Comparing Strategies To Estimate a Measure of Heteroscedasticity

Kimberly Henry, Internal Revenue Service, and
Richard Valliant, University of Michigan

Estimating totals is often a survey sampling objective. With a model-based approach, one factor that can affect the variance and bias of estimated totals is the superpopulation structure. We consider cases where a dependent variable's variance is proportional to some power of the independent variable. Various strategies that are conceivable in this case include: (1) selection of a pilot sample to make preliminary structural parameter estimates, (2) selection of a main sample based on either pilot results or educated guesses about population parameters, and (3) use of either a model-based or design-based estimator of the total. For various sample designs, sizes, and estimators, alternative strategies for estimating values of that variance power are compared for simulated population data. The strategies' effects on estimates of totals and their variances are then evaluated.

This paper is organized into six sections. After the introduction, the second section contains descriptions of our superpopulation model and generated populations. The third section includes our simulation setup details, while results are discussed in the fourth section. Conclusions, limitations, and future considerations are in the fifth section and references in the sixth section.

▸ Superpopulation Model and Generated Populations

Model Theory

Given a study variable of interest Y and an auxiliary variable X, we consider a superpopulation with the following structure:

$$\begin{aligned} E_M(y_i \mid x_i) &= \beta_0 + \beta_1 x_i \\ Var_M(y_i \mid x_i) &= \sigma^2 x_i^{\gamma} \end{aligned} \qquad (2.1)$$

The x_i's are assumed to be known for each unit i in the finite population. The exponent γ in model (2.1)'s conditional variance has been referred to as a *measure of heteroscedasticity* (Foreman, 1995), or *coefficient of heteroscedasticity* (Brewer, 2002). This parameter is of interest since a reasonable γ estimate produces nearly optimal sample designs and estimators of totals and their variances (Theorem 4.2.1, Valliant, Dorfman, and Royall, 2000).

Applications of models like (2.1) include companies using cost segregation to report depreciable assets on their Internal Revenue Service Tax Form 1120 (e.g., Allen and Foster, 2005 and Strobel, 2002) and comparing inventory data values versus actual values (e.g., Roshwalb, 1987 and Godfrey et al., 1984).

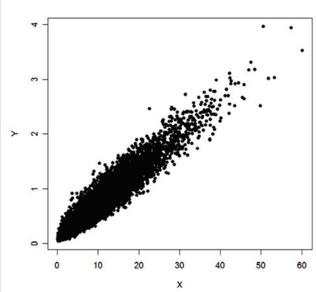
Given generated population data, our goal is to use various strategies to draw samples and estimate γ from them, then examine the impact of these strategies on the estimation of totals and their variances.

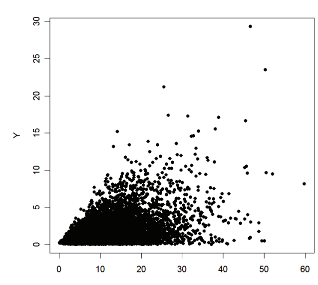
Generated Populations

We created two unstratified versions of the population described in Hansen et al. (1983, denoted HMT hereafter), since it follows model (2.1). We chose γ equal to 3/4 and 2 for populations of 10,000 units. Figures 1 and 2 show the population X, Y for each generated population (note a difference in Y-scales):

Figure 1—Generated Populations

$\gamma = 3/4$ $\qquad\qquad$ $\gamma = 2$





The first population has a relatively strong dependence between y and x, while the second one has a much weaker relationship. Note that these populations have a small non-zero intercept, which resulted in some model-based estimators being biased in the earlier HMT study.

▸ Simulation Setup

This section describes the details of our simulation study, including working models, sample designs, simulation strategies, and the method of estimating γ .

Models

Using Valliant et al.'s (2000) notation, we based estimators of totals on the following two working models:

$$M(1,1:x^{\gamma}) \quad (3.1)$$

$$M(x^{\gamma/2},x^{\gamma}:x^{\gamma}) \quad (3.2)$$

Model (3.1) is the correct working model, i.e., the one equivalent to model (2.1). Model (3.2) is associated with the following superpopulation structure:

$$\begin{aligned} E_M(y_i \mid x_i) &= \beta_{1/2}x_i^{\gamma/2} + \beta_1 x_i^{\gamma} \\ Var_M(y_i \mid x_i) &= \sigma^2 x_i^{\gamma} \end{aligned} \quad (3.3)$$

Working model (3.3) is called the minimal model (Valliant et al., 2000, p. 100) associated with the above conditional variance. If (2.1) were unknown, but the intercept is small, working model (3.3) may be a reasonable starting place for determining a sample size.

When the variance of y_i is proportional to x_i^{γ} and $E_M(y_i \mid x_i)$ is a linear combination of auxiliaries, one of which is x_i^{γ}, two important optimality results hold: (1) The selection probabilities that minimize the anticipated variance of the general regression (GREG) estimator are proportional to x_i^{γ} (Särndal, Swensson, and Wretman, 1992, sec. 12.2); and (2) The optimal model-based sample will have a certain type of weighted balance that also depends on x_i^{γ} (Valliant et al., 2000, sec. 4.2.1). An optimal, weighted balanced sample can be approximated by a probability-proportional-to- x_i^{γ} sample, denoted *pp*($\sqrt{x^{\gamma}}$).

There is often a huge incentive to use optimal samples and estimators in the applications we consider due to high data collection costs. In a cost segregation study, for example, experts may be needed to assign capital goods to depreciation classes (e.g., 5, 7, 15, or 39-year). Assessments can be time-consuming and expensive; so, the smaller the sample size that yields desired precision, the better.

Sample Designs

For each unit i in the population, we consider four without replacement (*wor*) sample designs:

(1) *srswor*: simple random sampling.
(2) *ppswor*: the Hartley-Rao (1962) method with probabilities of selection proportional to a measure of size (MOS).
(3) *ppstrat*: strata are formed in the population by cumulating an MOS and forming strata with equal total size. An srswor of one unit is selected from each stratum.
(4) *wtd bal*: weighted balanced sampling. *Ppswor* samples using an MOS are selected that satisfy particular conditions on the population and sample moments of x_i .

For each of these designs, we drew 1,000 samples of 100 and 500 units. When the MOS is $\sqrt{x^{\hat{\gamma}}}$, the *ppstrat* design approximates optimal $pp(\sqrt{x^{\gamma}})$ selection and *wtd bal* $\sqrt{x^{\gamma}}$ sampling. It is similar to "deep stratification" (e.g., Bryant et al., 1960; Cochran, 1977, pp. 124-126; Sitter and Skinner, 1994), which is used in accounting applications (Batcher and Liu, 2002). More specific details on these designs are given in pages 66-67 of Valliant et al. (2000).

Strategies

The strategies we examined consisted of selecting a pilot study to get a preliminary estimate of γ followed by a main sample or only selecting a main sample. Both options were crossed with the possibility of rounding γ or not. Thus, our main comparisons concern four strategies:

A: draw a $pp(\sqrt{x})$ pilot of 50 units, estimate γ , and select a main sample using *pp* ($\sqrt{x^{\hat{\gamma}}}$), *ppstrat* ($\sqrt{x^{\hat{\gamma}}}$), and *wtd bal* ($\sqrt{x^{\hat{\gamma}}}$) samples.
B: draw *srswor*, *ppswor* ($\sqrt{x}$) , *ppstrat* ($\sqrt{x}$), and *wtd bal* ($\sqrt{x}$) main samples only and estimate γ in each.
C: strategy A, rounding $\hat{\gamma}$ to the nearest one-half.
D: strategy B, rounding $\hat{\gamma}$ to the nearest one-half.

By definition, there is no *srswor* used for strategies A and C. Also, B and D correspond to assuming $\gamma = 1$

for selecting the *ppswor*, *ppstrat*, and *wtd bal* samples, which does not match our population γ's, but will be a reasonable advance choice for sampling in many populations. We consider the rounding in C and D to see if reducing variability in the $\hat{\gamma}$'s leads to improved estimates of totals and variances.

Estimation of γ

To estimate γ, following Roshwalb (1987), we iteratively fit a given working model and regressed the log of the squared residuals on $\log(x)$ as follows:

$$\log(r_i^2) = \alpha + \gamma \log(x_i) ,$$

and repeated the process until $\hat{\gamma}$ stabilized.

For all strategies, if $\hat{\gamma} \le 0$, then it was forced to one, which corresponds to $pp(\sqrt{x})$ sampling. Rejected alternatives included forcing $\hat{\gamma} = 0$, implying homoscedasticity, or dropping these samples, both of which are unrealistic. Table 1 shows the number of these occurrences for the $\gamma = 3/4$ population (there were less than 5 cases for each strategy for the $\gamma = 2$ population). Also, for all strategies, if $\hat{\gamma} > 3$, then it was forced to equal three to avoid unreasonably large $\hat{\gamma}$'s. Table 2 contains the number of these occurrences for the $\hat{\gamma} = 2$ population (there were none of these cases for the $\gamma = 3/4$ population).

In Table 1, strategies A and B's numbers are the number of negative $\hat{\gamma}$'s. For C and D, the numbers include cases where small positive $\hat{\gamma}$'s were rounded down to zero. The numbers in parentheses are the number of negative $\hat{\gamma}$'s. The rounding used for C and D leads to fewer negative estimates than in A and B, but rounding does not offer overall improvement. Strategies B and D produced fewer negative $\hat{\gamma}$'s than A and C since B and D use 100 and 500 units, as opposed to pilot samples of size 50 in A and C. Also, depending on the strategy, there were at least three times as many negative $\hat{\gamma}$'s using model (3.2) versus using (3.1).

Table 1—Number of Times $\hat{\gamma} \to 1$, $\gamma = 3/4$ Population

Strategy	*Design*	$M(1,1:x^{\gamma})$		$M(x^{\gamma/2}, x^{\gamma} : x^{\gamma})$	
		pilot n=50		*pilot n=50*	
A	*ppswor*	52	67	159	171
	ppstrat	56	56	164	199
	wtd bal	60	59	167	181
C	*ppswor*	157 (18)	134 (28)	263 (98)	243 (122)
	ppstrat	129 (20)	150 (25)	256 (83)	275 (114)
	wtd bal	136 (24)	142 (24)	252 (63)	267 (105)
		n=100	*n=500*	*n=100*	*n=500*
B	*srswor*	8	0	68	3
	ppswor	16	0	93	5
	ppstrat	11	0	81	5
	wtd bal	12	0	92	3
D	*srswor*	43 (2)	0	158 (40)	30 (0)
	ppswor	67 (2)	0	179 (52)	43 (1)
	ppstrat	53 (2)	0	191 (50)	23 (0)
	wtd bal	59 (2)	0	184 (52)	34 (0)

In Table 2, Strategies B and D produced fewer large $\hat{\gamma}$'s than A and C. Rounding in C and D also produced fewer large $\hat{\gamma}$. There were at least twice as many large $\hat{\gamma}$'s when using model (3.1) versus model (3.2).

Table 2—Number of Times $\hat{\gamma} \to 3$, $\gamma = 2$ Population

Strategy	*Design*	$M(1,1:x^{\gamma})$		$M(x^{\gamma/2}, x^{\gamma} : x^{\gamma})$	
		pilot n=50		*pilot n=50*	
A	*ppswor*	73	73	21	21
	ppstrat	61	51	22	18
	wtd bal	81	63	28	24
C	*ppswor*	39	46	9	6
	ppstrat	32	36	8	8
	wtd bal	27	32	14	10
		n=100	*n=500*	*n=100*	*n=500*
B	*srswor*	7	0	2	0
	ppswor	7	0	2	0
	ppstrat	12	0	1	0
	wtd bal	5	0	3	0
D	*srswor*	2	0	0	0
	ppswor	2	0	1	0
	ppstrat	3	0	0	0
	wtd bal	2	0	1	0

Estimation of Totals

We consider three kinds of estimators for totals: the *Horvitz-Thompson* (HT) estimator, *best linear unbiased predictors* (BLUP), and *general regression estimators* (GREG). The HT estimator is given by

$$\hat{T}_{\pi} = \sum_{i \in s} y_i / \pi_i ,$$

where π_i is the probability of selection for unit *i*.

The general form of the BLUP estimator is

$$\hat{T} = \sum_{i \in s} y_i + \sum_{i \notin s} \mathbf{x}_i' \hat{\boldsymbol{\beta}} ,$$

where $\mathbf{x}_i' \hat{\boldsymbol{\beta}}$ is the prediction for y_i using the working model and set of units in the population that are not in the sample (denoted by $i \notin s$) and $\hat{\boldsymbol{\beta}}$ is estimated using the sample units ($i \in s$). For example, following Valliant et al.'s (2000) notation, the BLUP using the correct model is

$$\hat{T}(1,1:x^{\gamma}) = \sum_{i \in s} y_i + \sum_{i \notin s} \mathbf{x}_i' \hat{\boldsymbol{\beta}},$$

where $\hat{\boldsymbol{\beta}} = (\mathbf{X_s}'\mathbf{V_{ss}^{-1}}\mathbf{X_s})^{-1}\mathbf{X_s}'\mathbf{V_{ss}^{-1}}\mathbf{y_s}$, $\mathbf{X_s}$ is an nx2 matrix with rows $(1, x_i)$, $\mathbf{V_{ss}} = diag(x_i^{\gamma})$, and $\mathbf{y_s}$ is the n-vector of sample data.

The general form of the GREG estimator is

$$\hat{T}_{GR} = \sum_{i \in s} g_i y_i ,$$

where g_i is the "g-weight" for unit i (Särndal et al., 1992).

These estimators combined with the two working models and true value of γ and estimates of γ lead to nine totals. For model (3.1), we have $\hat{T}(1,1:x^{\gamma})$, and $\hat{T}(1,1:x^{\hat{\gamma}})$, $\hat{T}_{GR}(1,1:x^{\gamma})$, and $\hat{T}_{GR}(1,1:x^{\hat{\gamma}})$. The estimators $\hat{T}(x^{\gamma/2}, x^{\gamma}:x^{\gamma})$, $\hat{T}(x^{\hat{\gamma}/2}, x^{\hat{\gamma}}:x^{\hat{\gamma}})$, $\hat{T}_{GR}(x^{\gamma/2}, x^{\gamma}:x^{\gamma})$, and are $\hat{T}_{GR}(x^{\hat{\gamma}/2}, x^{\hat{\gamma}}:x^{\hat{\gamma}})$ for model (3.2). $\hat{T}_{\pi}$ is the ninth. Note that the true γ is not available in any real situation; estimators computed using γ serve as a comparison standard for the other choices.

Variance Estimation

For the HT estimator, the variance estimator is:

$$\mathrm{var}_0(\hat{T}_{\pi}) = \left(1 - n/N\right)\frac{n}{n-1}\sum_{i \in s}\left(y_i/\pi_i - 1/n\sum_{i \in s} y_i/\pi_i\right)^2 .$$

This variance expression assumes with replacement sampling, but uses the finite population correction adjustment $1 - n/N$ to approximately account for *wor* sampling. Since the sampling fractions are small, the bias is negligible (Wolter, 1985, sec. 2.4.5).

The following is the *basic model variance* estimate for the BLUP estimators:

$$\mathrm{var}_1(\hat{T}) = \sum_{i \in s} a_i^2 r_i^2 + \left(\sum_{i \notin s} x_i^{\gamma}\right)\left(\sum_{i \in s} x_i^{\gamma}\right)^{-1} \sum_{i \in s} r_i^2 ,$$

where a_i is the "model weight" involving x_i in the working model and r_i is the residual for unit i.

We also include a robust *leverage-adjusted variance* estimate for the BLUP's:

$$\mathrm{var}_2(\hat{T}) = \sum_{i \in s} \frac{a_i^2 r_i^2}{1 - h_{ii}} + \left(\sum_{i \notin s} x_i^{\gamma}\right)\left(\sum_{i \in s} x_i^{\gamma}\right)^{-1} \sum_{i \in s} r_i^2 ,$$

where h_{ii} is the leverage for unit i. The identical second term in both model variances accounts for variability in population units not in the sample.

For the GREG's, we include the following variance estimators (e.g., see Valliant, 2002, expression 2.4):

$$\mathrm{var}_3(\hat{T}_{GR}) = \left(1 - \frac{n}{N}\right)\sum_{i \in s} \frac{g_i^2 r_i^2}{\pi_i^2}$$

$$\mathrm{var}_4(\hat{T}_{GR}) = \left(1 - \frac{n}{N}\right)\sum_{i \in s} \frac{g_i^2 r_i^2}{\pi_i^2 (1 - h_{ii})} .$$

The same variances were used for all sample designs, except for the *ppstrat* design-based variances for the HT and GREG estimators, where successive pairs of sample units were grouped, variances were calculated within each stratum, and strata variances were cumulated. Since both working models were specified over all strata, the model variance formulae var_1 and var_2 were used for samples selected using *ppstrat* sampling in estimating the variance of the BLUP.

▶ Simulation Results

Estimates

We calculated the average $\hat{\gamma}$ over each set of 1,000 samples drawn from both populations. Results are only summarized here.

When $\gamma = 3/4$, strategies B and D had more nearly unbiased estimates than A and C due to the smaller pilot sample sizes in the latter two. The rounding in strategies C and D made the average $\hat{\gamma}$'s further from the true value, since $\hat{\gamma}$'s close to three-fourths were either rounded down to one-half or up to one.

When $\gamma = 2$, the average $\hat{\gamma}$'s were closer to the true values. There was not much difference between the average $\hat{\gamma}$'s for the pilot study strategies A and C versus the no-pilot strategies B and D. The rounding also did not make much of a difference. Using the correct model (3.1) rather than (3.2) resulted in $\hat{\gamma}$'s closer to the true value, as might be expected.

Total and Variance Estimates

Our primary focus is how estimating γ effects estimates of totals and their variances. Tables 5 and 6 at the end of this paper include the root mean square error (RMSE) and 95-percent confidence interval (CI) coverage of each of the nine total estimators based on samples of size 100 drawn from the $\gamma = 3/4$ and $\gamma = 2$ populations, respectively (similar generalizations held for the samples of size 500, which are omitted due to length). Both tables are organized such that the HT estimates are first, followed by the BLUP and GREG totals produced using the true γ value (which resulted in identical results for strategies B and D), then those that used $\hat{\gamma}$'s. Relative biases (Relbias) are not shown in the tables but are briefly mentioned below.

For the $\gamma = 3/4$ population, where the true total is 7,174.74, all estimators were approximately unbiased over the 1,000 samples since the largest Relbias value was -0.41 percent for $\hat{T}(x^{\hat{\gamma}/2}, x^{\hat{\gamma}} : x^{\hat{\gamma}})$ using strategy B and *wtd bal* ($\sqrt{x}$) samples. For all strategies, using the correct working model (2.1) versus model (3.2) resulted in lower Relbias and RMSE values and CI coverage closer to 95 percent, though differences are not drastic. With model (3.2), using the GREG estimator resulted in improvements in all three measures over the equivalent BLUP estimators. Comparing strategies, there are slight improvements in the Relbias, RMSE, and CI coverage of strategy B over A and D over C, so that using the small pilot studies does not lead to any improvements. While the rounding of the pilot $\hat{\gamma}$'s in C offers improvements in the measures over A's, that is not the case in strategies B and D. For the sample designs, results from the *ppstrat* samples seem to be most favorable. For these populations, *wtd bal* sampling based on $\sqrt{x}$ in the main sample for Strategy B is suboptimal since the variance of neither population is proportional to *x*. Nonetheless, *ppstrat* ($\sqrt{x}$) is still reasonably efficient. As expected in these types of populations, the RMSE's when sampling by *srswor* are uniformly worse than those for the other designs in strategies B and D.

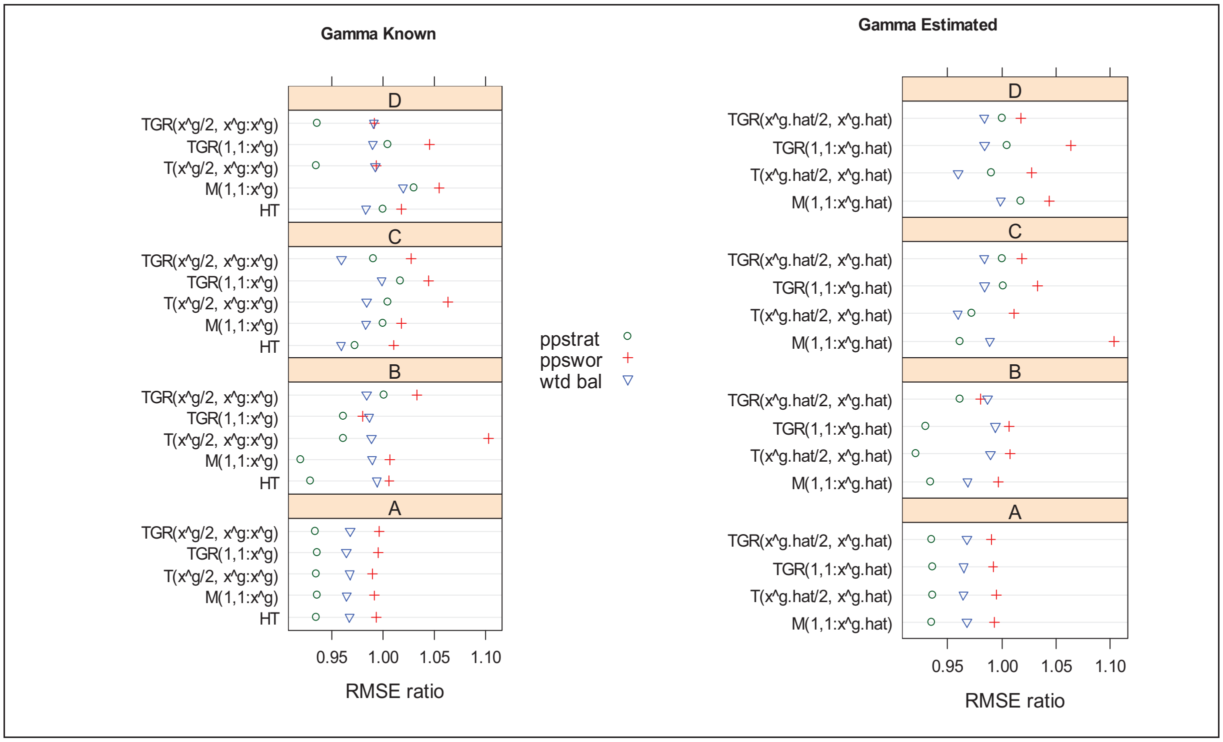
For the $\gamma = 2$ population, which had a total of 14,304.74, the largest Relbias value was 1.29 percent. Again, using the correct working model led to improved results, in terms of lower Relbias and RMSE values and CI coverage closer to 95 percent; there are slight gains in using the GREG estimator with model (3.2). Here, there is a notable (but not drastic) drop in the overall CI coverages compared to the $\gamma = 3/4$ population, the lowest being 91.7 percent. The most striking difference in RMSE values are the gains achieved with the pilot strategies over the corresponding nonpilot ones. For example, the RMSE for the combination ($\hat{T}_{GR}(1,1 : x^{\hat{\gamma}})$, A, *ppstrat*) is 1,186.76, while the RMSE for ($\hat{T}_{GR}(1,1 : x^{\hat{\gamma}})$, B, ppstrat) is 1,289.02. That is, using a pilot leads to an RMSE that is about 92.1 percent of that of using no pilot.

Figure 2 on the following page displays the ratios for the $\gamma = 2$ population of RMSE's of the various estimators and sampling plans to the RMSE of the combination of $\hat{T}_{GR}(1,1 : x^{\hat{\gamma}})$, B, *ppstrat*, with estimated γ for *n* = 100. This combination was selected as the reference since (a) *ppstrat* is a popular plan in practice, and (b) the GREG estimator $\hat{T}_{GR}(1,1 : x^{\hat{\gamma}})$ is one that is used by conservative practitioners because it is approximately design-unbiased while still taking advantage of the *y*-*x* relationship. The left and right panels show the ratios for estimators that use the true γ and an estimated γ. When the true gamma is used in estimation, but a pilot study is conducted to determine how to select the main sample, the most efficient method of sampling is *ppstrat*. In the (*ppstrat*, pilot) case, all estimators have about the same RMSE.

The right-hand panel gives the more realistic comparisons among combinations that could be used in practice. Conducting a pilot study with strategy A (no rounding) followed by a *ppstrat* ($\sqrt{x^{\hat{\gamma}}}$) main sample yielded a 4- to 8-percent reduction in RMSE compared to the reference combination described above. Rounding in strategy C reduces the gains from doing a pilot. Weighted balance on an estimated γ has no advantage over the reference combination.

If no pilot is conducted (strategies B and D), then *wtd bal* ($\sqrt{x}$) is the most efficient scheme, but *ppstrat* ($\sqrt{x}$) is very competitive. The rounding in strategy D leads to virtually the same results as B. Among the estimators, the model-based choice $\hat{T}(x^{\hat{\gamma}/2}, x^{\hat{\gamma}} : x^{\hat{\gamma}})$ and the GREG $\hat{T}_{GR}(x^{\hat{\gamma}/2}, x^{\hat{\gamma}} : x^{\hat{\gamma}})$ are somewhat worse than the others, although differences are not extreme.

Figure 2—Ratios of RMSE's for estimators and sampling plans to the RMSE for $\hat{T}_{GR}(1,1:x^{\hat{\gamma}})$, B, ppstrat, with estimated γ, $\gamma = 2$ population, n=100.



In all cases, unrestricted *ppswor* sampling was the poorest performer, regardless of whether γ was known or estimated.

▶ General Conclusions, Limitations, and Future Considerations

We investigated some alternative strategies for sampling and estimation in populations where there is one target variable *y*, whose total is to be estimated, and one auxiliary *x*, which is known for every unit in the population. The variance of y is known to increase as *x* increases, but the exact form of the variance is unknown to the sampler. Modeling the variance as $Var_M(y_i \mid x_i) = \sigma^2 x_i^{\gamma}$ is assumed to be a good approximation to reality. We studied three options that might be considered for this type of problem: design of a pilot sample, design of a main sample, and selection of an estimator.

We obtained ambiguous results on whether a pilot study, designed to get a preliminary estimate of γ, would be worthwhile. For our versions of the HMT population, the smaller pilot studies gave more negative $\hat{\gamma}$'s and more biased ones on average. In the less variable population we studied, conducting a pilot did not consistently give lower root mean square errors for the totals than using only a main sample with an educated guess about the size of γ. Rounding $\hat{\gamma}$'s, to the nearest half was not particularly helpful or harmful in estimating totals. Small root mean square error improvements came from reducing the variability in the $\hat{\gamma}$'s, in strategies C and D, for the less variable population ($\gamma = 3/4$), but the opposite was true in the more variable population ($\gamma = 2$). Thus, when the focus is on estimating γ, a pilot study and rounding are not useful. But, if the focus is on estimating totals, a pilot, possibly with rounding, may offer slight MSE improvements, depending on the population variability.

Among the sampling plans we considered, stratification based on cumulative $\sqrt{x^{\hat{\gamma}}}$ or $\sqrt{x}$ rules, denoted *ppstrat* here, were both reasonably efficient. The use of *wtd bal* samples based on $\hat{\gamma}$'s was not effective in reducing the root mean square errors of totals.

Table 3—Root Mean Square Error and 95-Percent Confidence Interval Coverage using Design-Based (D), Basic Model (B), and Leverage-Adjusted Model (L) Variances, $\gamma = 3/4$, n=100 for All Strategies

Strategy	**A**			**B**				**C**			**D**			
Design	*ppswor*	*ppstrat*	*wtd bal*	*srswor*	*ppswor*	*ppstrat*	*wtd bal*	*ppswor*	*ppstrat*	*wtd bal*	*srswor*	*ppswor*	*ppstrat*	*wtd bal*
$\hat{T}_\pi$														
RMSE	259.02	138.93	134.86	471.34	207.98	139.56	138.40	259.02	138.93	138.44	471.34	207.98	139.56	138.40
95% CI - D	94.9	94.4	100.0	94.0	95.4	94.5	99.8	95.1	94.4	100.0	94.0	95.4	94.5	99.8
$\hat{T}(1,1:x^\gamma)$														
RMSE	140.33	138.29	133.28	148.36	138.29	138.77	136.98	140.33	138.29	136.27	148.36	138.29	138.77	136.98
95% CI – B	94.1	94.4	94.8	94.1	94.8	94.4	94.6	94.1	94.4	94.6	94.1	94.8	94.4	94.6
95% CI – L	94.4	94.7	94.9	94.4	95.2	94.6	95.0	94.4	94.7	95.0	94.4	95.2	94.6	95.0
$\hat{T}_{GR}(1,1:x^\gamma)$														
RMSE	140.13	138.28	134.21	149.10	138.55	139.52	137.50	140.13	138.28	136.92	149.10	138.55	139.52	137.50
95% CI – D	94.1	94.3	94.8	94.4	94.5	94.3	94.7	94.1	94.3	94.5	94.4	94.5	94.3	94.7
95% CI – L	94.4	94.5	95.1	94.6	94.8	95.1	95.1	94.4	94.5	95.1	94.6	94.8	95.1	95.1
$\hat{T}(x^{\gamma/2},x^\gamma:x^\gamma)$														
RMSE	146.93	138.18	135.24	155.89	144.42	139.94	139.78	146.93	138.18	139.14	155.89	144.42	139.94	139.78
95% CI – B	94.1	95.3	95.4	94.0	94.8	95.1	95.2	94.1	95.3	95.5	94.0	94.8	95.1	95.2
95% CI – L	94.4	95.7	95.6	94.3	94.9	95.1	96.0	94.4	95.7	95.6	94.3	94.9	95.1	96.0
$\hat{T}_{GR}(x^{\gamma/2},x^\gamma:x^\gamma)$														
RMSE	147.85	138.22	135.15	155.59	144.24	139.49	139.05	147.85	138.22	139.02	155.59	144.24	139.49	139.05
95% CI – D	94.3	94.3	95.4	94.0	94.8	94.4	95.2	94.3	94.3	95.5	94.0	94.8	94.4	92.5
95% CI – L	94.4	94.6	95.6	94.3	95.5	94.9	95.3	94.4	94.6	95.6	94.3	95.5	94.9	95.3
$\hat{T}(1,1:x^{\hat{\gamma}})$														
RMSE	137.87	142.82	137.67	149.61	138.72	139.39	137.70	137.85	139.73	141.16	149.92	138.80	139.12	137.23
95% CI – B	94.3	94.2	94.1	93.8	95.0	94.5	94.8	95.3	94.6	94.9	93.8	94.8	94.6	94.5
95% CI – L	94.5	94.8	94.7	94.3	95.0	95.1	94.9	95.6	94.8	95.2	93.9	95.0	95.0	94.8
$\hat{T}_{GR}(1,1:x^{\hat{\gamma}})$														
RMSE	138.15	143.07	137.98	149.22	138.54	139.54	137.47	137.19	139.78	141.25	149.26	138.51	139.52	137.48
95% CI – D	94.5	94.5	94.9	94.4	94.5	94.4	94.8	95.4	94.9	94.9	94.5	94.5	94.4	94.9
95% CI – L	94.7	94.7	95.1	94.7	94.8	95.0	95.0	95.6	95.0	95.2	94.6	94.8	95.0	95.0
$\hat{T}(x^{\hat{\gamma}/2},x^{\hat{\gamma}}:x^{\hat{\gamma}})$														
RMSE	156.26	145.03	150.80	168.46	165.68	153.98	153.12	148.19	144.89	143.34	159.06	147.89	141.51	142.63
95% CI – B	94.9	95.9	95.3	94.0	93.7	95.3	95.0	94.6	95.7	95.0	94.2	94.3	95.2	95.0
95% CI – L	94.9	96.0	95.4	94.4	94.0	95.4	95.2	94.7	95.9	95.5	94.5	94.5	95.6	95.2
$\hat{T}_{GR}(x^{\hat{\gamma}/2},x^{\hat{\gamma}}:x^{\hat{\gamma}})$														
RMSE	158.14	143.42	150.75	168.24	154.78	141.81	144.09	153.69	141.72	143.14	158.61	145.58	139.32	140.97
95% CI – D	94.5	94.5	95.3	93.9	95.0	94.4	95.7	94.0	94.7	95.0	94.0	94.6	94.4	95.5
95% CI – L	94.7	94.8	95.4	94.3	95.3	95.0	95.8	94.4	95.2	95.6	94.6	94.8	94.9	95.7

Table 4—Root Mean Square Error and 95-Percent Confidence Interval Coverage using Design-Based (D), Basic Model (B), and Leverage-Adjusted Model (L) Variances, $\gamma = 2$, n=100 for All Strategies

Strategy	**A**			**B**				**C**			**D**			
Design	*ppswor*	*ppstrat*	*wtd bal*	*srswor*	*ppswor*	*ppstrat*	*wtd bal*	*ppswor*	*ppstrat*	*wtd bal*	*srswor*	*ppswor*	*ppstrat*	*wtd bal*
$\hat{T}_\pi$														
RMSE	1280.39	1205.96	1247.28	1693.63	1331.49	1291.02	1268.66	1280.39	1205.96	1279.58	1693.18	1331.49	1291.02	1268.66
95% CI - D	92.6	94.4	94.2	91.7	94.2	93.2	94.0	92.6	94.4	93.6	91.7	94.2	93.2	94.7
$\hat{T}(1,1:x^\gamma)$														
RMSE	1278.38	1206.63	1243.64	1334.49	1302.87	1253.61	1236.75	1278.38	1206.63	1277.71	1335.55	1302.87	1253.61	1236.75
95% CI – B	93.1	94.0	93.6	92.9	94.3	93.9	94.6	93.1	94.0	92.5	92.9	94.3	93.9	95.3
95% CI – L	93.5	94.1	94.0	93.2	94.3	94.0	94.7	93.5	94.1	93.2	93.2	94.3	94.0	95.4
$\hat{T}_{GR}(1,1:x^\gamma)$														
RMSE	1282.87	1206.52	1243.09	1482.90	1312.25	1289.11	1268.04	1282.87	1206.52	1277.18	1484.84	1312.25	1289.11	1268.04
95% CI – D	93.2	94.0	93.7	92.3	93.5	93.2	93.6	93.2	94.0	93.7	92.3	93.5	93.2	94.6
95% CI – L	93.5	94.2	93.9	92.6	93.7	93.3	93.8	93.5	94.2	93.9	92.6	93.7	93.3	94.6
$\hat{T}(x^{\gamma/2},x^\gamma:x^\gamma)$														
RMSE	1276.01	1205.81	1247.71	1594.75	1345.67	1311.31	1287.59	1276.01	1205.81	1279.80	1595.21	1345.67	1311.31	1287.59
95% CI – B	92.4	94.4	93.7	92.8	94.2	93.2	94.4	92.4	94.4	93.7	92.8	94.2	93.2	94.2
95% CI – L	92.8	94.5	93.9	93.3	94.6	94.2	94.6	92.8	94.5	93.9	93.3	94.6	94.2	94.2
$\hat{T}_{GR}(x^{\gamma/2},x^\gamma:x^\gamma)$														
RMSE	1284.85	1204.91	1248.23	1544.90	1370.79	1295.43	1268.39	1284.85	1204.91	1279.76	1545.43	1370.79	1295.43	1268.39
95% CI – D	92.6	94.3	93.7	93.1	94.1	93.0	93.5	92.6	94.3	92.7	93.1	94.1	93.0	94.7
95% CI – L	93.0	94.6	93.9	93.7	94.1	93.3	93.8	93.0	94.6	93.2	93.7	94.1	93.3	94.8
$\hat{T}(1,1:x^{\hat{\gamma}})$														
RMSE	1297.04	1198.12	1281.47	1358.54	1324.69	1276.81	1237.23	1267.80	1246.20	1252.81	1364.39	1327.10	1280.05	1236.12
95% CI – B	93.4	96.1	92.2	91.8	93.3	93.2	94.1	93.4	94.8	93.6	91.8	93.2	93.2	95.2
95% CI – L	93.3	96.3	92.4	91.8	93.5	93.2	94.2	93.7	95.1	93.8	92.0	93.5	93.3	95.3
$\hat{T}_{GR}(1,1:x^{\hat{\gamma}})$														
RMSE	1298.17	1186.76	1275.54	1482.33	1311.79	1289.02	1268.02	1272.01	1230.47	1239.79	1483.95	1311.59	1289.02	1267.99
95% CI – D	93.3	95.7	92.0	92.3	93.7	93.3	93.6	93.4	94.8	93.6	92.2	93.6	93.3	94.6
95% CI – L	93.4	96.1	92.3	92.5	93.8	93.3	93.7	93.6	95.3	93.8	92.5	93.6	93.3	94.6
$\hat{T}(x^{\hat{\gamma}/2},x^{\hat{\gamma}}:x^{\hat{\gamma}})$														
RMSE	1263.57	1239.37	1271.88	1536.34	1359.13	1327.38	1314.23	1293.69	1242.82	1250.31	1541.92	1366.02	1332.87	1315.43
95% CI – B	93.9	93.3	93.0	91.7	93.7	93.0	93.3	94.1	93.8	93.4	91.6	93.6	93.4	94.2
95% CI – L	94.4	93.8	93.5	92.0	93.9	93.5	93.9	94.4	93.9	93.4	91.9	94.1	93.8	94.5
$\hat{T}_{GR}(x^{\hat{\gamma}/2},x^{\hat{\gamma}}:x^{\hat{\gamma}})$														
RMSE	1421.62	1239.20	1274.71	1531.53	1347.67	1295.43	1276.47	1275.41	1243.26	1249.69	1536.62	1350.59	1298.09	1279.59
95% CI – D	93.7	93.4	93.5	91.8	93.5	93.0	93.1	94.1	93.9	93.5	91.7	93.8	93.0	94.5
95% CI – L	93.8	93.6	93.7	92.3	94.0	93.4	93.6	94.2	94.0	93.5	92.2	93.8	93.3	94.8