# Further Simulation Results on the Distribution of Some Survey Variance Estimators

Elizabeth T. Huang and William R. Bell*
Statistical Research Division, U.S. Census Bureau, Washington, D.C. 20233

**Abstract**
A relatively new aspect to small area estimation with area level models involves modeling direct survey variances for small areas to improve them. Here, as one aspect of this, we consider the distribution of some survey variance estimators – linearization, Fay's successive difference replication variance estimator, the jackknife, and the random group. We use simulations to examine whether the variance estimators might be assumed to approximately follow a scaled chi-squared distribution, and if so, with what value of the degrees of freedom? We do this for variances of estimated proportions from simple random samples of various sizes, with data generated from various distributions (Poisson, and Bernoulli), and with an artificial population constructed from American Community Survey data. This builds on previous work where we considered Fay's successive difference replication variance estimator for means computed from simple random samples.

**Key Words:** sampling variance, degrees of freedom, chi-squared distribution, sample size

## 1. Introduction

Small area estimation with area level models requires variance estimates of the sampling errors in the direct survey point estimates being modeled. Direct sampling variance estimates for small areas are likely to be unstable due to the small sample sizes, which suggest modeling the variance estimates to improve them. Statistical models to improve sampling variance estimates have been investigated by Otto and Bell (1995); Huff, Eltinge, and Gershunskaya (2002); Cho et al. (2002); Eltinge, Cho, and Hinrichs (2002); Gershunskaya and Lahiri (2005); Maples, Bell, and Huang (2009); and Maples (2010). Most of these papers assumed the direct sampling variance estimates were unbiased and followed scaled chi-squared distributions with known or estimated degrees of freedom. Huff, Eltinge, and Gershunskaya (2002) also considered a lognormal distribution, and Gershunskaya and Lahiri (2005) modeled log variance estimates focusing on the first two moments without explicitly specifying a distribution. Also, Otto and Bell (1995) and Eltinge, Cho, and Hinrichs (2002) modeled sampling covariance matrices, and so used the Wishart distribution, the multivariate generalization of the chi-square. Other papers that developed models assuming a chi-squared distribution for sampling variance estimates, though not with the explicit aim of improving the variance estimates, include Arora and Lahiri (1997) and You and Chapman (2006).

*Disclaimer: This report is released to inform interested parties of ongoing research and to encourage discussion. The views expressed on statistical, methodological, technical, or operational issues are those of the authors and not necessarily those of the U.S. Census Bureau.

The chi-squared (or Wishart) assumption is particularly important to the models of Otto and Bell (1995); Maples, Bell, and Huang (2009); and Maples (2010). Their models incorporated random small area effects for the variances, assumed distributed as inverse Gamma, and these distributional assumptions lead to an empirical Bayes smoothing of the direct variance estimates. The resulting shrinkage estimates are weighted averages of the direct variance estimates and fitted generalized variance function (GVF) values, with the weights depending on the degrees of freedom of the chi-squared distribution and the inverse Gamma parameter that determines the precision of the random variance effects. (Gershunskaya and Lahiri (2005) obtained an empirical Bayes smoothing of log variance estimates without distributional assumptions, which required direct estimation of the variances of the log variance estimates.)

The importance of the chi-squared distributional assumption to sampling variance modeling naturally leads to the questions of (*i*) under what circumstances does the chi-squared distribution provide a reasonable approximation, and (*ii*) when it does, what are the degrees of freedom, and how do they vary with sample size and other characteristics of the survey? Huang and Bell (2009) used simulations to address these questions in a particular case, namely, the application of Fay's successive difference replication variance estimator (Fay and Train 1995) to estimating variances of means of simple random samples. There we used simulations to examine the distributions of the variance estimates of means from simple random samples of various sizes from various distributions (normal, Poisson, and Bernoulli). In this paper, we extend this research to examine the distributions of several different variance estimators applied to estimates of ratios (proportions). The survey variance estimators examined here are the Taylor series linearization approach, Fay's successive difference replication variance estimator, the jackknife variance estimator, and the random group variance estimator (using 10 groups).

Section 2 of this paper reviews the four variance estimators of proportions used in the simulation study. Section 3 presents results obtained with simple random sampling from populations simulated using Bernoulli and Poisson distributions, and from an artificial population constructed by pooling 2005 American Community Survey (ACS) data over several counties from the state of Maryland. We examine bias of the variance estimators, how well their distribution is approximated by a chi-squared distribution, and the degrees of freedom of the chi-squared approximation. Results are obtained for a range of sample sizes from $n$ = 2 to $n$ = 760. Section 4 offers some conclusions.

## 2. Variance Estimators of Proportions under Simple Random Sampling

We assume a finite population of size $N$ units. The population parameter $\theta$ to be estimated is a proportion expressed as the ratio of two population totals $Y$ and $X$:

$$\theta = \frac{Y}{X} = \frac{\sum_1^N y_i}{\sum_1^N x_i}.$$

The sample estimate of $\theta$ from a simple random sample of size $n$ is $\hat{\theta} = \sum_{i=1}^n y_i / \sum_{i=1}^n x_i$, where, from now on, we let $i$ = 1,…, $n$ denote the units in the sample. In our simulations, $x_i$

and $y_i$ will be either indicator or count variables for unit *i*. In particular, with the ACS data, $x_i$ is the number of individuals who are children of school age (5-17) in household *i*, and $y_i$ is the number of individuals being both of school age and in poverty in household *i*. Thus, in this application, $\theta$ becomes the population proportion of school-age children in poverty (the 5-17 poverty ratio), and $\hat{\theta}$ the sample based estimate of the 5-17 poverty ratio.

The four variance estimates we consider, as applied to estimate the variance of a proportion, are as follows. We let $f = n/N$ denote the sampling fraction.

1) Taylor series linearization (Cochran (1977, pp. 31-32 and p. 155), compare to Wolter (1985, p. 236)):

$$v_L = \frac{1-f}{n\bar{x}^2(n-1)} \sum_{i=1}^{n} (y_i - \hat{\theta}\, x_i)^2$$

2) Fay's successive difference replication variance estimator (Fay and Train 1995):

$$v_F = 4R^{-1}(1-f) \sum_{r=1}^{R} \left(\hat{\theta}_r - \hat{\theta}_0\right)^2$$

   where *R*, the number of replicates, is a multiple of 4, and

$$\hat{\theta}_r = \textstyle\sum_{i=1}^{n} f_{ir} w_i\, y_i / \sum_{i=1}^{n} f_{ir} w_i\, x_i \qquad r = 1, \dots, R$$

$$\hat{\theta}_0 = \textstyle\sum_{i=1}^{n} w_i\, y_i / \sum_{i=1}^{n} w_i\, x_i$$

$$f_{ir} = 1 + (2)^{-3/2} a_{i+1,r} - (2)^{-3/2} a_{i+2,r} \qquad i = 1, \dots, n.$$

   In the above expressions, $w_i$ is the survey weight for sampled unit $i = 1, \dots, n$. For simple random sampling as considered here, $w_i = N/n$. Also, $A = [a_{ij}]$ is an $R \times R$ Hadamard matrix (Wolter 1985, pp. 320-352), which has elements $a_{ij} = \pm 1$, and has the property that $AA' = R \times I_R$. In the ACS, the number of replicates used is $R = 80$, and this is the value of $R$ we shall use here. Note above that row 1 of the matrix A (which is all ones) is not used. In application to ACS, the 41$^{st}$ row is also not used, and we shall proceed this way here as well. For sample sizes $n > R - 2$, we must re-use rows of *A* in defining the $f_{ir}$ for $i > R - 2$, a point discussed by Huang and Bell (2009).

3) The jackknife variance estimator (Wolter 1985, pp. 172-173):

$$v_J = \frac{1}{k(k-1)} \sum_{\alpha=1}^{k} \left(\hat{\theta}_\alpha - \bar{\hat{\theta}}\right)^2$$

   where the sample has been divided into *k* groups each of size *m*, with $n = mk$, and the "pseudo-value," $\hat{\theta}_\alpha$, is defined as

$$\hat{\theta}_\alpha = k\hat{\theta} - (k-1)\hat{\theta}_{(\alpha)}$$

where $\hat{\theta}_{(\alpha)}$ is the estimated ratio obtained by omitting the $\alpha^{th}$ group from the sample, and $\bar{\bar{\theta}} = k^{-1}\sum_{j=1}^{k}\hat{\theta}_\alpha$. (Wolter also mentions a more conservative version of $v_J$ obtained by replacing $\bar{\bar{\theta}}$ by the full sample estimate $\hat{\theta}$.) We shall use the "delete one" version of the jackknife, for which $m$ = 1 and $k$ = $n$. We also use the following modification to $\hat{\theta}_{(\alpha)}$ suggested by Wolter (1985, p. 173)

$$\hat{\theta}^*_{(\alpha)} = \hat{\theta} + (1-f)^{1/2}\left(\hat{\theta}_{(\alpha)} - \hat{\theta}\right).$$

Wolter suggested this modification for cases where the sampling fraction $f$ is not negligible, as occurs for the largest sample sizes we consider.

4) The random group variance estimator (Wolter 1985, pp. 23–33):

$$v_{RG} = \frac{1}{k(k-1)}\sum_{j=1}^{k}\left(\tilde{\theta}_j - \bar{\bar{\theta}}\right)^2$$

where the sample is divided into $k$ groups each of size $m$, $\tilde{\theta}_j$ is the estimate of $\theta$ formed from the data in group $j$, and $\bar{\bar{\theta}} = k^{-1}\sum_{j=1}^{k}\tilde{\theta}_j$. The form of $v_{RG}$ is obviously similar to that of $v_J$, the difference being that for $v_{RG}$ the groups from which $\tilde{\theta}_j$ are computed are disjoint. (As with $v_J$, there is a more conservative version of $v_{RG}$ which replaces $\bar{\bar{\theta}}$ with the full sample estimate $\hat{\theta}$.) Also, as with $v_J$, we use the following modification of $\tilde{\theta}_j$:

$$\tilde{\theta}^*_j = \hat{\theta} + (1-f)^{1/2}\left(\tilde{\theta}_j - \hat{\theta}\right).$$

In the simulation study we use random groups of size 10. To avoid complications, we compute $v_{RG}$ only for sample sizes $n$ which are integer multiples of 10.

Previous study of sampling variance estimators such as the above has focused mostly on their bias, and sometimes their variance or mean squared error, rather than on their distribution. Regarding the bias of $v_L$, Cochran (1977, p. 162) states that "With small samples, say $n$ < 30 and $C_{xx}$ $(= Var(y_i)/\bar{Y}^2)$ large, it has long been suspected that the large-sample formulas given for the variance of a ratio and its estimate are underestimates." Regarding $v_J$, Cochran (1977, p. 179) states that the jackknife variance estimate of a ratio becomes unbiased either for fixed $k$ or for $k$ = $n$ as $n$ becomes large. Wolter (1985, pp. 156-160) noted that, "The jackknife variance estimator was correct asymptotically. In finite samples, however, it tends to incur an upward bias of order $1/k^2$. But for linear functionals, the jackknife variance estimator is unbiased." We will see both these properties confirmed later in the simulation results.

## 3. Simulation Study of Variance Estimators of Estimated Proportions from Various Populations: Bernoulli, Poisson, ACS Data

To study the properties of the four variance estimators presented in Section 2, we created several artificial populations from which we could repeatedly draw samples. The populations included simulated data with $N$ = 10,000 observations from Bernoulli and (conditional) Poisson distributions. We also constructed a population by pooling ACS data from several counties in Maryland. From each of these populations we drew a large number (10,000) of simple random samples without replacement (srs wor) for each of a set of sample sizes *n*, computed the four variance estimates of proportions from each sample, and examined properties of the variance estimators for a given sample size over the simulations.

We restricted our analyses to simulated samples for which the denominator of the estimated proportion was positive. The various sample sizes *n* used in the simulations varied for the four variance estimators. For $v_L$ and $v_F$, we used 56 sample sizes ranging from *n* = 2 to *n* = 760. For $v_J$, we used 32 sample sizes ranging from *n* = 2 to *n* = 610. For $v_{RG}$, we used 27 sample sizes ranging from *n* = 2 to *n* = 610 in multiples of 10, so we would have equal numbers of sampled units in each of the 10 random groups.

Let $v$ denote any of the four variance estimators. We focus our analysis of the simulation results on the following three properties of the variance estimators, computed as summary statistics of $v$ across the simulations for each sample size used.

1) Percent relative bias of $v$, defined as

$$\text{Percent RelBias } (v) = 100 \times \frac{E(v) - Var(\hat{\theta})}{Var(\hat{\theta})}$$

where the true variance, $Var(\hat{\theta})$, is estimated by $\widehat{Var}(\hat{\theta}) = \frac{1}{K-1}\sum_{\ell=1}^{K}\left(\hat{\theta}_\ell - \bar{\hat{\theta}}\right)^2$, $\hat{\theta}_\ell$ is the estimated proportion from simulated sample $\ell$, $\bar{\hat{\theta}}$ is the mean across the simulations of the $\hat{\theta}_\ell$, $E(v)$ is estimated by $\bar{v} = K^{-1}\sum_{\ell=1}^{K} v_\ell$ where $v_\ell$ is the variance estimate from simulated sample $\ell$, and $K$ is the number of simulated samples drawn for which $\sum_{i=1}^{n} x_{li} > 0$ (the denominator of $\hat{\theta}_\ell$).

2) The "degrees of freedom," *d*, of $v$, is approximated by the Satterthwaite approximation (Ames and Webster 1991) via

$$d = \frac{2}{RelVar(v)}$$

where $RelVar(v) = Var(v)/[E(v)]^2$, the relative variance of $v$, is estimated by $\widehat{Var}(v)/\bar{v}^2$, where $\widehat{Var}(v) = \frac{1}{K-1}\sum_{\ell=1}^{K}(v_\ell - \bar{v})^2$ is the variance of the individual sample variance estimates $v_\ell$ over the simulations. If the distribution of $v$ is actually proportional to a chi-squared distribution, then *d* as given above is indeed its degrees of freedom. When $v$ is not exactly chi-squared, then $d$ as given above can be used as the degrees of freedom for a chi-squared approximation. Even if this approximation is not good, $d$ can still be thought of as a measure of precision of $v$ (given its connection to $RelVar(v)$ and its square root, the coefficient of variation of $v$.)

3) To measure how closely the distribution of $d \times v/E(v)$ matches a $\chi_d^2$ distribution, we use an unnormalized version of the Komogorov-Smirnov (K-S) statistic (Rao 1973, p. 421):

$$\text{K-S} = \sup_v |G(v) - F(v)|$$

where $G(v)$ is the empirical cumulative distribution function (c.d.f) of $d \times v/E(v)$ over the simulations, and $F(v)$ is the corresponding c.d.f. of the $\chi_d^2$ distribution. We use K-S merely to get an approximate measure of the difference between the true distribution (approximated by $G(v)$ from the large number of simulations) and the approximating $\chi_d^2$ distribution. We use .10 as a rough criterion value for K-S to indicate when the $\chi_d^2$ approximation seems reasonable. If K-S is substantially less than .10 we would be very satisfied with the $\chi_d^2$ approximation. When K-S is substantially more than .10 we would regard the $\chi_d^2$ approximation as inadequate. Values of K-S near .10 are marginal.

## 3.1. Simulation results from two Bernoulli populations

We simulated an artificial population of size $N$ = 10,000 from independent Bernoulli distributions by defining the variables $x_i$ , $z_i$ for $i$ = 1,…, $N$ as follows:

$$x_i = \begin{cases} 1 & \text{with probability } p_1 \\ 0 & \text{with probability } 1 - p_1 \end{cases}$$

$$z_i = \begin{cases} 1 & \text{with probability } p_2 \\ 0 & \text{with probability } 1 - p_2. \end{cases}$$

We use the values $p_1$ = 0.3 and $p_2$ = 0.08, 0.25. Letting $y_i = x_i \times z_i$, the finite population proportion $\theta$ will be approximately equal to $p_2$. We shall thus identify the two cases as $\theta$ = 0.08 and $\theta$ = 0.25 (which is at least approximately true).

We summarize the simulation results by plotting the three evaluation statistics (K-S, degrees of freedom, and percent relative bias) versus the sample sizes in Figure 1. The three graphs in the left column in Figure 1 are for the population with $\theta$ = 0.08; the three graphs in the right column are for $\theta$ = 0.25. Each graph contains four curves corresponding to the four variance estimates. The solid black line is for the Taylor series linearization variances ($v_L$), the red dashed line is for Fay's replication variance ($v_F$), the blue dot-dash line is for the jackknife variance ($v_J$), and the green dashed line is for the random group variance ($v_{RG}$). We summarize the results as follows:

The K-S values (the top two graphs in Figure 1) show that the chi-squared distribution is not a very good approximation for "small" sample sizes for any of the variance estimators. For $\theta$ = 0.08, we need a sample size of about 100 or more for the K-S values to be less than or equal to 0.1 for all four variance estimators. For $\theta$ = 0.25, we need sample sizes of about 40, 30, 50 and 90 for $v_L$, $v_F$, $v_J$, and $v_{RG}$, respectively, for the K-S values to be less than or equal to 0.1. Apart from peaks in the K-S values in all cases for small sample sizes, the K-S values mostly decline with increasing *n*, showing that the chi-squared approximation improves as sample size increases. We also note some sections on both graphs where the random group variance has larger K-S values reflecting that, for these sample sizes, the chi-squared approximation is not as good for $v_{RG}$ as it is for the other three variance estimators. The graphs also give some indication that, when *n* is large enough for the chi-squared approximation to be fairly good, it is slightly better for Fay's variance estimator than for the other three.

The graphs of the degrees of freedom (middle two graphs of Figure 1) show generally increasing values reflecting better relative precision (lower CVs) of the variance estimators as *n* increases. The lone exception is for $v_{RG}$ when $\theta$ = 0.25, for which the curve levels out at around 9 for *n* > 40. The leveling out at 9 is not surprising given that the form of $v_{RG}$ suggests its maximum degrees of freedom for any sample size would be $k-1$ (= 9 here). For $\theta$ = 0.08, the degrees of freedom for $v_{RG}$ does not seem to level out, but then it does not reach 9 even up to *n* = 760. For $v_L$ and $v_J$, the degrees of freedom are virtually identical, and they increase linearly with *n*, though at a slower rate for $\theta$ = 0.08 than for $\theta$ = 0.25. (Note that even for $\theta$ = 0.25, the degrees of freedom remain substantially less than $n-1$, the value that would obtain for estimates of means with normal data.)

For $v_F$ the degrees of freedom are less than for $v_L$ and $v_J$, though more than for $v_{RG}$. Also, we see some curvature, with a suggestion that the degrees of freedom of $v_F$ may be leveling out somewhere past *n* = 760. For comparison, Huang and Bell (2009) noted that, when using $v_F$ to estimate variances of sample means with normal data, the degrees of freedom increased linearly (roughly according to 2*n*/3) up to about *n* = 78, then increased more slowly until reaching a maximum of about 74. This behavior was generally expected, given some analyses done of the form of $v_F$, and since the use of 80 replicates suggested a maximum degrees of freedom of slightly less than 80 (given that we drop two rows of the Hadamard matrix). For estimating variances of means of nonnormal data, the degrees of freedom of $v_F$ were lower.

The graphs of the percent relative biases of the variance estimators (bottom two graphs in Figure 1) show large negative biases for the smallest sample sizes for all four variance estimators. For $v_L$ and $v_F$, the biases (which are similar) rapidly diminish as *n* increases, and appear effectively negligible for *n* > 50 or so. For $v_J$, there are substantial positive biases for a narrow range of relatively small sample sizes, after which the biases decline to negligibility, which is perhaps reached slightly later than is the case for $v_L$ and $v_F$. For $v_{RG}$ there is a similar pattern of large negative biases for the smallest *n* transitioning to substantial positive biases for moderate *n*, followed by a decline of the biases to negligibility as *n* further increases. This pattern is much delayed, however, compared to the similar pattern for $v_J$.

### 3.2 Simulation results from a (conditional) Poisson population

The (conditional) Poisson population is generated from two variables $z_i$ and $x_i$ according to the following distributions (which are independent over $i$):

$$z_i = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{cases}$$

$$x_i \sim \begin{cases} \text{Poisson}(\mu_1) & \text{when } z_i = 1 \\ \text{Poisson}(\mu_0) & \text{when } z_i = 0 \end{cases}$$

For the simulations we set $p = 0.08$, $\mu_1 = 0.49$, and $\mu_0 = 0.44$. Letting $y_i = x_i \times z_i$, the finite population proportion $\theta$ will be

$$\theta = \frac{\sum_1^N y_i}{\sum_1^N x_i} \approx \frac{\mu_1 p}{\mu_1 p + \mu_0 (1-p)} \approx .088.$$

This set-up was motivated as a plausible model for poverty data for which $z_i$ would be an indicator variable of poverty status for the $i^{\text{th}}$ household (with households here having poverty proportion $p = 0.08$), while $x_i$ represents the number of school-age children in the household (which follows a Poisson distribution whose mean is allowed to be different for households in poverty versus those not in poverty). While the resulting population proportion (poverty ratio) $\theta$ is similar to that for the first Bernoulli case, the distribution of the data here will be different.

Results from the simulations using the above model are plotted in the left column of Figure 2. (The right column of Figure 2 contains results obtained with the artificial population constructed from ACS data discussed in the next subsection.) The arrangement and labeling of the graphs is the same as that for Figure 1. We summarize the results shown in the left column of Figure 2, making comparisons to the left column graphs in Figure 1, as follows.

The K-S graph from the Poisson population is very similar to the K-S graph from the Bernoulli(0.08) population shown in Figure 1. There are large K-S values for all four variance estimators for small sample sizes, but the K-S values decline with increasing $n$. The K-S values are less than 0.10 for $n > 80$ for all four variance estimators. This is somewhat better than for the Bernoulli(0.08) case of Figure 1, which might be explained by the slightly larger value of .088 for $\theta$.

The patterns of the degrees of freedom plots in the left columns of Figures 1 and 2 are very similar, with the relations between the four variance estimators the same in both plots. However, the degrees of freedom for the conditional Poisson case are much lower than for the Bernoulli(0.08) case, except that for $v_{RG}$ the degrees of freedom are about the same between the two populations. However, $v_{RG}$ has the lowest degrees of freedom by far in both populations.

The bias plot in the left column of Figure 2 is very similar to that for the Bernoulli(0.08) case of Figure 1, except that the positive bias of $v_{RG}$, over the range of sample sizes where this is noticeable, is not quite as severe as in the Bernoulli(0.08) case.

### 3.3 Simulation results for an artificial population constructed from ACS 2005 poverty data

We pooled ACS 2005 sample data from Maryland's five largest county equivalents (Anne Arundel County, Baltimore County, Montgomery County, Prince George's County, and Baltimore City) to define an artificial population. This involved data from 19,264 households. The population proportion of interest ($\theta$) is the poverty ratio of age 5-17 related children. We drew 10,000 simple random samples of households from this artificial population with the various sample sizes (number of households) mentioned in Section 2. For each household we have $x_i$, the number of children in the household (related to the household head), and an indicator $z_i$ of whether the household was in poverty. Then, $y_i = x_i \times z_i$ equals zero if the household is not in poverty, and equals $x_i$ if it is in poverty. We could thus calculate the population poverty ratio $\theta$ for this artificial population (which was 0.085), and calculate the estimate $\hat{\theta}$ for each sample, as well as the four estimates of $Var(\hat{\theta})$. The three evaluation statistics for this case are plotted in the second column of Figure 2, which can be compared to both the first column of Figure 2 and the second column of Figure 1. We summarize the results as follows.

The K-S plot for this population is fairly similar to that for the conditional Poisson population, and hence also to that for the Bernoulli(0.08) population. One main difference is that the K-S statistics for the ACS data are a little worse, overall, than the others, in that they require somewhat larger sample sizes (about $n \geq 120$) until they decline to .10 and lower. Additionally, the K-S values for $v_{RG}$ are more similar to those of the other variance estimators than was the case for the conditional Poisson and Bernoulli(0.08) populations.

The degrees of freedom for $v_L$, $v_F$, and $v_J$ increase more slowly with increasing $n$ than is the case for the conditional Poisson population, which we noted showed slower increases in degrees of freedom than for the Bernoulli(0.08) population. With the ACS data, as with the other two populations, the degrees of freedom of $v_L$ and $v_J$ are quite similar, while they are lower for $v_F$. (This difference is less, however, with the ACS data than it is with the other two populations.) For $v_{RG}$, the rate of increase in the degrees of freedom is similar for all three of these populations, though its degrees of freedom are also easily the lowest among the four variance estimators for all three populations (never exceeding 6 with the ACS data).

The pattern of the bias plots with the ACS data is broadly similar to that of the bias plots of the conditional Poisson and Bernoulli(0.08) populations, but with some differences worth noting. The most noticeable difference is that, with the ACS data, the negative bias of $v_{RG}$ is more severe and persists to larger sample sizes, not approaching zero until near $n = 200$. Also, there is no "overshoot" in the bias of $v_{RG}$, that is, it does not increase into positive values for $n > 200$, but instead stays reasonably close to zero. The positive bias of $v_J$, however, persists to larger sample sizes (about to $n$ = 100) than was the case with the conditional Poisson and Bernoulli(0.08) populations. Finally, the negative biases of $v_L$ and $v_F$ diminish towards zero more slowly than was the case for the other two populations, with small negative biases persisting even for $n$ > 200.

## 4. Conclusions

In general, for sample sizes $n$ > 100, the simulation results are mostly supportive of using a scaled chi-squared approximation to the distributions of the four variance estimators of estimated proportions (with some cautions about extrapolating the conclusions very far outside the range of the simulations.) For sample sizes much smaller than 100, it appears that the chi-squared approximation may not be very good. The chi-squared approximation tended to fare better for larger values of the population proportion within the range considered here. Notice, though, that here we considered only values of the population proportion less than 0.5. For values greater than 0.5, the results might reverse and show poorer approximation by the chi-squared distribution as the proportion approaches 1.0.

The degrees of freedom, $d$, differed among the four variance estimators and across the various populations considered. Thus, $d$ generally increased with sample size $n$, but at different rates across the various populations and among the different variance estimators. For a fixed sample size and a given variance estimator, $d$ tended to be larger for populations with a larger population proportion (though this could be reversed for population proportions exceeding 0.5). For the artificial populations considered here, the degrees of freedom of the linearization and jackknife variance estimators were close and increased the fastest with the sample size $n$. The degrees of freedom of Fay's successive difference replicate variance estimator increased more slowly with $n$ and always remained below the number of replicates (80) being used. The degrees of freedom of the random group variance estimator increased the slowest with $n$ and did not exceed nine (the number of random groups used minus 1).

Biases of three of the four variance estimates of ratios were generally small for sample size $n$ > 100, the exception being the random group variance estimator, whose bias remained non-negligible for even larger values of $n$. For very small sample sizes, all four variance estimates have substantial negative bias. The bias patterns of the linearization variance estimator and Fay's variance estimator were similar, while there are positive biases for the jackknife variance for small, but not the smallest, $n$.

Obvious possibilities for future research would include extending the results to consider other variance estimators, other population distributions or other population parameters, and other sampling schemes.

## References


Ames, Michael H. and Webster, John T. (1991), "On Estimating Approximate Degrees of Freedom," *The American Statistician*, 45, 45-50.

Arora, Vipin and Lahiri, Partha (1997), "On the Superiority of the Bayesian Method Over the BLUP in Small Area Estimation Problems," *Statistica Sinica*, 7, 1053-1063.

Cho, Moon, Eltinge, John, Gershunskaya, Julie and Huff, Larry (2002), "Evaluation of Generalized Variance Function Estimators for the U.S. Current Employment Survey," *Proceedings of the American Statistical Association, Survey Research Methods Section*, Alexandria, VA: American Statistical Association, 534-539.

Cochran, W. G. (1977) *Sampling Techniques, 3rd ed.*, New York: John Wiley and Sons.

Eltinge, John, Cho, Moon, and Hinrichs, Paul (2002), "Use of Generalized Variance Functions in Multivariate Analysis," *Proceedings of the American Statistical Association*, *Survey Research Methods Section*, Alexandria, VA: American Statistical Association, 904-912.

Gershunskaya, Julie B. and Lahiri, Partha (2005), "Variance Estimation for Domains in the U.S. Current Employment Statistics Program," *Proceedings of the American Statistical Association*, *Survey Research Methods Section*, 3044-3051.

Fay, Robert E. and Train, George F. (1995), "Aspects of Survey and Model-Based Postcensal Estimation of Income and Poverty Characteristics for State and Counties," *Proceedings of the American Statistical Association, Government Statistics Section*, 154-159.

Huang, Elizabeth T., and Bell, William R. (2009), "A Simulation Study of the Distribution of Fay's Successive Difference Replication Variance Estimator," *Proceedings of the American Statistical Association, Survey Research Methods Section*, [CD-ROM], Alexandria, VA: American Statistical Association.

Huff, Larry, Eltinge, John and Gershunskaya, Julie (2002), "Exploratory Analysis of Generalized Variance Function Models for the U.S. Current Employment Survey," *Proceedings of the American Statistical Association, Survey Research Methods Section*, Alexandria, VA: American Statistical Association, 1519-1524.

Maples, Jerry J. (2010), "Using Small Area Modeling to Improve Design-Based Estimates of Variance for County Level Poverty Rates in the American Community Survey," Research Report, Statistical Research Division, U.S. Census Bureau, forthcoming.

Maples, Jerry J., Bell, William R., and Huang, Elizabeth T. (2009), " Small Area Variance Modeling with Application to County Poverty Estimates from the American Community Survey," *Proceedings of the American Statistical Association, Survey Research Methods Section*, [CD-ROM], Alexandria, VA: American Statistical Association.

Otto, Mark C., and Bell, William R. (1995), "Sampling Error Modelling of Poverty and Income Statistics for States," *Proceedings of the American Statistical Association, Government Statistics Section*, pp. 160-165.

Rao, C. R. (1973), *Linear Statistical Inference and its Applications, 2nd ed.*, New York: Wiley.

You, Yong and Chapman, Beatrice (2006), "Small Area Estimation Using Area Level Models and Estimated Sampling Variances," *Survey Methodology*, 32, 97-103.

Wolter, Kirk M. (1985), *Introduction to Variance Estimation*, New York: Springer-Verlag.

Figure 1. The K−S, D.F.,and Bias of 4 Variance Estimates of Estimated Proportions−Bernoulli

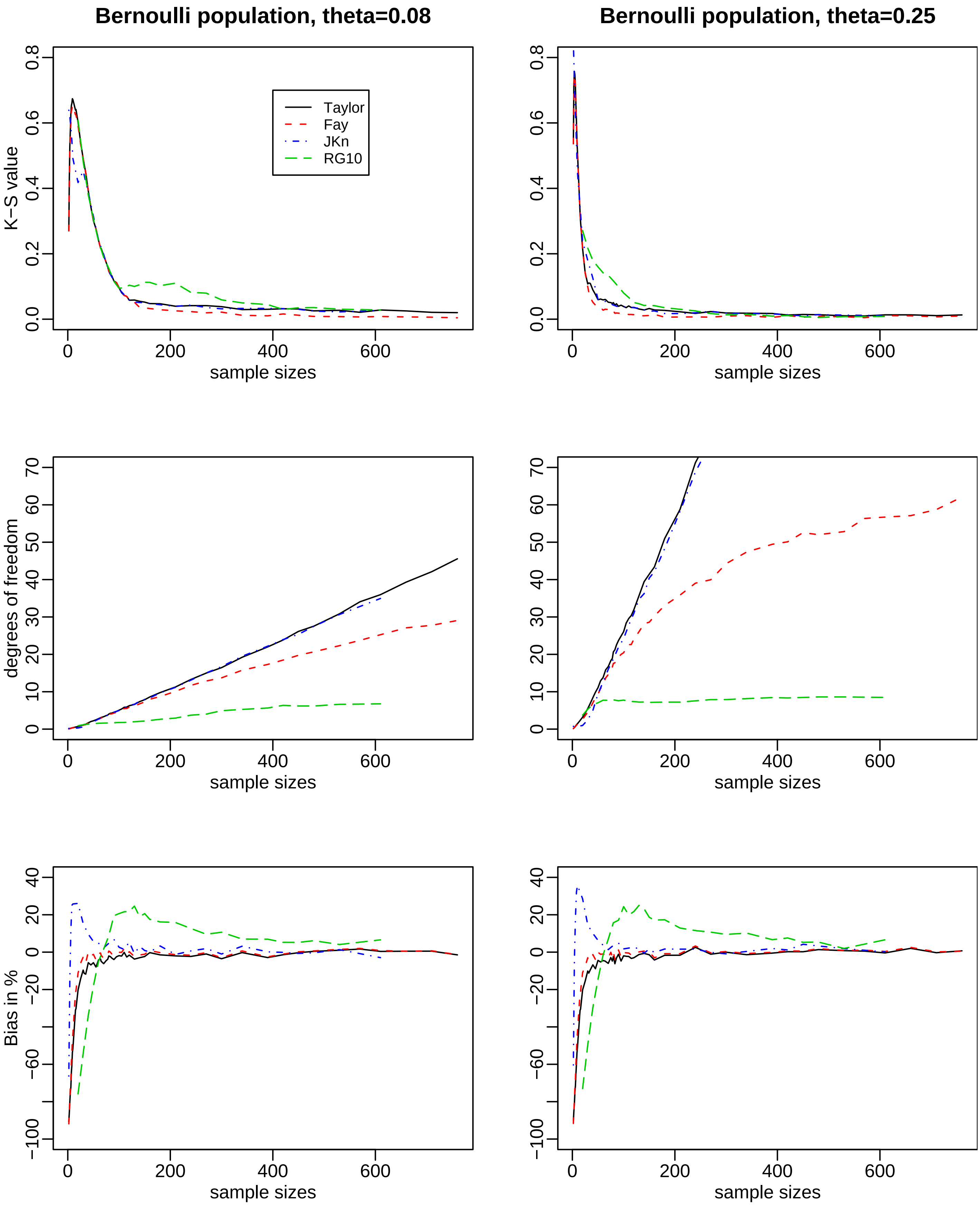

Figure 2. The K−S, D.F.,and Bias of 4 Variance Estimates of Estimated Proportions