic (e.g., race); and
k = number of categories in the characteristic (e.g., there are three race groups).
Then, in general terms,

$$MALPE = (1/k)*[\sum_i (F_i - CEN_i)/CEN_i] \quad (9)$$

and

$$\text{Weighted MALPE (WMALPE)} = \sum [(CEN_i/\sum CEN_i)*((F_i - CEN_i)/CEN_i)] \quad (10)$$

We can re-express the preceding definitions of MALPE and WMALPE as follows.

$$MALPE = [\sum(F_i/(k*CEN_i))] - 1 \quad (11)$$

and WALPE as $[(\sum F_i)/(\sum CEN_i)] - 1$ (12)

Note that

$$1 + WMALPE = [(\sum F_i)/(\sum CEN_i)] \quad (13)$$

So, 1+ WMALPE can be viewed as simply the sum of the forecasted population for a given characteristic divided by the total census population.

The re-expression found in (12) represents a short-cut way to calculate WMALPE, which is to subtract 1 from the ratio of the sum of the forecasted population to the total census population.

As an example, we know that 1+ WMALPE = 1+ (-0.0358) = 0.9642. This is equal to the ratio of the sum of the forecasted population (for either race or geography) to the total census population, which in our example is 1329/1369 = 0.9642. Thus, MALPE = -0.0358 = .9642 – 1.

Using the re-expressions in equations (12) and (13), the following relationship can be seen between WMALPE and MALPE:

$$WMALPE = MALPE + [(\sum F_i)/(\sum CEN_i)] - [\sum(F_i/(k*CEN_i))] \quad (14)$$

$$MALPE = WMALPE - [(\sum F_i)/(\sum CEN_i)] + [\sum(F_i/(k*CEN_i))] \quad (15)$$

WMALPE can be viewed as MALPE plus the sum of the forecasted population divided by the total census population from which is subtracted the sum of the (1 = 1 to k) ratios of estimated populations to the census populations with the latter weighted by a constant, k Thus, a WMALPE for a given characteristic (e.g., race) equals the WMALPE for another characteristic (e.g., geography). The WMALPEs for different characteristics also equal their joint MALPE, which simultaneously is equal to the TALPE.

Note that if we have another characteristic (geographic areas), we can modify our notation accordingly

Let F_j = Forecast counts for population for another characteristic (e.g., geography);

CEN_j = Census counts for the same characteristic (e.g., geography); and

L = number of categories in the characteristic (e.g., there are 4 geographic areas).

So that, equations (1) through (6) can be expressed in terms of the “jth” characteristic and its “L” categories.

3.2 The Relationship between RMALPE & GMALPE and their Relationship to the Total Algebraic Percent Error (TALPE)

To be more specific, again using the two example characteristics of race and geography (but noting that this result can be generalized to more than 3 columns and 4 rows, as demonstrated in the appendix), we have the following.

Since

$$RMALPE = WMALPE - (\sum F_i)/(\sum CEN_i) + \sum(F_i/(k * CEN_i)) \qquad (16)$$

Then

$$WMALPE = RMALPE + (\sum F_i)/(\sum CEN_i) - \sum(F_i/(k * CEN_i)) \qquad (17)$$

and since

$$GMALPE = WMALPE - (\sum F_j)/(\sum CEN_j) + \sum(F_j/(L * CEN_j)) \qquad (18)$$

Then

$$WMALPE = GMALPE + (\sum F_j)/(\sum CEN_j) - \sum(F_j/(L * CEN_j)) \qquad (19)$$

It follows that

$$RMALPE = GMALPE + \sum F_j)/(\sum(CEN_j) - \sum(F_j/(L * CEN_j)) - (\sum F_i)/(\sum CEN_i) + \sum(F_i/(k * CEN_i)) \qquad (20)$$

and that

$$RMALPE = GMALPE - \sum(F_j/(L * CEN_{j)}) + \sum(F_i/(k * CEN_i)) \qquad (21a)$$

and

$$GMALPE = RMALPE + \sum(F_j/(L * CEN_j)) - \sum(F_i/(k * CEN_i)) \qquad (21b)$$

We can interpret these relationships as follows. RMALPE is equal to GMALPE - (the sum of the ratios of the forecasts by each geographic area to: the product of the corresponding census numbers for each geographic area and the number (l) of column marginal's (the 4 geographic categories)) + (the sum of the ratios of the forecasts by each racial category to: the product of the corresponding census numbers for each racial category and the number (k) of row marginal's (the 3 racial categories).

Table 8. Empirical Examples of the Relationships Regarding MALPE*

WMALPE=	RMALPE -	$[\sum(F_i/(k*P_i))]$ +	$\{[(\sum F_i)/(\sum P_i)]$
-0.0358 =	-0.0273 -	0.9727 +	0.9642
RMALPE =	WMALPE -	$[\sum(F_i/(k*P_i))]$ +	$[(\sum F_i)/(\sum P_i)]$
-0.0273 =	-0.0358 -	0.9727 +	0.9642
WMALPE=	GMALPE -	$[\sum(F_j/(l*P_j))]$ +	$[(\sum F_j)/(\sum P_j)]$
-0.0358 =	-0.0421 -	0.9579 +	0.9642
GMALPE =	WMALPE +	$[(\sum F_i)/(\sum P_i)]$ -	$[\sum(F_i/(k*P_i))]$
-0.0421 =	-0.0358 +	0.9579 -	0.9642

*Some numbers do not algebraically sum exactly due to rounding.

GMALPE is equal to RMALPE + (the sum of the ratios of the forecasts by each geographic area to: the product of the corresponding census numbers for each geographic area and the number (l) of column marginal's (the 4 geographic categories)) - (the sum of the ratios of the forecasts by each racial category to: the product of the corresponding census numbers for each racial category and the number (k) of row marginal's (the 3 racial categories).These relationships are displayed in Table 8 using the data for our hypothetical forecast.

As can be seen in comparison with the values shown in Table 7, the values in Table 8 for WMALPE derived from RMALPE and GMALPE are the same and from this we can see that TALPE = WMALPE, which means that:

$$\begin{aligned} TALPE &= WMALPE \\ &= RMALPE + [(\sum F_i)/(\sum CEN_i)] - [\sum(F_i/(k_* CEN_i))] \\ &= GMALPE + [(\sum F_j)/(\sum CEN_j)] - [\sum(F_j/(L_* CEN_j))] \\ &= -0.0358 \end{aligned}$$

4. MAPE and Weighted MAPE

As was the case in the preceding discussion of MALPE, note that in the calculations shown for the MAPEs in Table 6 the two MAPEs are "not weighted." That is, the absolute proportionate errors (APEs) underlying each of them are calculated without regard to the proportion of the "actual" numbers they are associated with (the data shown in Table 2). Suppose, however, we do weight them. The two weighted MAPEs are calculated as follows:

WEIGHTED RACE MAPE (WRMAPE) = 0.0358 =

$$[(.0412)_*(996/1369)] + [(.0230)_*(261/1369)] + [(.0179)_*(112/1369)$$

Note that WRMAPE is found by summing the product of the "total" absolute relative errors for each of the three race groups and the corresponding population weight across all four geographic areas. This corresponds to how RMAPE was calculated, given that we are weighting each of the three race group's "total" relative error instead of dividing the sum by three (which is equivalent to giving them all the same "weight.").

WEIGHTED GEOG MAPE (WGMAPE) = 0.0358 =

$$[(.0668)_*(434/1369)] + [(.0089)_*(671/1369)] + [(.0262)_*(97/1369)] + [(.0719)_*(167/1369)]$$

WGMAPE is found by summing the product of the "total" absolute relative errors for each of the four geographic areas and the corresponding population weight across all three race groups. Finally we have

WRMAPE = WGMAPE = TOTAL APE (TAPE) = 0.0358 = $[\,|(1320-1369)/1369|\,]$

That is, once the two MAPEs have been "weighted," we see they are equal not only to one another, but also to the Total APE, as shown in Table 9.

Table 9. Summary Absolute Error by Characteristic

Characteristic	Error
WEIGHTED RACE MAPE	0.0358
WEIGHTED GEOG MAPE	0.0358
TOTAL ABS ERROR	0.0358

As was the case for MALPE, these weighted MAPEs provide a key to understanding the relationship between the two unweighted MAPEs and the relationship between each of them and the Total APE.

Let F_i = Forecast counts for population on a given characteristic (e.g., race);

CEN_i = census counts in a the same characteristics (e.g., race); and

k = number of categories for the characteristic (e.g., there are three race groups).

Then, in general terms,

$$\text{MAPE} = (1/k)*[\sum |((F_i\text{-}CEN_i)/CEN_i)|] \quad (22)$$

and weighted MAPE,

$$\text{WMAPE} = \sum [(CEN_i/\sum CEN_i)*(|(F_i\text{-}CEN_i)/CEN_i)|] \quad (23)$$

We can re-express the preceding definitions of MAPE and WMAPE as follows.

$$\text{MAPE} = |[\sum(F_i/kCEN_i)] - 1| \quad (24)$$

$$\text{and WMAPE} = |[(\sum F_i)/(\sum CEN_i)] -1| \quad (25)$$

Note that

$$1 + \text{WMAPE} = [(\sum F_i)/(\sum CEN_i)]| \quad (26)$$

So, as was the case for WMALPE, 1+ WMAPE can be viewed as simply the sum of the forecasted population for a given characteristic divided by the total census population.

The re-expression found in (13) represents a short-cut way to calculate WMAPE, which is to subtract 1 from the ratio of the sum of the forecasted population to the total census population. As an example, we know that 1+ WMAPE = 1+ (-0.0358) = 0.9642. This is equal to the ratio of the sum of the forecasted population (for either race or geography) to the total census population, which in our example is 1329/1369 = 0.9642. Given this, we know that MAPE = - 0.0358 = .9642 – 1.

Using the re-expressions in (22) and (23) the following relationship is revealed between WMAPE and MAPE.[2]

$$\text{WMAPE} = \text{MAPE} + |(\sum F_i)/(\sum CEN_i) - \sum(F_i/(k*CEN_i))| \quad (27)$$

$$\text{MAPE} = \text{WMAPE} - |(\sum F_i)/(\sum CEN_i) + \sum(F_i/(k*CEN_i))| \quad (28)$$

Thus, WMAPE can be viewed as MAPE + the absolute value of the difference between two terms: (1) the sum of the Forecasted Population counts on a given characteristic (e.g., race with k categories) divided by the sum of Census Population counts on the same characteristic; and (2) the sum of the (k) ratios of Forecasted Populations counts to the Census Population counts, with the latter weighted by the constant, k (the number of categories for the characteristic in question). MAPE can be viewed as the difference between WMAPE, on the one hand, and on the other, the absolute value of the sum of two terms: (1) the sum of the Forecasted Population counts on a given characteristic (e.g., race with k categories) divided by the sum of Census Population counts on the same characteristic; and (2) the sum of the (k) ratios of Forecasted Populations counts to the Census Population counts, with the latter weighted by the constant, k (the number of categories for the characteristic in question).

Note that if we have another characteristic (geographic areas), we can modify our notation accordingly

Let F_j = Forecast counts for population for another characteristic (e.g., geography);

CEN_j = Census counts for the same characteristic (e.g., geography); and

L = number of categories in the characteristic (e.g., there are 4 geographic areas).

So that, equations (1) through (6) can be expressed in terms of the "jth" characteristic and its "L" categories.

As Table 9 suggests, the two WMAPEs taken on two characteristics (e.g., race and geography) are not only equal to one another, but also the Total APE. This should extend to n characteristics, such that all n WMAPEs = one another = the Total APE. We follow this up in the appendix, where we provide a proof that it does generalize to n characteristics. Thus, it appears that we also have general statements for the relationship between a MAPE for a given characteristic and its equivalent WMAPE. With this knowledge in hand, we can now specify the relationships between MAPEs taken on two characteristics, and the Total APE. Moreover, since the two weighted MAPEs (e.g., WRMAPE and WGMAPE) are the same, we can simply refer to "WMAPE" in relation to either dimension.

5. The Relationship between RMAPE & GMAPE and Their Relationship to TAPE

To be more specific, again using the two example characteristics of race (with i=1 to *k* categories) and geography (with j=1 to *l* categories), we have the following relationships for the MAPEs on each of the two characteristics, (RMAPE = RACE MAPE and GMAPE = GEOG MAPE).

Since

$$RMAPE = WMAPE - (\sum Fi)/(\sum CENi) + \sum(Fi/kCENi) \quad (29)$$

Then

$$WMAPE = RMAPE + (\sum Fi)/(\sum CENi) - \sum(Fi/kCENi) \quad (30)$$

and since

$$GMAPE = WMAPE - (\sum Fj)/(\sum CENj) + \sum(Fj/lCENj) \quad (31)$$

Then

$$WMAPE = GMAPE + (\sum Fj)/(\sum CENj) - \sum(Fj/lCENj) \quad (32)$$

It follows that

$$RMAPE = GMAPE + \sum Fj)/(\sum CENj) - \sum(Fj/lCENj) - (\sum Fi)/(\sum CENi) + \sum(Fi/kCENi) \quad (33)$$

$$RMAPE = GMAPE - \sum(Fj/lCENj) + \sum(Fi/kCENi) \quad (34a)$$

and

$$GMAPE = RMAPE + \sum(Fj/lCENj) - \sum(Fi/kCENi) \quad (34b)$$

We can interpret these relationships as follows. RMAPE is equal to GMAPE - (the sum of the ratios of the forecasts by each geographic area to: the product of the corresponding census population numbers for each geographic area and the number (l) of column marginal's (the 4 geographic categories)) + (the sum of the ratios of the forecasts by each racial category to: the product of the corresponding census population numbers for each racial category and the number (k) of row marginal's (the 3 racial categories). GMAPE is equal to RMAPE + (the sum of the ratios of the forecasts by each geographic area to: the product of the corresponding census population numbers for each geographic area and the number (l) of column marginal's (the 4 geographic categories)) - (the sum of the ratios of the forecasts by each racial category to: the product of the corresponding census population numbers for each racial category and the number (k) of row marginal's (the 3 racial categories).These relationships are displayed in Table 10 using the data for our hypothetical forecast.

Table 10. Empirical Examples of the Relationships Regarding MAPE*

WMAPE=	RMAPE +	[∑(Fi/(k*Pi))] -	{[(∑Fi)/(∑Pi)]
0.0358 =	0.0273 +	0.9727 -	0.9642
RMAPE =	WMAPE -	[∑(Fi/(k*Pi))] +	[(∑Fi)/(∑Pi)]
0.0273 =	0.0358 -	0.9727 +	0.9642
WMAPE=	GMAPE +	[∑(Fj/(l*Pj))] -	[(∑Fj)/(∑Pj)]
0.0358 =	0.0421 +	0.9579 -	0.9642
GMAPE =	WMAPE -	[(∑Fi)/(∑Pi)] +	[∑(Fi/(k*Pi))]
0.0421 =	0.0358 -	0.9579 +	0.9642

*Some numbers do not algebraically sum exactly due to rounding.

As can be seen in comparison with the values shown in Table 10, the values in Table101 for WMAPE derived from RMAPE and GMAPE are the same and from this we can see that TAPE = WMAPE, which means that:

$$\begin{aligned} TAPE &= WMAPE \\ &= RMAPE + [(\sum Fi)/(\sum CENi)] - [\sum (Fi/(k_* CENi))] \\ &= GMAPE + [(\sum Fj)/(\sum CENj)] - [\sum (Fj/(l_* CENj))] \\ &= 0.0358 \end{aligned}$$

6. Discussion

As noted earlier in the discussion of the hypothetical data set and its two characteristics, race and geography, RMAPE is equal to GMAPE - (the sum of the ratios of the forecasts by each geographic area to: the product of the corresponding census population numbers for each geographic area and the number (l) of column marginal's (the 4 geographic categories)) + (the sum of the ratios of the forecasts by each racial category to: the product of the corresponding census population numbers for each racial category and the number (k) of row marginal's (the 3 racial categories). GMAPE is equal to RMAPE + (the sum of the ratios of the forecasts by each geographic area to: the product of the corresponding census population numbers for each geographic area and the number (l) of column marginal's (the 4 geographic categories)) - (the sum of the ratios of the forecasts by each racial category to: the product of the corresponding census population numbers for each racial category and the number (k) of row marginal's (the 3 racial categories).These relationships are displayed in Table 10 using the data for our hypothetical forecast. Either of the preceding two "interpretation" statements concerning the relationship between RMAPE and GMAPE is likely to lead to glazed-over eyes on somebody who asked why the MAPEs in Table 6 are different. The same would apply to the two relationships for each of the other summary measure, MALPE.

Here is an alternative way to answer someone who asks why they are different and, given these differences, how they are related to one another and to the Total error. We express this in terms of MAPE, with the understanding that the same statement applies to MALPE, given appropriate substitutions:

> "We can answer this question by looking first at the Total APE, which is .0358. One might think at first glance that it is an average of RMAPE (.0273) and GMAPE (.0421), but it is not. The relationship between, RMAPE and GMAPE, and in turn, the relationship of both of them to Total APE turns out to be a bit more complicated in that they involve the implicit weights that are used when one takes a simple arithmetic average, as is the case with MAPE. The implicit weight for each of the three race categories used in calculating RMAPE is 1/3 and the implicit weight for each of the four geographic areas used in calculating GMAPE is 1/4. That is, the forecast errors found for the categories within a given dimension are given equal weights. These implicit weights do not correspond to the actual weights, which for the former are the proportions of the population in each of the three race categories and for the latter are the proportions of the population in each of the four geographic areas. When these "actual weights" are used then RMAPE is equal to GMAPE, and in turn both are both equal to the Total APE: .0358. There is nothing wrong about using the implicit weights in that the resulting MAPEs provide useful information about the average error for each dimension. However, it also is useful to know that when actual weights are used, all of the dimensional MAPEs are not only the same, but also equal to the Total Absolute Percent Error."

7. Summary

This paper has examined an issue that appears to be hitherto unexplored by demographers and others who routinely evaluate the accuracy of population forecasts and estimates, namely the lack of an apparent relationship between the same summary measures of error when taken across multiple characteristics. In a non-rigorous manner it has demonstrated that there is a relationship among each of two important summary measures of error, MALPE and MAPE, respectively, when used to evaluate forecasts or estimates done across multiple characteristics in terms of bias and precision. That is if MALPE is taken on one dimension (e.g., age), it has a relationship to a MALPE taken across another dimension (e.g., race). The same finding holds for MAPE. This means in evaluating a forecast that includes more than one characteristic, one can express MALPE (or MAPE) as found for one characteristic (e.g., race) in terms of each of the other characteristics (e.g., age and geography). This finding allows one to explain the reason for differences in a summary error measure used to evaluate a forecast done for more than one characteristic. The findings are informative and suggest that the results are generalizable. Following this suggestion, a general formal general proof is available from the author, both for MALPE and for MAPE. This proof supports the non-rigorous treatment found in the main body of the text.

References

[1] Chi, G., and P. Voss (2011). "Small area population forecasting: Borrowing strength across space and time." *Population, Space, and Place*, 17(5): 505-520.

[2] Chu, S. (1974). "On the use of regression methods in estimating regional population." *International Statistical Review*, 42 (1): 17-28.

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

[4] Long, J. (1995). "Complexity, accuracy, and utility of official population projections. " *Mathematical Population Studies*, 5(3): 203-216.

[5] Rayer, S. (2008). "Population forecast errors: A primer for planners." *Journal of Planning Education and Research*, 27(4): 417-430.

[6] Shahidullah, M., and M. Flotow (2005). "Criteria for selecting a suitable method for producing post-2000 county population estimates: A case study of population estimates in Illinois." *Population Research and Policy Review*, 24 (3): 215-229.

[7] Smith, S. and T. Sincich (1992). "Evaluating the forecast accuracy and bias of alternative population projections for states." *International Journal of Forecasting*, 8(3): 495-508.

[8] Smith, S., J. Tayman, and D. A. Swanson (2013). *A Practitioners Guide to State and Local Population Projections*. Springer B.V. Press. Dordrecht, Heidelberg, London, and New York.

[9] Swanson, D. A. and J. Tayman (2012). *Subnational Population Estimates*. Springer B.V. Press. Dordrecht, Heidelberg, London, and New York.

[10]Swanson, D., J. Tayman, and C. Barr (2000). "A note on the measurement of accuracy for sub-national demographic estimates." *Demography*, 37(2): 193-201.

[11] Swanson, D. A., J. Tayman, and T. Bryan (2011). "MAPE-R: A Rescaled Measure of Accuracy for Cross-Sectional, Sub-national Forecasts." *Journal of Population Research*, 28(2-3): 225-243.

[12] Tayman, J. (1996). "The accuracy of small area population forecasts based on a spatial interaction modeling system." *Journal of the American Planning Association*, 62(1): 85-98.

[13] Tayman, J., and D. Swanson (1999). "On the validity of MAPE as a measure of population forecast accuracy." *Population Research and Policy Review*, 18(4): 299 -232.

[14] Tayman, J., D. Swanson, and C. Barr (1999). "In search of the ideal measure of accuracy for sub-national demographic forecasts." *Population Research and Policy Review*, 18(5): 387-409.