Title

New insights on the impact of coefficient instability on ratio-correlation population estimates

Authors
Tayman, Jeff
Swanson, David A

Publication Date
2016

Peer reviewed

NEW INSIGHTS ON THE IMPACT OF COEFFICIENT INSTABILITY ON RATIO-CORRELATION POPULATION ESTIMATES

Jeff Tayman (Correspondence author)
Department of Economics
University of California San Diego
9500 Gilman Drive
La Jolla, California 92093-0508 USA
Tel: 1-858-581-6710
Email: jtayman@ucsd.edu

David A. Swanson
Department of Sociology
University of California Riverside
Riverside, California 92521 USA
Tel: 1-951-827-4373
E-mail: **David.swanson@ucr.edu**

Abstract

In this study we examine the regression-based ratio-correlation method and suggest some new tools for assessing the magnitude and impact of coefficient instability on population estimation errors. We use a robust sample of 904 counties from 11 states and find that: (1) coefficient instability is not a universal source of error in regression models for population estimation and its impact is less than commonly assumed; (2) coefficient instability is not related to bias, but it does decrease precision and increase the allocation error of population estimates; and (3) unstable coefficients have the greatest impact on counties under 20,000 in population size. Our findings suggest that information about the conditions that affect coefficient instability and its impact on estimation error might lead to more targeted and efficient approaches for improving population estimates developed from regression models.

Key Words: population estimation; ratio-correlation method; coefficient instability

Introduction

The regression approach for estimating population has a long history beginning with Snow's [34] seminal paper. This method, most often used for county population estimates, involves relating changes in population to changes in one or more symptomatic indicators [1]. Symptomatic indicators relate to changes in population such as vital events, employment, school enrollment, voter registration, and tax returns. While variations have been developed, the most common regression-based approach for estimating populations is the "ratio-correlation" method introduced and tested by Schmitt and Crosetti [28] and Crosetti and Schmitt [3]. Comparative analysis has shown the ratio correlation method is one of the most accurate approaches for estimating population [1, 9, 23, 30, 15]. Swanson [39] has also observed that the ratios of change used in the ratio-correlation model provide some of the benefits associated with "stationarity," an important characteristic associated with a good time series model. Typically, regression-based equations for population estimation are cross-sectional, use 30 to 250 observations, and contain two to four symptomatic indicators.

Given good quality input data, the accuracy of the ratio-correlation and other regression-based methods largely depends on the validity of the underlying assumption that the observed relationships between the symptomatic indicators and population in the past intercensal period (e.g. 2000 to 2010) will be the same in the postcensal period (e.g., 2010 to 2020) [1, 6, 21, 4: 173]. That changes in the coefficients relating the symptomatic indicators and population between the estimation and postcensal periods, or coefficient instability, will transmit error into postcensal estimates is not debatable. However, research into the magnitude of coefficient instability and its effect on population estimate error is far from conclusive. Some studies have found that coefficient instability is a significant issue, while others have found this not to be the case. Moreover, attempts to alter methods and procedures for dealing with coefficient instability have generally led to marginal improvements to population estimates.

In this study, we provide expanded and updated analyses and suggest some new tools for assessing the magnitude and impact of coefficient instability on population estimation errors. We use the ratio-correlation technique, a robust sample of counties, and a variety of measures and analytical techniques to address the following three questions:

(1) What is the extent and magnitude of coefficient instability?

(2) What is the impact of coefficient instability on the bias, precision, and allocation error of population estimates?

(3) What is the impact of coefficient instability on estimate error relative to the size and growth rate of counties?

This study differs in several ways from previous research. First, we analyze a sample of 904 counties from 11 states. To our knowledge, this is the largest and most diverse sample ever used to study regression models for population estimation, which have focused on case studies of counties from a single state. While valuable, these case studies are limited in their generalizability and unique conditions within a state can have a substantial impact on the behavior of ratio-correlation models through time [20]. Second, we offer new measures of the magnitude of coefficient instability derived directly from changes in the regression coefficients between the estimation and postcensal periods. Third, we examine the three main dimensions of estimate error (precision, bias, and allocation error). Studies of coefficient instability have largely focused on the accuracy of estimates, as measured by the Mean Absolute Percent Error (MAPE), and have not addressed estimate bias. Tayman and Schafer [43] is the only study that we are aware of that has examined allocation error in the context of coefficient instability. Finally, we go beyond the usual approach of using aggregate data to compare typical errors for areas with different groupings of a characteristic (e.g., size and growth rate) by constructing statistical models based on data from individual counties [44]. Disaggregated statistical models can help identify patterns that cannot be observed through aggregated data analysis.

Literature review

It has been argued that further improvements in accuracy in regression models of population estimation are not likely until ways for adjusting or mitigating the impact of coefficient instability are developed [24], and many refinements have been suggested to deal with this problem. Ericksen [6, 7] proposed a method of postcensal estimation which combined the symptomatic indicators with sample survey data from the Current Population Survey. Swanson [37] presented a mildly restrictive approach using a theoretical causal ordering and principles from path analysis to modify the regression-coefficients in the postcensal period. Other methods for dealing with coefficient instability have included the use of dummy variables in the regression model [25], estimating separate models for different geographic stratifications [27], using differences in ratios as opposed to ratio of ratios over the estimating period [26, 29, 36], and using logarithmic transformations of the variables [42].

Multicollinearity can also affect coefficient instability in that estimated coefficients may change radically in response to small changes in the model or data and can also create difficulties in assessing the statistical significance of coefficients [4: 165-175]. Attempts to mitigate the impact of muticollinearity on population estimates include the averaging of estimates of univariate regression models [22] and ridge regression [35].

Improvements in estimate accuracy by these refinements have not been uniformly significant and the basic form of the ratio-correlation model has proven robust over a wide range of conditions [42, 43]. However, we are aware of only three studies that have either measured the extent of coefficient instability in the ratio-correlation model and its variants or ascertained its impact on the resultant population estimates.

In the first of these three studies, O'Hare [23] analyzed 1970 estimates of total population from ratio- and difference-correlation models for counties in Michigan. These models were constructed over the 1950 and 1960 time period using school enrollment (grades 1-8), auto registration, sales tax revenue, and vital events as symptomatic indicators. Using correlation matrices to measure coefficient instability, he found that the difference correlation model was more stable over time and yielded slightly more accurate estimates than the ratio-correlation model with MAPEs of 4.5 and 4.7, respectively.

In the second study, Mandell and Tayman [18] analyzed 1970 estimates of total population from eight ratio-correlation and eight difference-correlation models for counties in Florida. These alternate models constructed over the 1950-1960 and 1960-1970 time periods were based on different combination of symptomatic indicators that included sales tax revenues, school enrollment (grades 1-8), vital events, labor force population, occupied housing units, and families receiving aid for dependent children. Mandell and Tayman criticized the use of the correlation coefficient to measure coefficient instability because it consists of the unstandardized regression slope and the standard deviations of the independent and dependent variables. Therefore, one cannot know whether the change or lack of change in a correlation coefficient is due to coefficient instability, differences in the variability of model variables, or both. Given the drawback of the correlation coefficient to measure coefficient instability, they proposed the use of the F-statistic based on the Chow [2] test as a direct measure to quantify the change in a set of regression coefficients from one time period to another.

Mandell and Tayman [18] found that the ratio-correlation models had more stable coefficients than the difference correlation models and 7 of the 8 ratio-correlation models had lower MAPEs. This was in sharp contrast to O'Hare's [23] findings. They also found a strong positive relationship (Spearman's *Rho* = 0.72) between the F-statistic and MAPE for the ratio-correlation models, but a negligible relationship in the difference-correlation models (Spearman's *Rho* = -0.07). The findings for the ratio-model lent considerable support to the assertion that coefficient instability has a great impact on the accuracy of ratio-correlation models for population estimation. However, the lack of association found in the difference models cast some doubt as to the pervasiveness of the effect of coefficient instability on estimate accuracy and that other factors should be considered as well.

In the third study, Tayman and Schafer [43] conducted a series of experiments that analyzed 1980 estimates of total population from ratio-correlation models for counties in Washington State. These models were constructed over the 1960 and 1970 time period using school enrollment (grades 1-8), voter registration, and employment as symptomatic indicators. They created six different tests by varying the decade of estimation (1960-1970 and 1970-1980) and by using estimated or actual values for the symptomatic indicators. Measurement error was introduced into the symptomatic indicators using predicted values derived from regression models for each symptomatic indicator. These six tests allowed the examination of the relative impacts of coefficient instability, symptomatic indicator measurement error in the estimating equation, and postcensal symptomatic indicator measurement error.

Tayman and Schafer [43] found that coefficient instability and measurement error in the estimating equation contributed little to estimate error, while poorly measured postcensal symptomatic indicators had, by far, the great impact of estimate error. The MAPE for the standard model (no coefficient drift or measurement error) was 2.8 compared to a MAPE 3.0 for the model containing just coefficient drift, to a MAPE of 2.9 for the model just containing measurement error in the estimating equation, and to a MAPE of 3.8 for the model containing measurement error in the postcensal symptomatic indicators. These findings called into question the prevailing thought that reducing coefficient instability would be the principal mechanism for achieving greater accuracy in ratio-correlation models.

Sample

Most of our 904-county sample comes from states that use the ratio-correlation model in their official population estimates. As such, we evaluate models and variables currently used in practice. We began with a list of contacts for the 50 states from the U.S. Census Bureau's Federal State Cooperative Program for Population Estimates. Based on a review of agency web sites and telephone conversations, we identified 9 states that fit our criteria (California, Colorado, Illinois, North Carolina, Oregon, Texas, Virginia, Washington, and Wisconsin). We augmented this sample by adding Alaska and Nevada in part because of data availability, but also because these states have several unique characteristics including a relatively small number of counties, distinct settlement patterns (heavily rural with a few large urban centers), and economies largely dependent on a single industry (natural resources and gaming).

Characteristics of the sample are shown in Table 1. The 904 counties comprise 29% of all counties and 37% of the 2000 population in the United States. The overall sample has larger counties and faster percentage change on average than all counties, with average 2000 population sizes of 114,888 and 89,491 and average percent population change between 2000 and 2010 of 7.7% and 5.2%, respectively. Median population sizes are closer, with 25,249 for the sample compared to 24,556 for all counties (data not shown). Median percent changes still show the sample counties have faster growth between 2000 and 2010 (5.7% versus 3.2%). The faster growth rate for the overall sample is due in large part to a smaller percent of counties with declines (27%) compared to 36% for all counties.

Table 1. Population Size and Growth Rate Characteristics of the Sample

				Percent Change Distribution							
	No. of Counties	2000 Pop.[a]	2000-10 Percent Change[a]	< -5.0%	-4.9 - -0.1%	0.0 - 2.9%	3.0 - 4.9%	5.0 - 9.9%	10.0 - 14.9%	15.0 - 24.9%	25.0+%
All Counties	3,141	89,491	5.2%	17%	19%	14%	9%	16%	10%	9%	7%
Sample Counties	904	114,888	7.2%	12%	15%	11%	10%	19%	12%	12%	9%
Alaska	29	21,618	2.4%	28%	17%	10%	10%	7%	14%	7%	7%
California	58	583,994	10.3%	2%	3%	12%	16%	26%	12%	24%	5%
Colorado	63	68,033	8.4%	17%	10%	6%	8%	13%	14%	21%	11%
Illinois	102	121,760	1.8%	19%	41%	16%	5%	10%	2%	3%	5%
Nevada	17	117,544	13.9%	12%	6%	12%	6%	18%	6%	18%	24%
North Carolina	100	80,465	13.1%	0%	7%	10%	6%	27%	15%	19%	16%
Oregon	36	95,040	6.9%	8%	14%	6%	14%	28%	19%	8%	3%
Texas	254	82,094	7.0%	17%	14%	11%	11%	16%	10%	10%	10%
Virginia	134	52,825	9.3%	10%	12%	14%	10%	16%	10%	13%	13%
Washington	39	151,132	12.4%	3%	3%	3%	10%	28%	28%	23%	3%
Wisconsin	72	74,496	3.9%	7%	19%	14%	17%	28%	13%	1%	1%

[a] Average across counties.

There is substantial variability in size and growth rate between the 11 states (see Table 1). Average 2000 population sizes range from 21, 618 in Alaska to 583,994 in California. Sixty-nine percent of the counties in Alaska have less than 10,000 persons compared to 5% in North Carolina (data on the population size distribution is not shown). California, North Carolina, and Wisconsin are the other states where fewer than 10% of their counties contain less than 10,000 people, while almost one-half of the counties in Nevada fall into this category. Only 3% of the counties in Alaska have a population of 100,000 or more. Virginia has the second lowest percent of counties with 100,000 or more (10%) while fewer than 15% of the counties in Nevada and Texas have 2000 populations this large. California has the largest share of counties with 100,000 or more persons (60%). Washington has the second highest share of counties of this size (28%), followed by Oregon (25%) and North Carolina (23%).

The average percent change between 2000 and 2010 ranges from 1.8% for counties in Illinois to 13.9% for counties in Nevada. Alaska and Wisconsin are the other states where county growth rates average less than 5%. On average, counties in Nevada and North Carolina have the largest percent increases with 13.9% and 13.1%, respectively. California and Washington are the other states where county growth rates average more than 10%. Sixty percent of the counties in Illinois decline between 2000 and 2010. Alaska has the second largest percent of declining counties (45%), followed by Texas (31%). Less than 10% of the counties experience population decline in California, North Carolina, and Washington. Forty-two percent of the counties in Nevada experience population change of 15% or more. In North Carolina and Colorado more than 30% of their counties show changes at this level. Wisconsin and Illinois have the smallest share of counties with large percent changes at 2% and 8%, respectively, followed by Oregon (11%) and Alaska (14%).

Methods [1]

Ratio-Correlation Model

In the ratio-correlation method, the population and symptomatic indicators are measured as ratios that represent shares or proportions (e.g., county population / state population and county employment / state employment) at each census point. The change in these censal ratios is measured by dividing the censal ratio from the latest census by the corresponding censal ratio from the earlier census. Ratio-correlation estimates are based on a regression equation estimated using variables from the last two decennial censuses points that precede the date of the postcensal estimate. An estimate is then derived by solving the equation using values of the symptomatic indicators in the postcensal estimation year. When estimating counties, the ratio-correlation model requires an independent estimate of the state population. This study uses the "official" state population estimates for 2010 produced by each state agency in the sample. A detailed description of the ratio correlation model can be found in [41: 167-168].[2]

A variety of symptomatic indicators were obtained for the counties in each state (see Table 2). Separate ratio-correlation models are estimated for the counties in each state using 1990 and 2000 as the estimation period. These models are used to estimate the 2010 total population for each county, which is evaluated against the 2010 census population. Our objective is to estimate the "best" model for each state using standard criteria for selecting appropriate variables (e.g., significance tests using $\alpha = 0.05$; strength of relationship; examination of residual statistics and plots, and tests for multicollinearity). The variables in the final models are shown in bold in Table 2.

Following Tayman and Schafer [43], we evaluate the magnitude and impact of temporal instability by producing two sets of 2010 county estimates. The first set is based on ratio-correlation models estimated using the 1990 and 2000 decade (Model90-00). For the second set, we re-estimate the models using 2000 and 2010 as the estimation period to derive coefficients that reflect the simulated postcensal period (Model00-10). The "best" models turn out to include

Table 2. Regression Model Independent Variables

	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6
Alaska	**Permanent Fund Residents**					
California	**License Drivers**	**Sch. Enrollment**	Housing Units	Births	Deaths	
Colorado	**Registered Voters**	**Sch. Enrollment**	**QCEW Employment**	Registered Vehicles	Births	Deaths
Illinois	**Registered Vehicles**	**Sch. Enrollment**	**Fed Tax Exemptions**	**Births**		
Nevada	**Sch. Enrollment**	QCEW Employment	Labor Force	Births	Deaths	
North Carolina	**Registered Vehicles**	**Sch. Enrollment**	**Births**			
Oregon	**Medicare Enrollment**	**Sch. Enrollment**	**State Tax Exemptions**	Registered Voters	Births	
Texas	**Registered Vehicles**	**Sch. Enrollment**	**Registered Voters**	**Births**	Deaths	
Virginia	**License Drivers**	**Sch. Enrollment**	**Housing Units**	**Births**	Deaths	
Washington	**Registered Voters**	**Sch. Enrollment**	**Registered Vehicles**	Births	Deaths	
Wisconsin	**License Drivers**	**Sch. Enrollment**	**State Tax Exemptions**			

Notes: 1. **Bold** = included in the final model.
2. School enrollment grades 1 through 8.

the same variables found in the 1990 and 2000 estimation, although the relative ranking of their slopes may change. For example, in Colorado the unstandardized slope for school enrollment (0.502) in Model90-00 is the largest of the three variables, but its value declines to 0.278 in Model00-10 and is the smallest (data not shown). The second set of 2010 population estimates is developed using the estimated coefficients from the Model00-10 while keeping all other information the same as in the first set of estimates. Any differences in errors between the two sets of estimates is due solely to coefficient changes between the estimation and simulated postcensal periods, which allow a direct examination of the magnitude and impact of coefficient instability on population estimates not confounded by other factors. Details of the regression results for Model90-00 and Model00-10 for each state are available from the authors.

Bryan [1] notes that the use of multiple and differing variables can compromise the comparison of ratio-correlation estimates between different subnational areas. There may be advantages in using the same variables for the models in each state (e.g., comparing the effect and importance of public school enrollment on population changes across different states). Enforcing homogeneous model specifications for each state does not seem warranted in this context. Our aim is to develop ratio-correlation models with sound and defensible statistical properties that exploit the maximum amount of useful information to optimize the accuracy in the simulated postcensal estimates. We believe using homogeneous model specifications may likely yield suboptimal estimates for some, if not all, states and add an extraneous source of error into the analysis.

Measures of Coefficient Instability

As noted earlier, Mandell and Tayman [18] proposed an alternative to the correlation coefficient for measuring coefficient instability in regression models for population estimation; namely, a method within covariance analysis that tests the temporal stability of a set of regression coefficients. The Chow test [2] uses the residual sum of squares (RSS) from two time periods (e.g., 1990-2000 and 2000-2010) and from a pooled regression (e.g., 1990-2010) to calculate an F-statistic (F_{Chow}). While the usual application of the Chow test is to statistically test the Ho: No structural change in the regression coefficients between the two time periods, Mandell and Tayman [18] also used F_{Chow} independently as an empirical index of coefficient instability with a smaller F_{Chow} indicating less coefficient instability.

While useful, F_{Chow} has important drawbacks as an empirical measure of coefficient instability. First, it lacks an intuitive interpretation and is unfamiliar to many users and producers of population estimates. Second, F_{Chow} is a global test of coefficient instability and does not provide information for individual coefficients in models with more than one symptomatic indicator.

We propose alternatives to the F_{Chow} based on a direct comparison between the regression slopes for the independent variables (symptomatic indicators) that address its shortcomings. The absolute value of the percentage change in the unstandardized regression slopes (*apchb*) between the first and second time periods (i.e., 1990-2000 and 2000-2010) measures the instability of individual coefficients:

$$apchb_v = \left| (b_{v,t=2} - b_{v,t=1}) / b_{v,t=1} \times 100 \right|$$

where, *v* is the independent variable; *b* is the unstandardized regression coefficient; and *t* is the time period.

We take the absolute value because it is the magnitude of the slope change for each variable that is important in capturing coefficient instability and not the direction of that change.

We also conduct a statistical hypothesis test of the difference between the individual regression coefficients using a pooled-regression with a dummy variable and interaction terms. The dummy variable is 0 for the 1990-2000 period and 1 for the 2000-2010 period. It represents the difference in the intercepts between Model90-00 and Model00-10. The interaction term(s) are the product of the dummy variable and each independent variable in the model and represent difference in the regression coefficients between the two models. The *p-values* for the interaction term(s) are used for evaluate the Ho: No difference in the individual regression coefficients.

We also evaluate two global measures of coefficient instability. The first measure, called the instability index (*isi*), is the arithmetic average over the *apchb*:

$isi = \sum apchb_v$,

where, $\sum$ is the sum over *v*.

The second is a weighted instability index (*wtisi*) that uses the slope from the first time period ($b_{v,t=1}$) as the weight:

$wtisi = \sum(b_{v,t=1} \times apchb_v) / \sum(b_{v,t=1})$.

The idea behind *wtisi* is a large percentage change can result from a relatively small numeric change where the initial value is low. Basing the weight on the size of the slope will lessen the influence of small numeric and large percentage changes on the average.

Measures of Estimate Error

We analyze several commonly used measures that capture three dimensions of estimate error—accuracy or precision, bias, and allocation error [41: 268-273]. Error is defined as the difference between the simulated 2010 population estimate and the 2010 census count for each county. The mean absolute percent error (MAPE) measures estimate accuracy in which positive and negative errors do not offset each other. It shows the average percentage difference between the estimated an actual population whether individual estimates were too high or too low. The mean algebraic percent error (MALPE) is a measure of bias in which positive and negative values offset each other. A positive MALPE reflects the tendency for the estimates to be too high on average and a negative MALPE reflect the tendency for the estimates to be too low on average.[3]

The measures described above are based on the error for a particular geographic area. Another perspective views the misallocation of the estimates across geographic space, in our study counties. Allocation error is pertinent for estimation procedures like the ratio-correlation model that nest population from a larger geographic area into smaller areas and use the postcensal population estimate from the larger area as a control. A number of measures can be used to measure allocation error [5, 19]. For this study we use the Index of Dissimilarity (IOD), a popular measure for evaluating postcensal demographic and economic estimate allocation error [8, 31: 425-427, 38]. The IOD compares the percentage distribution of the estimated county population with the percentage distribution in the census and calculates the percentage that the estimated distribution would have to change to match the census distribution. The IOD ranges from 0 to 100, with 0 indicating indentical percentage distributions and 100 indicating complete disparity between the estimated and census distributions.

Other Methods

We employ several other methods to analyze the impact of coefficient instability on estimation error for the total population. In addition to a descriptive comparison of the accuracy, bias, and allocation error for Model90-00 and Model00-10 for counties in each state, we conducted a one-tailed paired observation t-test of Ho: $\mu_{1990\text{-}2000} > \mu_{2000\text{-}2010}$ for the MAPEs and MALPEs [18]. A paired observation t-test is appropriate because we assume the errors for each county are related over time and cross-sectionally between counties. To measure the strength of the association between coefficient instability and estimate error, we use Spearman's Rank-Order correlation coefficient (*Rho*).[4] The *Rho's* relate F_{Chow}, *isi*, *wtisi* to the MAPE, MALPE, and IOD for the 11 sample states.

To examine the impact of coefficient instability on errors by population size and growth rate, we construct statistical models based on the estimates for individual counties [44].[5] The dependent variable is a binary indicator assigned to each county based on a comparison of the absolute percent errors between Model00-10 and Model90-00, where 0 indicates a lower error in Model00-10 and 1 indicates a larger error in Model00-10. We first examine a univariate crosstabulation of the binary indicator against population size and growth rate categories and then use a binary logistic model to examine the combined effects of size and growth rate. [16]. Size is measured as the natural logarithm of the 2000 population and growth rate is measured as the 1990-2000 percent change in population.

Analysis

Magnitude of Coefficient Instability

We begin the analysis examining the magnitude of coefficient instability for each state. Table 3 shows the F_{Chow} and the *p-value* of the test hypothesis of no change in the set of regression coefficients along with the instability index (*isi*) and weighted instability index (*wtisi*). The measure of individual coefficient instability (*apchb*) and their *p-values* for the state-specific ratio-correlation models is shown in the Appendix.

The measures based on the Chow test indicate that 5 of 11 states show little coefficient instability with F_{Chow} values less than 2.0 and *p-values* indicating acceptance of the null hypothesis. The greatest coefficient instability occurs in Texas and Illinois that have by far the largest F_{Chow} values of 18.6 and 21.9, respectively. The F_{Chow} values show considerable variability, ranging from 0.7 in Alaska to 21.9 in Texas, much larger than the range found by Mandell and Tayman [18] in eight ratio-correlation models for Florida (3.3 to 6.4). This comparison suggests that the context of the ratio-correlation model (i.e., the state where the model is estimated) has a greater impact on the magnitude of coefficient instability compared to the effect of different sets of variables within the same context.

Table 3. Regression Coefficient Instability Measures[a]

	Change in Coefficients[b]		Chow Test	
Counties in	Average	Weighted Average	F-Statistic	P-value[c]
Alaska	14.2%	14.2%	0.724	0.49
California	7.0%	6.3%	0.973	0.41
Colorado	49.0%	39.3%	1.508	0.20
Illinois	71.8%	60.0%	18.568	<0.01
Nevada	30.5%	30.5%	4.474	0.02
North Carolina	14.6%	14.4%	0.245	0.91
Oregon	39.6%	35.6%	2.699	0.04
Texas	93.1%	87.4%	21.919	<0.01
Virginia	46.5%	44.9%	1.701	0.14
Washington	80.0%	44.4%	3.228	0.02
Wisconsin	21.9%	14.2%	2.372	0.06

[a] Based on ratio-correlation models for 1990-2000 and 2000-2010.
[b] Absolute % change in the unstandardized regression slopes.
[c] Ho: No change in the set of regression coefficients

There is a close relationship between the global stability measures based on the Chow test and the change measured directly from the regression slopes. The parametric correlation coefficients (*r*) relating F_{Chow}, *isi*, and *wtisi* range from 0.746 to 0.940 (data not shown). California, Alaska, and North Carolina have the lowest F_{Chow}, *isi*, and *wtisi* values and the highest *p-values.* The changes in all of their individual coefficients are modest and have large *p-values* (see Appendix). Illinois and Texas have the largest F_{Chow} and *isi* values and the largest and third largest *wtisi* values, respectively. In Texas all coefficients show substantial percent changes and have *p-values* < 0.05. In Illinois, the large and significant global instability is due to the school enrollment and tax exemptions variables. Births and vehicle registrations have relative low *apchb* and non-significant *p-values* (see Appendix).

There are a few inconsistencies between the measures of coefficient instability. Virginia and Colorado have relatively large *isi* and *wtisi* values, but their F_{Chow} statistics are relatively small with *p-values* that fail to reject the null hypothesis. In Virginia only one variable (driver licenses) has a large *apchb* (87.3%) and a significant *p-value* (0.031), while the changes in the other variables are not large enough to cause a statistically significant global change based on the Chow test (see Appendix). The situation is somewhat similar in Colorado, where only school enrollment shows a significant change and voters show a small, insignificant change. Interestingly, the employment variable has the largest *apchb* of any variable, but it is not significant according to its *p-value* (see Appendix). Wisconsin has relatively small *isi* and *wtisi* values, but its F_{Chow} statistic is relatively large with a *p-value* = 0.06 that rejects the null hypothesis of stability at $\alpha = 0.10$. Two of the three independent variables in the Wisconsin model (driver licenses and tax returns) have stable coefficients, but the school variable coefficient has a large percent change (55.8%) with a *p-value* = 0.056 that causes the Chow test to reject the hypothesis of stability (see Appendix).

For most states, there is not a great deal of difference between the *isi* and *wtisi, which indicates* consistency in the size of the unstandardized slopes in the 1990-2000 equation. The greatest differences are seen in Wisconsin and Washington, where the *wtisi* are smaller than the *isi* by 35% and 45%, respectively. In Wisconsin the *wtisi* reduces the influence of the 55.8% change in the school enrollment coefficient because its slope 0.163 is considerably smaller than the slopes of driver licenses (0.492 and 7.3% change) and tax returns (0.268 and 2.4% change) variables (see Appendix). The effect is even more dramatic in Washington, where the slope on voter registration variables changes by 181.7%. However, its slope of 0.093 is much smaller than those of registered vehicles (0.336 and 55.1% change) and school enrollment (0.398 and 3.3% change).

These findings underscore the value of having alternative measures of global coefficient instability and instability measures for individual coefficients to help identify the specific variable or variables most affecting the instability of the ratio-correlation model. They also show the value of analyzing both description changes and performing statistical tests on these changes.

Impact of Coefficient Instability on Estimate Error

Table 4 compares the errors from a model using coefficients estimated from the 1990-2000 period (Model90-00) to errors from a model using coefficients estimated from the 2000-2010 period (Model00-10) for the counties in each state. We expect that errors from Model00-10 to be smaller because the coefficients are estimated over the simulated postcensal period effectively eliminating the effect of coefficient instability on estimate error. Along with the measures of error identified previously, Table 4 also shows the *p-value* from a paired t-test on MAPEs and MALPEs.

In terms of accuracy or precision, the MAPEs from Model90-00 are larger in every state, except North Carolina, but there is considerable variation in the degradation of accuracy due to coefficient instability. Five states have numeric differences less than 0.25 (in an absolute sense) and fail to reject the null hypothesis of no difference at $\alpha = 0.05$. The impact of coefficient instability is most pervasive in Nevada and Texas with differences in MAPEs of 7.90 and 3.43, respectively. The MAPEs from Model00-10 are 60% and 42% smaller than the MAPEs from Model90-00 in these states, respectively. MAPEs for Oregon and Alaska have difference of

Table 4. Accuracy, Bias, and Allocation Error, Ratio Correlation Models 1990-2000 and 2000-2010

MAPE

Counties in	1990-2000	2000-2010	Numeric Diff.[a]	p-value[b]
Alaska	5.02	4.30	-0.72	0.052
California	2.21	2.19	-0.02	0.316
Colorado	7.38	6.00	-1.38	0.007
Illinois	4.07	2.67	-1.40	< 0.001
Nevada	14.14	6.24	-7.90	< 0.001
North Carolina	2.75	2.79	0.04	0.752
Oregon	3.57	2.90	-0.67	0.046
Texas	8.17	4.74	-3.43	< 0.001
Virginia	3.03	2.89	-0.14	0.086
Washington	3.18	2.94	-0.24	0.232
Wisconsin	2.53	2.38	-0.15	0.152
All Counties	5.07	3.62	-1.45	< 0.001

MALPE

Counties in	1990-2000	2000-2010	Numeric Diff.[a,c]	p-value[b]
Alaska	-2.91	-0.91	-2.00	< 0.001
California	-1.96	-1.96	0.00	0.512
Colorado	-3.82	-0.82	-3.00	< 0.001
Illinois	1.44	0.47	-0.97	0.021
Nevada	-14.14	-4.68	-9.46	< 0.001
North Carolina	-0.80	-1.14	0.34	0.999
Oregon	-0.83	1.08	0.24	0.999
Texas	6.69	1.37	-5.32	< 0.001
Virginia	0.39	0.19	-0.20	0.026
Washington	-0.26	-1.04	0.78	0.992
Wisconsin	0.56	0.03	-0.53	0.002
All Counties	1.26	0.04	-1.22	< 0.001

Index of Dissimilarity

Counties in	1990-2000	2000-2010	Numeric Diff.[a]
Alaska	1.27	1.22	-0.05
California	0.77	0.79	0.02
Colorado	2.22	2.09	-0.13
Illinois	1.36	1.16	-0.20
Nevada	1.25	0.52	-0.73
North Carolina	1.25	1.22	-0.03
Oregon	1.73	1.35	-0.38
Texas	1.77	1.46	-0.31
Virginia	1.22	1.25	0.03
Washington	0.88	0.85	-0.03
Wisconsin	0.81	0.97	0.16
All Counties	1.32	1.17	-0.15

[a] 2000-2010 Model - 1990-2000 Model.

[b] Ho: $\mu_{1990\text{-}2000} > \mu_{2000\text{-}2010}$, using paired t-test.

[c] Ignores the signs when different.

around -0.7 and Colorado and Illinois have differences of around -1.4. For all 904 counties, stable coefficients decrease the MAPE by-1.45 or -29%.

The picture is somewhat murkier for the impact of coefficient instability on estimate bias. In four states, bias is either the same (California) or increases under the Model00-10 (North Carolina, Oregon, and Washington) and the null hypothesis is not rejected. Similar to accuracy, bias is most affected by coefficient instability in Nevada and Texas. For all 904 counties, stable coefficients reduce the bias to near zero in Model00-10; it was 1.26 in Model90-00.

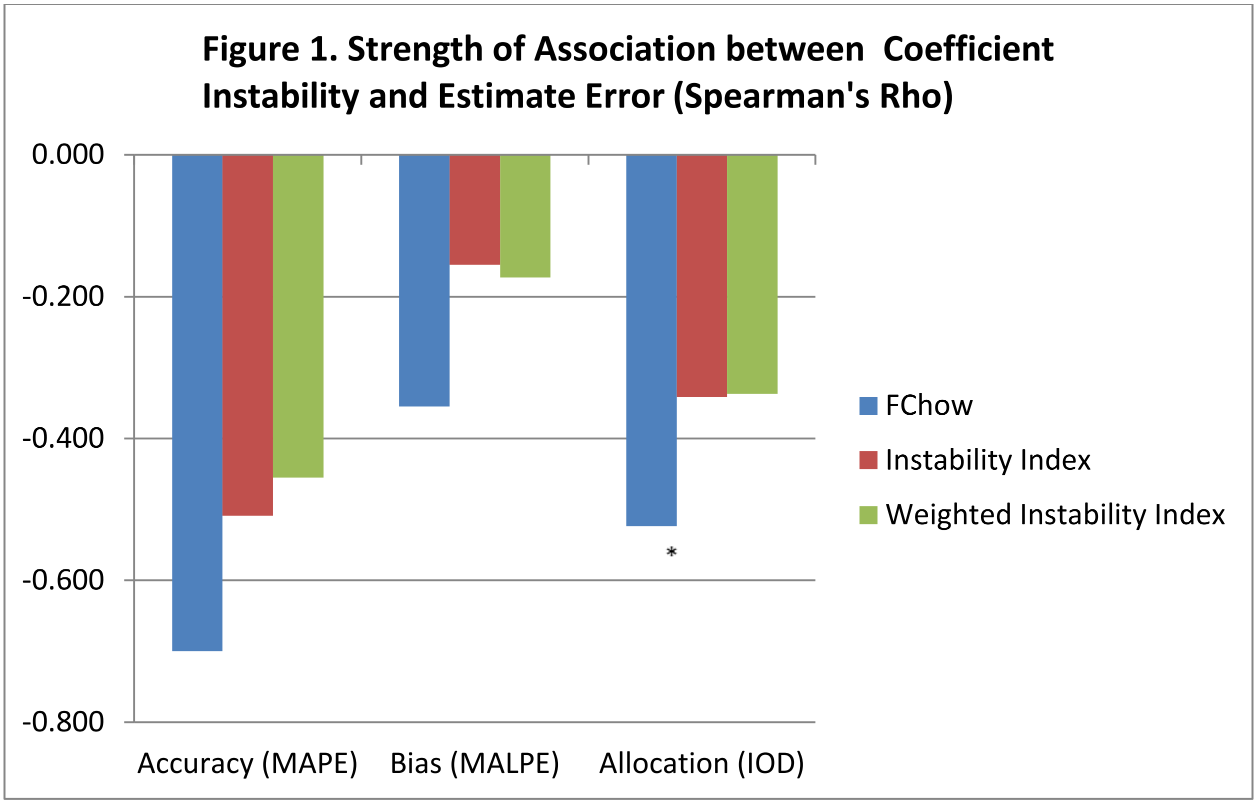
Allocation error is shown in the third panel of Table 4. The IOD is quite small for both models indicating that the ratio-correlation model has very low allocation error that is relatively insensitive to coefficient instability. The IOD increases slightly in Model00-10 in California and Virginia, and modestly in Wisconsin, but coefficient instability leads to modestly larger allocation errors in the other states; a lower IOD in Model00-10. The largest decrease in the IOD from Model90-00 to Model00-10 (-0.73) occurs in Nevada. Decreases of -0.20 or less are seen in five states and the IOD decreases by-0.38 in Oregon and -0.31 in Texas. For all 904 counties, the IOD is Model00-10 is lower by -0.15 compared to Model90-00.

The picture is somewhat murkier for the impact of coefficient instability on estimate bias. In four states, bias is either the same (California) or increases under the Model00-10 (North Carolina, Oregon, and Washington) and the null hypothesis is not rejected. Similar to accuracy, bias is most affected by coefficient instability in Nevada and Texas. For all 904 counties, stable coefficients reduce the bias to near zero in Model00-10; it was 1.26 in Model90-00.

Allocation error is shown in the third panel of Table 4. The IOD is quite small for both models indicating that the ratio-correlation model has very low allocation error that is relatively insensitive to coefficient instability. The IOD increases slightly in Model00-10 in California and Virginia, and modestly in Wisconsin, but coefficient instability leads to modestly larger allocation errors in the other states; a lower IOD in Model00-10. The largest decrease in the IOD from Model90-00 to Model00-10 (-0.73) occurs in Nevada. Decreases of -0.20 or less are seen in five states and the IOD decreases by-0.38 in Oregon and -0.31 in Texas. For all 904 counties, the IOD is Model00-10 is lower by -0.15 compared to Model90-00.

We have seen that the magnitude of coefficient instability and its impact on population estimate error varies from state to state. For example, the ratio-correlation model for Texas counties has the greatest degree of coefficient instability and also shows a large impact on estimate accuracy, bias, and allocation error. The model for North Carolina counties has little coefficient instability and, in turn, shows small impacts on estimate error. The model for Washington counties, however, has a substantial amount of coefficient instability, but its impact on estimation error is more muted.

What is the strength of the relationship between coefficient instability and estimate error? We address this question in Figure 1 that shows Spearman's Rho values between the three global measures of coefficient instability and three dimensions of estimate error for the 11 states. Estimate error variables are measured as the algebraic difference between Model00-10 and Model90-00 for each state. A larger difference (less negative to positive) indicates less impact of coefficient instability on estimate error. We anticipate an inverse relationship between the estimate error variables because for measures of coefficient instability larger values indicate greater coefficient instability.



As expected, increases in coefficient instability are associated with lower accuracy, greater bias, and greater allocation error. F_{Chow} shows the strongest association with all measures of error, while for *isi* and *wtisi* the *Rho's* are generally close in value. The *Rho* between the F_{Chow} and MAPE is the strongest (-0.700) and the relationships between other measures of coefficient instability and various dimensions of estimate error are small to moderate in strength. They range from -0.455 to -0.509 for accuracy, -0.173 to -0.355 for bias, and -0.337 to -0.524 for allocation error. These results supports the claim by Tayman and Schafer [43] that coefficient instability does impact on estimate error, but its impact is more muted than would be suggested by theory and the literature. Moreover our findings suggest that coefficient instability has the least impact on estimation bias, a larger impact on allocation error, and the greatest impact on estimate accuracy

Impact of Coefficient Instability on Estimate Error by Size and Growth Rate

Finally, we address the relationship between estimate error resulting from coefficient instability and the size and growth rate of counties. To our knowledge, this topic has not received any attention in past studies of coefficient instability. Table 5 contains the crosstabulation of the binary indicator that compares the absolute percent error (*ape)* between Model90-00 and Model00-10 for each county against size and growth rate categories. In the overall sample of 904 counties, Model00-10 that eliminates coefficient instability has a lower *ape* almost 61% of time. Still, in a sizable number of counties, Model00-10 creates larger errors compared to Model90-00 based on parameters from the decade prior to the simulated postcensal period. Size has a weak but stronger and statistically significant relationship compared to growth rate (*p value* = 0.02 vs *p value* = 0.50). The measures of association are roughly 2.5 times larger for size than for growth rate.

Model00-10 out-performs in 67.4% of the counties <5,000 persons and declines to 60.4% for counties 20,000-49,999. For counties with populations 50,000 or more, the percent is near 50%, meaning that Model00-10 and Model90-00 have roughly the same performance on this criteria. The odds ratios and *p-values* show the only statistically significant effect relative to the reference groups are in counties with less than 20,000 persons. For example, the odds ratio for the smallest counties (0.526) indicates that odds of Model00-10 having a higher error than Model90-00 are roughly half the odds of counties with over 200,000 persons. There is much less variation across growth rate categories in the percent of counties where Model00-10 performs best. The variation ranges from 57.1% for counties that grew 20.0-49.9% to 65.6% for counties that declined by more than five percent. None of the odds ratios for any growth rate category is statistically significant.

Table 5. Cross-tabulation of 2000-2010 Model Error by Size and Growth Rate

	2000 Population						
	< 5,000	5,000 - 19,000	20,000 - 49,999	50,000-99,999	100,000 - 199,999	200,000+	Total
Lower Error[a]	67.4%	66.2%	60.4%	51.3%	56.0%	52.1%	60.6%
Odds Ratio[b]	0.526	0.555	0.712	1.030	0.854	n/a	
p-value	0.032	0.014	0.164	0.913	0.610		
Sample Size	95	290	235	113	75	96	904
Chi-Square	13.299	p = .021					
Kendal's Tau-c	0.126						
Spearman's r	0.114						

	1990-2000 Growth Rate						
	< -5.0%	-4.9 - 4.9%	5.0 - 9.9%	10.0 - 19.9%	20.0 - 34.9%	35%+	Total
Lower Error[a]	65.6%	64.7%	60.6%	57.3%	57.1%	63.3%	60.6%
Odds Ratio[c]	0.905	0.942	1.124	1.287	1.300	n/a	
p-value	0.770	0.821	0.677	0.321	0.331		
Sample Size	64	204	137	246	163	90	904
Chi-Square	4.368	p = .498					
Kendal's Tau-c	0.049						
Spearman's r	0.044						

[a] Compared to the absolute percent error from 1990-2000 model.
[b] 200,000+ is the reference group.
[c] 35.0%+ is the reference group.

For a more detailed look at the relationship between size, growth rate, and error, due to coefficient instability, Table 6 shows the statistics from the binary logistic regression models. Models 1 and 2 are univariate models with size and growth rate as the independent variable. Model 3 includes size and growth rate together and Model 4 adds a size and growth rate interaction term to Model 3.

Table 6. Binary Logistic Regressions of 2000-2010 Model Error[a]

Model Estimates	Model 1	Model 2	Model 3	Model 4
Constant	-1.894 (0.000)	-0.470 (0.000)	-1.919 (0.000)	-2.080 (0.000)
Size Slope[b]	0.142 (0.002)		0.146 (0.002)	0.162 (0.004)
Growth Rate Slope[c]		0.003 (0.470)	-0.001 (0.790)	0.012 (0.615)
Size × Growth Rate Slope				-0.001 (0.579)
-2 Log Likelihood	1,202.2	1,211.6	1,202.1	1,201.8
Cox & Snell r^2	0.011	0.001	0.011	0.011
Nagelkerke r^2	0.015	0.001	0.015	0.015

Notre: p-values in parenthesis.

[a] Dependent Variable Coding:

0 = 2000-2010 model lower absolute percentage error than 1990-2000 model

1 = 2000-2010 model higher absolute percentage error than 1990-2000 model

[b] Natural logarithm of 2000 population.

[c] % change in population 1990-2000.

The binary logistic regression models further confirm the findings in the cross-tabulation. Size is the only variable to have a significant impact on the Model00-10 binary variable; although the magnitude of the impact is weak. There is no effect from either growth rate or the interaction of size and growth rate. The two r^2 measures for Model 1 containing only age are higher than Model 2 that contains only growth rate and, in addition, are identical with the two multivariate models. The -2 Log Likelihood value for Model 4 (1,201.8) is marginally smaller than the value for Model 1 (1,202.2), but the difference is not statistically significant ($P[\chi^2\ (2) > 0.6] = 0.74$).

The coefficient for age in Model 1 is 0.142 and its corresponding odds ratio is 1.153 ($\exp^{0.142}$). The positive slope indicates that a larger population size increases the likelihood that Model 00-10 will have a larger absolute percent error than Model90-00. In particular, for every one-unit increase in the natural logarithm of population there is about a 15% increase in the odds of Model00-10 having a larger error than Model90-00. Based on the coefficients in Model 4, we see that for a one unit increase in the natural logarithm of population, there is a 16.2% increase in the odds; for a one unit increase in the growth rate a 1.2% increase in the odds; and for a one-unit increase in the interaction term a -0.001% decrease in the odds, which further illustrates the greater effect of population size.

Using Model 1, we also can compute the probability of Model00-10 having a larger absolute percent error than Model90-00 (prob(larger)) for counties of a given size using the following formula [16: 10]:

Prob(larger) = exp(-1.894 + (0.142* ln(pop)) / 1+ exp(-1.894 + (0.142* ln(pop)).

For example, prob(larger) values are 0.215 and 0.596 for counties with the smallest (67) and largest (9,519,338) populations in the sample.

Conclusions

It has long been thought that coefficient instability is the major source of error in ratio-correlation and other regression models used in population estimation. To some extent this is true, but cannot be taken as an absolute. One should not assume that regression coefficients change substantially between the estimation and postcensal periods. Using various measures of coefficient instability, we found that five of the 11 sample states had only marginal changes in their regression coefficients. In the other states, there was considerable variability in the magnitude of their coefficient instability; Illinois and Texas exhibited by far the greatest change in their coefficients. This study points out the value of using more robust samples to study regression models of population estimates. Results gained from a single state, the usual approach used to study coefficient instability, may not be applicable to other contexts. Our findings would have been quite different had we analyzed for example just North Carolina (stable coefficients) or Texas (unstable coefficients).

Importantly, we find there is an impact of coefficient instability on estimation error. We found the greatest effect was on the accuracy of population estimates; in general, greater coefficient instability was associated with larger absolute percent errors. The association between the measures of coefficient instability and estimate accuracy over the 11 states was moderate in strength, suggesting a more muted impact than is commonly assumed. We evaluated two other dimensions of estimation error—bias and allocation error. The degree of coefficient instability was not a good predictor of bias. In some cases, the MALPE was greater in the model that controlled for coefficient instability (Model00-10) and the association between coefficient instability and bias was negligible. The ratio-correlation method was especially robust in terms of allocation error. Greater coefficient instability was generally associated with larger allocation errors, but the IODs were very small for both ratio-correlation models for all states.

Does coefficient instability have a different effect on estimate error depending on a county's size and growth rate? We examined this question using models for individual county errors. Using crosstabulation and binary logistic regression techniques, we analyzed the likelihood for individual counties that the model controlling for coefficient instability (Model00-10) would have a lower absolute percent error than the model estimated prior to the postcensal period (Model90-00). We found overall that Model00-10 did outperform Model90-00 according to this criterion, but Model00-10 did have larger errors in 40% of the counties. A small but significant relationship exists between size and the likelihood of outperformance by Model00-10, but no relationship is seen with growth rate. Specifically, a smaller population size increases the likelihood that Model00-10 will have a smaller absolute percent error. The likelihood of outperformance by Model00-10 was greatest and statistically significant in counties under 20,000 persons.

We suggested an individual measure of coefficient instability based on absolute percent difference between each regression coefficient in ratio correlation models based on two time periods (i.e., 1990-2000 and 2000-2010) along with a method of testing the statistical significance of individual coefficient changes, and two global measures based on the average of the absolute value of the percent change in the regression slopes. We refer to these global measures as the instability index (*isi*) and weighted instability index (*wtisi*). We believe *isi* and *wtisi* are useful additions to the F_{Chow} statistic. All three measures provided reasonable estimates of global coefficient instability and are closely related. F_{Chow} can be used to statistically test the hypothesis of no change in a set of regression coefficients. *Isi* and *wtisi* have a more intuitive and familiar interpretation the than $F_{Chow.}$ We prefer *wtisi* because its weight, based on the size of the regression coefficient, will lessen the influence of small numeric and large percentage changes on the average.

We believe there is value of having alternative measures of global coefficient instability and an instability measures for individual coefficients to help identify the specific variable or variables most affecting the instability of the ratio-correlation model. In addition, measuring both description changes and conducting statistical tests on these changes provide a more comprehensive assessment of coefficient instability.

This research calls into question the findings that various approaches for mitigating the impact of coefficient instability (e.g., averaging univariate models, variable transformation, and stratification) generally showed marginal improvements in estimation error. These studies analyzed counties in a single state and did not ascertain the degree of coefficient instability before applying modifications to the ratio-correlation model. For example, in situations where the regression coefficients were relatively stable, improvements from modifications that did mitigate the impact of coefficient instability would likely not be detected. Conversely, in situations where regression coefficients were dramatically unstable and the effect on error of the modification was marginal suggest that modification was not effective in mitigating the impacts of coefficient instability.

It would be useful to apply the modifications for improving regression-models of population estimate in counties within states where the coefficients show, respectively, stability, moderate instability, and high instability. The results of such an analysis would provide additional insights into the relationship between coefficient instability and population estimation error, and perhaps clarify the circumstances that would benefit most from modifications to the basic ratio-correlation model. Coefficient stability may well be the major source of error in regression models of population estimates in some contexts, but not in others. Information about the conditions that affect coefficient instability [20] and their impact on estimation error might lead to a more targeted and efficient approach for improving population estimates developed from regression models.

It would also be worthwhile to investigate whether additional insights may be obtained in regard to coefficient instability and its effects on population estimation accuracy by modifying the fundamental form of the ratio-correlation model using methods such as hierarchical linear and random slope models [10, 26]. Hierarchical models allow the introduction of random and/or fixed effects that account for between-area effects not accounted for by the symptomatic indicators. Random slope models may provide another way to examine coefficient instability and potentially allow the modeling those effects into the postcensal decade. The efficacy of these multi-level modeling techniques may be better suited to multiple nested geographic area applications of the ratio-correlation model such as when counties of multiple states are combined into a single model rather than as individual models for each shown in this paper. Ratio-correlation models are estimated using ordinary least square regression techniques that ignore any spatial dependencies in the data, which may lead to unstable parameter estimates and unreliable significance tests [46]. It would be useful then to examine spatial correlation issues and spatial regression models in the context of population estimation

Attempts at adjusting the fundamental form of the ratio-correlation model may come at a price, however. As discussed in Endnote 1, the fundamental form of the ratio-correlation method allows for substantive interpretations of coefficients that are very useful and their interpretations may be lost entirely or become less straight-forward under some modifications. We suggest keeping this trade-off in mind as to the practical application of alternative regression model forms for population estimation.

Endnotes

1. The methods and results shown in this study were produced using IBM SPSS Statistics Version 22 and Microsoft Excel 2010 on an Intel Core I7-3770 CPU desktop computer running at 3.40GHz with 8GB of memory, running Windows 7 64-bit Home Premium operating system.

2. Swanson and Prevost [40] note that the ratio-correlation method has some useful features not always recognized in discussions of this method and its accuracy. Unstandardized coefficients in a ratio-correlation model, including the intercept term, always sum to 1.00 or approximately so; therefore, each coefficient can be interpreted as a "proportionate weight" associated with a given symptomatic indicator and its change over time. The intercept term also represents a proportionate weight associated with an estimate of the current population in a given subarea based on the proportion of population relative to the parent area found in the subarea at the beginning of the postcensal decade; and the ratio-correlation model is fundamentally hierarchical in nature. Retaining this fundamental structure allows a user to assess the "proportionate weights" applied to the symptomatic estimators and their changes as well as the proportionate weight represented in the intercept term.

3. Measures based on the average are affected by extreme values. Robust or resistant statistics are alternatives to the average because they focus on the main body of the data and attempt to minimize the impact of outlying observations [14, 17]. We also conducted the analysis using the Tukey biweight, a popular and widely used M-estimator [11, 13], as robust alternatives to the MAPE and MALPE. While Tukey biweight values were generally smaller than the corresponding MAPEs and MALPEs, due to the tendency for the error distributions to be right skewed [45], the patterns, relationships, and conclusions were not substantively affected by outlying errors.

4. We also analyzed the parametric correlation coefficient (r), but it was unduly influenced by one outlying observation for Nevada.

5. We also compared MAPEs and MALPEs from the Model90-00 and Model00-10 models for different size and growth rate groupings across all counties, the most common way to evaluate the impact of characteristics on estimation error. Our findings from this aggregate analysis were generally consistent with the literature on the relationship between size, growth rate, and estimate error [12, 15, 32, 33]. The reduction in accuracy introduced by coefficient instability was inversely related to size, with the greatest impact seen in counties with less than 20,000 persons. There appears to be no relationship between the impact of coefficient instability on bias and population size. In several size groups, bias was greater in the model that controlled for coefficient instability (Model00-10). For growth rate, we see a U-shaped pattern between the impact of coefficient stability on estimate accuracy and bias, with the greatest effects in counties with growth rates of less than 5%.

Appendix
Unstandardized Regression Coefficients Instability Measures

	Estimation Period		Change		
	1990-2000	2000-2010	Number	Percent	P-Value
Alaska					
Permanent Fund Residents	1.054	0.904	0.150	14.2%	0.329
California					
Driver's License	0.802	0.754	0.048	6.0%	0.579
School Enrollment	0.177	0.191	0.014	7.9%	0.758
Colorado					
Voter Registration	0.330	0.349	0.019	5.8%	0.859
School Enrollment	0.502	0.278	0.224	44.6%	0.018
Employment	0.147	0.289	0.142	96.6%	0.179
Illinois					
School Enrollment	0.254	0.026	0.228	89.8%	<0.001
Vehicle Registration	0.282	0.249	0.033	11.7%	0.663
Births	0.150	0.191	0.041	27.3%	0.466
Tax Exemptions	0.111	0.287	0.176	158.6%	0.011
Nevada					
School enrollment	1.024	0.712	0.312	30.5%	0.052
North Carolina					
Vehicle Registration	0.549	0.473	0.076	13.8%	0.380
School Enrollment	0.209	0.243	0.034	16.3%	0.477
Births	0.154	0.175	0.021	13.6%	0.675
Oregon					
Tax Exemptions	0.411	0.302	0.109	26.5%	0.395
School Enrollment	0.361	0.212	0.149	41.3%	0.089
Medicare Enrollment	0.112	0.169	0.057	50.9%	0.540
Texas					
Vehicle Registration	0.279	0.110	0.169	60.6%	0.002
Voter Registration	0.125	0.279	0.154	123.2%	0.007
School Enrollment	0.206	0.451	0.245	118.9%	<0.001
Births	0.197	0.060	0.137	69.5%	<0.001
Virginia					
Driver's License	0.252	0.472	0.220	87.3%	0.031
Housing Units	0.614	0.417	0.197	32.1%	0.101
School Enrollment	0.121	0.094	0.027	22.3%	0.510
Births	0.068	0.098	0.030	44.1%	0.501
Washington					
Voter Registration	0.093	0.262	0.169	181.7%	0.064
Vehicle Registration	0.336	0.151	0.185	55.1%	0.099
School Enrollment	0.398	0.385	0.013	3.3%	0.872
Wisconsin					
Driver License	0.492	0.528	0.036	7.3%	0.762
Tax Exemptions	0.286	0.293	0.007	2.4%	0.955
School Enrollment	0.163	0.254	0.091	55.8%	0.056

References

[1] T. Bryan, "Population estimation," in *Methods and Materials of Demography, Second Edition*, J. S. Siegel and D. A. Swanson, Eds., ed Amsterdam: Elsevier, 2004, pp. 523-560.

[2] G. C. Chow, "Tests of equality between sets of coefficients in two linear regressions," *Econometrica*, vol. 28, pp. 591-605, 1960.

[3] A. H. Crosetti and R. C. Schmitt, "A method for estimating the inter-censal population of counties," *Journal of the American Statistical Association*, vol. 51, pp. 587-590, 1956.

[4] C. Dougherty, *Introduction to econometrics, Fourth edition*. Oxford: Oxford University Press, 2011.

[5] O. D. Duncan, R. P. Cuzzort, and B. Duncan, *Statistical geography: Problems in analyzing area data*. Glencoe: Free Press, 1961.

[6] E. P. Ericksen, "A method for combining sample survey data and symptomatic indicators to obtain population estimates for local areas," *Demography*, vol. 10, pp. 137-160, 1973.

[7] E. P. Ericksen, "A regression method for estimating population changes for local areas," *Journal of the American Statistical Association*, vol. 69, pp. 867-875, 1974.

[8] L. Fonseca and J. Tayman, "Postcensal estimates of household income distributions," *Demography*, vol. 26, pp. 149-160, 1989.

[9] D. Goldberg, V. R. Rao, and N. K. Namboodiri, "A test of the accuracy of ratio correlation population estimates," *Land Economics*, vol. 40, pp. 100-102, 1964.

[10] H. Goldstein, *Multilevel statistical models, 4th Edition*. Chichester: John Wiley, 2011.

[11] C. Goodall, "M-estimators of location: An outline of the theory," in *Understanding robust and exploratory data analysis*, D. C. Hoaglin, F. Mosteller, and J. W. Tukey, Eds., ed New York: John Wiley, 1983, pp. 339-403.

[12] G. Harper, C. Coleman, and J. Devine. (2003). *Evaluation of 2000 subcounty population estimates* [Online]. Available: http://www.census.gov/population/www/documentation/twps0070/twps0070.html.

[13] S. Heritier, E. Cantoni, S. Copt, and M. P. Victorua-Feser, *Robust methods in biostatistics*. Chichester: John Wiley, 2009.

[14] D. C. Hoaglin, F. Mosteller, and J. W. Tukey, "M-estimators of location: An outline of the theory," in *introduction to more refined estimators*, D. C. Hoaglin, F. Mosteller, and J. W. Tukey, Eds., ed New York: John Wiley, 1983, pp. 283-296.

[15] N. Hoque, "Evaluation of small area population estimates produced by housing unit, ratio-correlation, and component method II compared to 2000 Census counts," *Canadian Studies of Population*, vol. 39, pp. 91-108, 2010.

[16] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied logistic regression*. Hoboken: John Wiley, 2013.

[17] P. J. Huber and E. M. Ronchetti, *Robust statistics, Second edition*. New York: John Wiley, 2009.

[18] M. Mandell and J. Tayman, "Measuring temporal stability in regression models of population estimation," *Demography*, vol. 19, pp. 135-146, 1982.

[19] D. S. Massey and N. A. Denton, "The dimensions of residential segregation," *Social Forces*, vol. 67, pp. 281-315, 1988.

[20] J. N. McKibben and D. A. Swanson, "Linking substance and practice: A case study of the relationship between socio-economic structure and population estimation," *Journal of Economic and Social Measurement*, vol. 24, pp. 135-147, 1997.

[21] N. K. Namboodiri, "On the ratio-correlation and related methods of subnational population estimation," *Demography,* vol. 9, pp. 443-453, 1972.

[22] N. K. Namboodiri and N. M. Lalu, "The average of several regression estimates as an alternative to the multiple regression estimate in postcensal and intercensal estimates: A case study," *Rural Sociology,* vol. 36, pp. 187-194, 1971.

[23] W. P. O'Hare, "Report on a multiple-regression method for making population estimates," *Demography,* vol. 13, pp. 369-379, 1976.

[24] W. P. O'Hare, "Report on a multiple-regression method for making population estimates," *Demography,* vol. 17, pp. 341-343, 1980.

[25] D. E. Pursell, "Improving population estimates with the use of dummy variables," *Demography*, vol. 7, pp. 87-91, 1970.

[26] J. N. K. Rao, *Small area estimation*. Hoboken: Wiley Interscience, 2003.

[27] H. Rosenberg, "Improving current population estimates through stratification," *Land Economics,* vol. 44, pp. 331-338, 1968.

[28] R. C. Schmitt and A. H. Crosetti, "Accuracy of the ratio-correlation method for estimating postcensal population," *Land Economics,* vol. 30, pp. 279-281, 1954.

[29] R. C. Schmitt and J. M. Grier, "A method for estimating the population of minor civil divisions," *Rural Sociology,* vol. 31, pp. 355-361, 1966.

[30] W. J. Serow and J. H. Martin, "Estimating demographic characteristics using the ratio-correlation method," *Demography,* vol. 15, pp. 223-233, 1978.

[31] J. S. Siegel, *Applied demography: Applications to business, government, law, and public policy*. San Diego: Academic Press, 2002.

[32] S. K. Smith and S. Cody, "An evaluation of population estimates in Florida: April 2010," *Population Research and Policy Review,* vol. 23, pp. 1-24, 2004.

[33] S. K. Smith and S. Cody, "An evaluation of population estimates in Florida," Bureau of Economic and Business Research, Gainesville, 2011.

[34] E. C. Snow, "The application of the method of multiple correlation to the estimates of postcensal populations," *Journal of the Royal Statistical Society,* vol. 74, pp. 575-629, 1911.

[35] M. A. Spar and J. H. Martin, "Refinements to regression-based estimates of postcensal population characteristics," *Public Data Use,* vol. 7, pp. 16-22, 1979.

[36] D. A. Swanson, "An evaluation of ratio and difference regression methods for estimating small, highly concentrated populations," *Review of Public Data Use,* vol. 6, pp. 18-27, 1978.

[37] D. A. Swanson, "Improving accuracy in multiple regression estimates of county populations using principals of causal modeling," *Demography,* vol. 17, pp. 413-427, 1980.

[38] D. A. Swanson, "Allocation accuracy in population estimates: An overlooked criteria with fiscal implications," in *Small area population estimates and their accuracy*, U.S. Bureau of the Census, Ed., ed Washington, D.C.: U.S. Bureau of the Census, 1981, pp. 13-21.

[39] D. A. Swanson, "Advancing methodological knowledge within state and local demography: A case study," *Population Research and Policy Review,* vol. 23, pp. 379-398, 2004.

[40] D. A. Swanson and R. Prevost, "A new technique for assessing error in ratio-correlation estimates of population: A preliminary note," *Applied Demography,* vol. 1, pp. 1-4, 1985.

[41] D. A. Swanson and J. Tayman, *Subnational population estimates*. Dordrecht: Springer, 2012.

[42] D. A. Swanson and L. M. Tedrow, "Improving the measurement of temporal change in regression models used for county population estimates," *Demography,* vol. 21, pp. 373-381, 1984.

[43] J. Tayman and E. Schafer, "The impact of coefficient drift and measurement error on the accuracy of ratio-correlation population estimates," *The Review of Regional Studies,* vol. 15, pp. 3-10, 1985.

[44] J. Tayman, S. K. Smith, and S. Rayer, "Evaluating population forecast accuracy: A regression approach using county data," *Population Research and Policy Review,* vol. 30, pp. 235-262, 2011.

[45] J. Tayman and D. A. Swanson, "On the validity of the MAPE as a measure of population forecast accuracy," *Population Research and Policy Review,* vol. 16, pp. 299-322, 1999.

[46] M. D. Ward and K. S. Gleditsch, *Spatial regression models*. Thousand Oaks: Sage, 2008.