FAMILY ESTIMATES EMPIRICAL STUDY

Alfredo Navarro, Richard A. Griffin and Lawrence M. Bates
Bureau of the Census

I. INTRODUCTION

From the 1990 Census sample data, three types of estimates will be formed: population estimates, housing estimates, and family estimates[1]. Iterative Proportional Fitting (IPF) or "raking" ratio estimation will be used in 1990 to produce a weight for each person and housing unit in the 1990 Census sample. These weights are intended to be used to produce estimates for person and housing unit characteristics. Estimates for person and housing unit characteristics are then obtained by summing the weights associated with each person or housing unit possessing the characteristic of interest.

Traditionally, there has been a strong interest in census estimates of family characteristics. To obtain estimates of family characteristics, a family weight must be assigned to each family. Estimates of family characteristics are then obtained by summing the weights associated with the families possessing the characteristic of interest. In the 1980 Census, the weight of the householder of a family was used to produce family estimates. This is known as a principal person weighting method, where the weight of an individual person in the family is used as the weight of the family. A method analogous to the principal person weighting method is the method that uses the weight of the housing unit as the family weight. Another way of deriving a family weight involves using an "average" of the weights of the members of the family as the family weight rather than the weight of an individual family member [1].

In this study, we examined eight different methods for determining the family weight. The methods were compared based on several empirical statistics. These empirical statistics were calculated for a set of family characteristics in 120 (1980) weighting areas. Weighting areas are generally formed of contiguous portions of geography, which closely agreed with census tabulation areas within counties. Weighting areas are required to have a minimum sample population. The averages of the empirical statistics across weighting areas were obtained and compared. The Friedman Test was used as a screening tool to compare the methods[2]. A pseudo-measurement of bias was also calculated for several 100 percent family characteristics and a bias analysis was conducted using these data.

II. FAMILY WEIGHT DEFINITIONS

Three general methods for determining a family weight were considered. They are:

1. Use of principal person weights
2. Use of "average" person weights
3. Use of the housing unit weights

There are three types of principal person weights that were considered: the use of the householder as the principal person, the use of the female spouse as the principal person in married couple families, and the use of the "best covered" family member as the principal person. Four types of "average" weights were also considered. They were the simple arithmetic mean, a weighted mean, the geometric mean, and the harmonic mean of the weights associated with each family member. The use of the housing unit weight as the weight of the family was also examined.

A. Use of Principal Person Weights

In principal person weighting, the weight of a single "principal" person is chosen to represent the family and this person's weight becomes the family weight.

1. Use of the Householder as the Principal Person-Method 1.
   This procedure uses the householder as the principal person. It was used for the 1980 Census because it was both simple and practical. The householder can be either male or female. It is not clear whether it is a problem that the person designated as the householder is usually the male in married-couple families. However, surveys such as the Consumer Expenditure Survey apply an "adjustment factor" to the weight of the "principal person" if this happens to be a male. The rationale for this is that historically males have been undercovered by surveys and censuses.
2. Use of the Female Spouse as the Principal Person-Method 2.
   The Census usually has a higher coverage for women than for men.[2] How this affects the production of family estimates is unclear; however, the use of the female spouse (when applicable) as the principal person might produce better quality of family estimates than the householder's weight. This procedure involves using the female spouse as the principal person in married couple families and using the householder as the principal person in all other families.
3. Use of the "Best Covered" Family Member as the Principal

Person-Method 3.
This procedure uses the "best-covered" member in the family as the principal person. A person is considered to be the "best-covered" if that person is in the population group better represented in sample. Assuming perfect sampling and response, the choice of this weight should yield estimates with an insignificant level of bias.

B. Use of Average Person Weights

An alternative to using a principal person weight is to use an average weight of the person weights of all persons considered to make up a family. This average weight can then be used for tabulating family data.

1. Use of Arithmetic Mean of Person Weights-Method 4.
The arithmetic mean of person weights is defined as:

$$AM_i = \sum_{j=1}^{a_i} W_{ij}/a_i$$

where W_{ij} = the weight of the j^{th} person in the i^{th} family, and a_i = the number of family members in the i^{th} family.
The arithmetic mean is the expected value of the weight assuming a member of the family is selected with equal probability to represent that family.
Let P_{ij} be the sampling probability of selection of the jth person in the ith family and Wij as defined before. Further suppose $P_{ij} = (W_{ij})^{-1}$. Let X_i be the value of some characteristic of the i^{th} family. Note that the census sample is a cluster sample of persons and not of families.
An unbiased estimate of the aggregate value of the characteristic is [5]

$$\hat{X} = \sum_{i=1}^{n_f} \sum_{j=1}^{a_i} W_{ij} X_i/a_i = \sum_{i=1}^{n_f} (X_i/a_i) \sum_{j=1}^{a_i} W_{ij}$$

Note that $\sum_{j=1}^{a_i} W_{ij}/a_i = \bar{W}_i$ is the arithmetic mean of the weights of the i^{th} family.

The expected value of $\hat{X}$,

$$E(\hat{X}) = \sum_{i=1}^{N_f} \sum_{j=1}^{a_i} \frac{1}{W_{ij}} W_{ij} X_i/a_i$$

$$= \sum_{i=1}^{N_f} \sum_{j=1}^{a_i} X_i/a_i = \sum_{i=1}^{N_f} X_i = X$$

Thus, the use of the arithmetic mean of the weights of the family members produces an unbiased estimate of the total.

2. Use of a Weighted Mean of Person Weights-Method 5.
The weighted mean of person weights is defined as:

$$WM_i = \sum_{j=1}^{a_i} \frac{W_{ij}^2}{\left[\sum_{j=1}^{a_i} W_{ij}\right]}$$

The weighted mean is the expected value of the weight assuming a member of the family is selected with probability proportional to size. It is intuitively obvious that this method will yield a biased estimate of the total. Members of the family in population groups that are poorly covered by the census sample will be assigned a weight larger than the average. Families with members in these poorly covered groups will tend to have larger weights. One strategy to solve this problem would involve identifying race/age/sex groups substantially underrepresented in the census sample living with persons who are efficiently covered by the sample and have their weights adjusted less.

3. Use of Geometric Mean of Person Weights-Method 6. The geometric mean of person weights is defined as:

$$GM_i = \left[\prod_{j=1}^{a_i} W_{ij}\right]^{1/a_i}$$

Suppose W_f is the best (in terms of variance and bias) family weight and W_i is the weight of the i^{th} person in the family. Suppose there is a vector $\underset{\sim}{c} = (c_i)$, $(i=1,2,\ldots,a_i)$ such that $W_i C_i = W_f$ and in general assume that $\prod_{i=1}^{a_i} C_i = 1$. (i.e. if one person has a weight double the best weight another might have a weight one half the best weight).
It follows that

$$\prod_{i=1}^{a_i} W_i C_i = (W_f)^{ai}$$

and consequently

$$\prod_{i=1}^{a_i} W_i = (W_f)^{ai}$$

yielding

$W_f = (\prod_{i=1}^{a_i} W_i)^{1/ai}$, the geometric mean of the weights of the members of the family.
The geometric mean may correct for some of the biases when using the weighted mean of the weights as the weight of the family if the above model holds.

4. Use of Harmonic Mean of Person Weights
The harmonic mean of person weights is defined as:

$$HM_i = \left[\frac{a_i}{\sum_{j=1}^{a_i} 1/W_{ij}}\right]$$

The harmonic mean can be thought of as the inverse of the expected value of the selection probabilities if the representative family member is selected with equal probability.

$$(HM_i)^{-1} = \frac{1}{a_i} \sum_{j=1}^{a_i} 1/W_{ij}$$

The probability of selection for the i^{th} member of the j^{th} family is given by $P_{ij} = (W_{ij})^{-1}$.

The relative performance of these methods have been studied and although the harmonic and perhaps the geometric mean results in smaller biases than variations of the principal person method, there are obvious advantages in using a principal person type of strategy for family estimation [1].

C. Use of the Housing Unit Weight-Method 8.

This method uses the weight of the housing unit to tabulate family data. This procedure is analogous to the principal person weighting methods described in part II.A. The rationale here is that overall the coverage of occupied housing units is better than that of some demogaphically defined population groups and, therefore, this excellent coverage should be transferred to the estimation of family characteristics. The use of the housing unit weight should yield estimates of family characteristics with very small biases.

III. EMPIRICAL STUDY

For this study, family data from all housing units in 120 weighting areas were used to determine which method of assigning family weights is "best". Several criteria were used to determine the "best" method of assigning family weights.

The following empirical statistics were examined:

1. Variance
2. Coefficient of variation (CV)
3. Absolute Bias
4. Relative Bias
5. Relative Absolute Bias
6. Root Mean Square Error (RMSE)

Notation

Let X = Census Count.

$\hat{X}_i$ = Estimate using the i-th method.

then the

Absolute Bias = $|X-\hat{X}_i|$

Relative Bias = B/σ_i ; σ = standard error of the estimate

$$\text{Relative Absolute Bias} = \frac{|X-\hat{X}_i|}{\hat{X}_i}$$

$$\text{RMSE} = \sqrt{\text{VAR}(\hat{X}_i) + (X-\hat{X}_i)^2}$$

Sample data from the 1980 Census were used to determine which method of assigning family weights provides estimates with the lowest variance and coefficient of variation. Random groups variance estimates were computed for each of 85 data items for each of the 8 methods for assigning family weights. A quantitative and a qualitative analysis was performed. (This analysis is explained in more detail in the next section).

The 85 data items consist of 20 income items, 19 items of labor force, 20 poverty items, and 26 100 percent data items including items on family type, number of families by race of the householder and others.

To determine the method with the lowest bias and RMSE, we used 100 percent data from the 1980 Census. Sample estimates of 26 100 percent data items were computed.

IV. LIMITATIONS OF THE STUDY

In terms of the bias of the estimates, we used a "pseudobias" measurement. We defined bias as the difference between the estimate and the 100 percent count (the bias is equivalent to the "pseudobias" here if $\hat{X} = E(\hat{X})$, which is true only if the estimate has no sampling variability). Therefore, our findings should be taken as indicative or suggestive, but not conclusive. Futhermore, this assumption raises some questions as to the validity of using the empirical study findings to evaluate the bias-reduction properties of a particular method.

V. ANALYSIS

A. Qualitative Analysis - Friedman Test

The Friedman test was used to determine if there were any significant differences with respect to the empirical statistics provided by the eight different weighting methods. For each data item, a random groups variance estimate was computed for each method of assigning a family weight for each weighting area. These variance estimates were ranked within each weighting area from least to greatest. The Friedman test was then performed at the 5 percent significance level.

The same test was performed for the CV, bias, RMSE, the relative bias and the relative absolute bias.

B. Quantitative Analysis

1. Variance

In order to determine which weighting method provides estimates with the lowest variance, a quantitative examination of the data was performed.

One way in which we compared the variances involves comparing the mean of the variances for each data item in a pairwise fashion. For example, assume that for a given pair of weighting methods there are 50 rejections (out of 60 data items) of the null hypothesis (H_0: $\sigma_u=\sigma_v$) in the pairwise comparison tests. To assess which of the two methods provides estimates with the lower variance, we computed the average of the variances across all weighting areas for a particular data item for each method and compared them. We also looked at these averages for each pair of weighting methods for each data item.

A second scheme involved comparing the median of the variance estimates for each weighting method. This was done in a similar fashion to that described above for the comparison of the means. Looking at the results from both the Friedman test and the quantitative analysis we determined which weighting method is superior, if any.

2. Bias Analysis

An analysis similar to the variance analysis was performed to determine which method of assigning weights provides estimates with the lowest absolute bias.

Using the results of the Friedman test and the quantitative analysis, we determined which method of assigning family weights provides estimates with the smallest bias.

C. Additional Criteria

In addition to the variance and bias criteria, the methods were compared in terms of consistency between family estimates and estimates of individuals. The following relationships were investigated for each of the methods:

1. The total population in families plus the unrelated individuals should equal the total household population.
2. The estimated number of married couple families should be equal to the number of such families in the census.
3. Total income of families by size should equal the total income of individuals in families of the corresponding size.

Although there are many other important relations, we limited the number in this study to those three.

VI. SUMMARY OF RESULTS[3]

A. Nonparametric Analysis

As described in Section V.A.1, a nonparametric analysis was done in this study. The analysis was used to identify any method whose mean rank across individual weighting areas is either large enough or too small to be unlikely under the null hypothesis of no difference among methods. The Friedman Test was performed on the variance and CV of the estimates for all sample data items. The test was also performed on the variance, CV, bias, relative bias and RMSE of the estimates for the 100 percent data items. First, we discuss the results of the test on the sample characteristics and then the results based upon the 100 percent data items.

1. Variance

The analysis of the mean ranks indicate that the variance of the estimates produced using method 7 ranked significantly lower than that of the other methods for a high proportion of the total data items. In fact the harmonic mean method ranked significantly lower for 56 of the 59 sample data items and for 19 of the 26 100 percent data items. This same pattern was observed for each of the data item groups.

The weighted mean and the (surprisingly) householder weight generally yielded a fairly large number of data items having a significantly high mean rank for the variance. A similar pattern holds for the 100 percent data items. The weighted mean ranked significantly higher for 81 of the 85 data items. The best covered person method did not perform any better than any of the other "principal person" methods.

The null hypothesis of no difference between the methods was rejected for 48 of the 59 sample data items and 19 of the 26 100 percent data items. A pairwise comparison test was performed for these 67 characteristics. The analysis showed a statistical significant difference between method 7 and methods 1, 2, and 3 for 38 data items; 10 income data items, 16 items of labor force, and 12 items of poverty.

For each pair of methods and weighting area, the ratio of the variances was calculated for each of the 38 characteristics. The distribution of the ratios was created and the 25, 50 (Median), and 75 percent quartiles were examined. The data in the table below indicates that the variance for method 7 is somewhat lower (between 92-97 percent) than the varaince for the other methods for at least 50 percent of the weighting areas (see table next page).

2. Coefficient of Variation

The null hypothesis of no difference between the methods was only rejected for 19 of the 85 data items. The null hypothesis was only rejected for 2 items of income and 2 poverty data items. Generally, the mean ranks were very close to the expected mean rank of 4.5. However, method 8 (Housing Unit Weight) showed a consistently high mean rank while the arithmetic mean and the weighted mean yielded consistently low mean ranks. Method 2 yielded a fairly large number of data items with a mean rank greater than the expected value of the ranks.

For the labor force group, we performed a pairwise comparison test and found a statistical significant difference between method 8 and 4, for 4 out of the 5 characteristics for which the null hypothesis was rejected. In general, the mean ranks suggest that the estimates produced by method 4 are somewhat more reliable than family estimates produced using the weight of the housing unit.

3. Other empirical statistics.
The analysis discussed in this section was performed on the bias, relative absolute bias, and the RMSE of the estimates for 26 100 percent characteristics.
The null hypothesis of no difference between the methods was rejected for almost half of the data items.
The data on the mean ranks indicate that the bias for methods 4 and 6 ranked significantly lower than that of the other methods for a large proportion of data items. The weighted and harmonic mean yielded a high number of data items having a significantly high mean rank. The apparent reduction in variance by method 7 does not translate into a significant gain in terms of the reliability of the estimates. The analysis based on CV's did not allow us to detect any difference between the methods. These data suggest that method 7 produces estimates with a larger bias than methods 1, 4 and 6.
In general, method 8 performed well, however, for estimates on the number of householders by age categories showed a consistently high mean rank.
A similar pattern was observed for the relative absolute bias and the RMSE.
In summary, the nonparametric analysis indicated that the use of the harmonic mean produces estimates with somewhat smaller variance but larger bias than the other methods. The apparent gains in variance are not large enough to translate into observable gains in terms of reliability of the estimate.

B. Quantitative Analysis

The quantitative analysis was performed by comparing a set of empirical statistics (variance, CV, etc.) that resulted from the production of family estimates for each of the methods. Tables 8 to 19 summarize our findings by data item and by data item group.

1. Variance
The analysis of the data showed that method 7 can reduce the variance of the estimates by as much as 8 percent when compared to method 1. The reduction in variance for poverty data items is about 7.6 percent while for items on income is about 5 percent. Method 7 has the smallest mean variance for almost all of the data items. Method 8, the housing unit weight produce estimates with the largest variance. However, the increase in variance over method 1 is not of any practical concern (2.5 percent at most).

2. Coefficient of Variation
The reductions in variance by method 7 are not translated into lower average CV's. There is no numerical difference between the eight methods with respect to the CV's calculated at two decimal places.

3. Other Empirical Statistics
The analysis discussed in this section is based upon the bias, relative absolute bias, and RMSE of the estimates for 26 100 percent data items.
The results suggested that method 4 and 6 have very interesting bias-reduction properties. These methods can reduce the bias for some of the characteristics by as much as 40 percent. The reduction in bias across all data items was found to be about 11 percent (method 4 compared to method 1). The average bias for method 6 is slightly over 9 percent less than for method 1. The average bias for method 7 is about 4 percent larger than the average bias for method 1. The average median bias for method 4 and 6 are smaller than for method 1.
Although reduction in bias can be realized by method 4 and 6, they are not as significant as the data on the absolute bias seemed to suggest. The same pattern was observed for the RMSE.
We also compared method 4 and 6 with method 1 in terms of the bias and RMSE. For each of the data items and weighting areas the ratio.
R_1 = Bias (4)/Bias(1) and
R_2 = Bias (6)/Bias(1)
were calculated. Similar ratios were calculated for the RMSE. These ratios were averaged across all weighting areas for each of the data items[3]. The results are summarized in Figures 2 and 3 below. For a majority of the data items (19 out of 26) the median ratio of the bias between method 4 and method 1 was 1.0000. For 16 of the data items the median ratio between method 6 and 1 was 1.0000. For most of the remaining data items the median ratios were less than 1. Similar observations hold for the RMSE.
This analysis suggests that neither method 6 nor method 4 are clearly superior to method 1 in terms of bias and RMSE. For a majority of the data items the mean median ratio is exactly 1.0000.

C. Additional Criteria

The methods were also compared in terms of the three criteria listed in Section V.C. For each method, the appropriate estimate was computed for each weighting area. For instance, for criterion number 2, the estimate of married couple families was computed for each of the methods and weighting areas. The relative absolute difference between the estimate and the 100 percent count was computed for each of the weighting areas. The relative absolute difference is defined as:

$$RA_{ij} = \frac{|X_c - \hat{X}_{ij}|}{X_c}$$

where

X_c - census count

$\hat{X}_{ij}$ - I-th method estimate in the j-th weighting area.

The RA's were averaged across all weighting areas.

For instance method 1 produces an estimate of the total population in families that is off by 2.06 percent of the true count. Method 4 has the lowest mean relative absolute difference for 2 of the three criteria. These data suggest the bias-reduction properties of method 4 one more time. The performance of method 4 for estimating total population in families (in the census context) is clearly superior to the other methods. A note of interest is that for estimating total income by family size categories, method 1 is clearly superior to the other methods. The worst method is the housing unit weight, it yielded an increase of about 400 percent over the bias for method 1. Method 4, its closest competitor yielded an increase in the bias of 90 percent. Method 6 and 7 produced underestimates of the number of married couple families while method 5 produced an overestimate. In summary, the performance of method 1 is very good for all three criteria. However, methods 4 and 6 are superior (with respect to bias) to method 1 for estimating type of families and total population in families.

VII. CONCLUSIONS

In conclusion, no single method emerged as clearly superior to the others with respect to the comparison criteria used in this study. Some of the data examined suggests the potential of the arithmetic mean (and perhaps the geometric mean) as a bias-reduction method. The data also suggest that the apparent reduction in bias is realized without increasing the variance of the estimates. To fully assess the bias - reduction properties of the arithmetic mean (and perhaps the geometric mean) more empirical research in this area is necessary and recommended.

A surprising finding was the poor performance of the "best covered person" method when compared to other "principal person" methods.

These findings do not by any means justify the introduction of an additional weighting operation for family estimates for the 1990 Census, for what would probably be marginal gains in bias reduction.

ACKNOWLEDGEMENTS

The authors are grateful to Steve Jarvis for his careful preparation of the input files for this study. They also want to thank Ardi Milstead for her expertise and careful preparation of this manuscript.

FOOTNOTES

[1] A household constitues a family it there are two or more persons, including the householder, who are related by birth, marriage, or adoption.
[2] Net undercount in 1980 Census: 2.0 percent males and 0 percent females. Revised estimates are from: Evaluation and Research Reports, PHC80-E., pp. 17.
[3] These data are available upon request. Write to: Alfredo Navarro
Bureau of the Census
Washington, DC 20233

REFERENCES

[1] LEMAITRE, G. and DUFOUR, J. (1986). Comparative Study of Weighting Methods for Family Data, 1986.
[2] HETTMANSPERGER, Thomas P., Statistical Inference Based on Ranks, 1984.
[3] ALEXANDER, C.H. and ROEBUCK, M.J., Comparison of Alternative Methods for Household Estimation, 1986.
[4] ZIESCHONG, KIMBERLY D., Sample Weighting Methods and Estimation of Totals in the Consumer Expenditure Survey, 1989.
[5] HANSEN, HURWITZ and MADOW, Sample Survey Methods and Theory Vol. I, 2nd Edition, pp. 59-63.