Modeling Superpopulation Variance: Its Relationship to Total Survey Error

James R. Knaub, Jr.
US Dept. of Energy, Energy Information Administration, EI-53, Washington, DC 20585

Keywords:
regression model; performance measure; scatterplot edits; inherent variance; nonsampling error; sampling error; total survey error

Abstract:
In repeated surveys there generally are auxiliary/regressor data available for all members of the population, that are related to data collected in a current sample or census survey. With regard to modeling, these regressor data can be used to edit the current data through scatterplots, and to impute for missing data through regression. Another use for regressor data may be the study of total survey error. To do this, follow these steps: (1) stratify data by regression model application (the related scatterplots can be used for editing); (2) find predicted values for data not collected, if any, and (3) replace all data that are collected with corresponding predicted values. If model-based ratio prediction is used, variance proportionate to a measure of 'size,' then the sum of the predicted values equals the sum of the observed values they replace. (See "Fun Facts" near the end of this article: fact # 4.) The standard error of the total of the predicted values for every member of a finite population, divided by that total, and expressed as a percent, could be labeled as an estimated relative standard error under a superpopulation, or a model-based RSESP. This RSESP would be influenced by (1) the models chosen, (2) inherent variance, and (3) total survey error (sampling and nonsampling error). This article proposes this model-based RSESP as a survey performance indicator and provides background and examples using both real and artificial data.

Introduction:
One suggestion that the author made to the Subcommittee on Measuring and Reporting the Quality of Survey Data, Federal Committee on Survey Methodology, in January 1998, was to use the variance of the prediction error (Maddala (1992)) as an indicator of data quality for a given number that would not be a useless measure, such as the relative standard error of the total, when dealing with a census. The square root of this statistic is available from commercial software, and provides an easily accessible performance measure that is impacted by three factors: 1) the models chosen, 2) inherent variance, and 3) data quality. Applying this across the total of all members of a population, as described in the abstract, results in the RSESP, also influenced by the same factors. Scatterplot graphical editing illustrates the same information that the RSESP captures overall, and that the standard error of the individual prediction error (*e.g.*, STDI in SAS PROC REG) captures for an individual case. As will be described, these measures can be used to pick the best model available. (Note the first factor influencing these performance indicators.) They can also indicate when data are not understood well. (See the second factor.) Controlling for these factors, one can then obtain an idea of data quality. Discussing this at the Energy Information Administration (EIA), John Vetter suggested that it is only a sudden increase in the RSESP, when applied repeatedly to a periodic survey, that indicates a data quality problem. This is consistent with the use of RSE estimates to determine when data edits may not have been successfully employed. For more than a decade, estimated RSEs and scatterplot edits have been used in the Office of Coal, Nuclear, Electric and Alternate Fuels (CNEAF) within the EIA to indicate and locate problems with data quality. This works well when sampling, but RSE estimates are zero for a census, even when data quality is very low, and thus total survey error is large. In a recent presentation to the American Statistical Association's Committee on Energy Statistics (Waugh, Norman and Knaub (2003)), uses for RSE estimates were discussed.

In the body of this article, examples of these methods will be presented. Most of these examples have been published previously.

RSESP: A proposed indicator of total survey error for each census and sample survey**:**
Inference in survey statistics may be made founded upon design-based sampling, model-based sampling, combined model-based and design-based considerations, or model-based imputation for either a census or any kind of sample. In all cases, variance is measured in some manner consistent with the collection methodology, and that variance is then applied in some manner to the data that are 'missing,' whether by 'design,' or through nonresponse. If there are no data 'missing,' then the estimated variance for the total is 0. Thus, relative standard error, RSE, is 0%. Therefore, nonsampling error would go unnoticed for a census, when relying only on the RSE. If, however, the variance measured from the data collected is applied to the entire population, as if the population observed is one manifestation of a superpopulation, then we may discuss a measure of variability for the finite population that may be called a relative standard error under a superpopulation, RSESP. (See Knaub (2002), pages 21-23 in the detailed ***InterStat*** version.)

Under design-based inference (see Korn and Graubard(1998)), we sometimes only need to remove the finite population correction (fpc) factor, which had only allowed the estimated variance to be applied to the unobserved portion of the population, when estimating variance of the total. Under model-based inference, we need to consider the 'predicted' value corresponding to each member of a population for the data element of interest. Rather than considering only the predicted values for the unobserved data, all observed data should also be replaced by predictions when estimating the variance of the total, when estimating the RSESP. (Note that the model-based ratio estimate, with variance proportional to the measure of size, rather than OLS estimation, has appeared to be fairly robust and has appeared to make good common sense.)

In any case, as we learn to group a population into fairly homogeneous strata, to reduce variability, we may calculate the lowest RSESPs possible. If the lowest RSESP possible is 'large,' say, over 10%, we may be concerned about total error, including nonsampling error. The nature of the regressor data would be important in the model-based cases. The ability to group data that are reasonably homogeneous is important. For the design-based cases, this could mean strata with minimum within-stratum variances. (See Korn and Graubard(1998) for complications.) For the model-based case, this means grouping data such that all data in each group are modeled well by a single regression model per group. The minimum estimated RSESP would indicate the best overall grouping scheme found. (For model-based inference of establishments, it may sometimes be better to consider the minimum RSESP subject to leaving the ‘smallest’ establishments unobserved. It is often difficult to collect quality data on a frequent basis from the smallest establishments.) The estimated min(RSESP) at least partially reflects the impact of nonsampling error, as well as sampling error and error due to regression model failure, and may be the best single indicator of data quality if one considers all of those factors. When doing so, all data used in survey estimations are to be included.

So, define RSESP as the relative standard error under a superpopulation, indicating variability across a population for a given data element.

It should be noted that the model-based form of the RSESP is very closely related to scatterplot editing. Model-based estimation here has been influenced by the work of KRW Brewer, RM Royall and others, but confined here to a strictly model-based approach, to study data quality. (Of particular fundamental usefulness to the reader may be Royall(1970) and Brewer(2002).)

The EIA has conducted a number of annual census surveys for a number of years. They may not have all been very successful, yet they may go years without being addressed, since it is common to assume that these surveys have captured "the answer," with regard to each response, with virtually 100 percent accuracy. Scatterplots show simple relationships between data sets that might instruct data collectors. In turn, this indicates regression, which can be used to impute for nonresponse, *e.g.*, in early census results. In turn, the variance found can be applied across the population to produce an RSESP to indicate overall credibility of survey results. So, for handling nonresponse on a sample survey of any kind, or on a census, the model-based form of the RSESP can be part of an overall system for improving and measuring data quality.

Indicator for Total Error:

Relative Standard Error Under a Superpopulation, RSESP

Summary Table:

Type of sampling and inference	***Standard error within a group***	***Variance based on***	***RSE: estimated standard error for total divided by total, written in %***	***RSESP: (std err, applied to all members of population - observed or not) - used as follows***
Design-based (Randomization)	Standard error (design may complicate matters – see Korn and Graubard(1998))	Mean square distances of points from means	RSE based on sample, but applied only to unobserved members of population	-To refine sample design? -As an indicator of total survey error? More investigation would be needed.
Model-based (conditionality) (Also good for imputation for nonrespondents in design-based sampling and in a census)	Standard errors calculated within 'estimation group' ('EG' - Knaub(1999))	Residual errors and model coefficient errors – Individual prediction errors and MSE used	RSE based on sample, but applied only to unobserved members of population	-To pick estimation groups and models -As an indicator of total survey error

So, what if we took the variability estimated from the observed data and applied it to the entire population? To reiterate: For a design-based sample, for some designs, that would mean removing the finite population correction factor, fpc, so that fpc=1. In other cases there are more complications. (See Korn, EL and Graubard, BI (1998).) For regression modeling, that would mean calculating variance of a total by using predicted values for all members of the population, replacing observed with imputed (predicted) values. Summations would then be over N rather than N-n. (See Knaub (1999) for formulas.)

When collecting data, note the following: *There is always a trade-off between data quality and sample size (data quantity): A larger sample size may have substantially more nonsampling error. The lowest obtainable variance, calculated for the entire population, would be an indicator of the usefulness of the data observed.*

Using the best method found, if the estimated RSESP is still large, this would indicate possibly large nonsampling error, as RSE has often done, but this would even apply to a census.

Examples will follow, showing uses for model-based RSESP estimation. Only the model-based RSESP will be considered here.

Below is an excerpted table from Knaub (2003) showing the use of the model-based RSESP to assist in comparing the merits of a model that accounts, to an extent, for possible changes in fuels used at an electric-generation plant, to a model that does not account for such "fuel switching," when collecting data on generation. Multiple regressors were used. "TOTO" here represents the total of other regressors, so TOTO=XF2+XF3+XF4. The graphs that immediately follow this table are for the same data. Artificial data were used in this case, which allows an isolated effect to be studied without being 'masked' by other factors.

RSESP* Estimates for Artificial Data

Here RSESP is used to pick the most appropriate model

RSE = 0 - All data are observed.

Fuel Switching Allowed

(artificial data illustrated in graphs to follow)

Regressors	Sum(e_i^2)	var due to coefficients	total variance	RSESP
XF1, XF2, XF3, XF4	2,152,776,960	1,968,930,468	4,121,707,428	2.09%
XF1 and TOTO	3,061,120,000	2,432,898,000	5,494,018,000	2.41%
XF1 only	17,430,527,000	14,742,935,000	32,173,462,000	5.84%
TOTO only	15,178,169,000	15,298,390,000	30,476,559,000	5.68%

RSE = 0 - All data are observed.

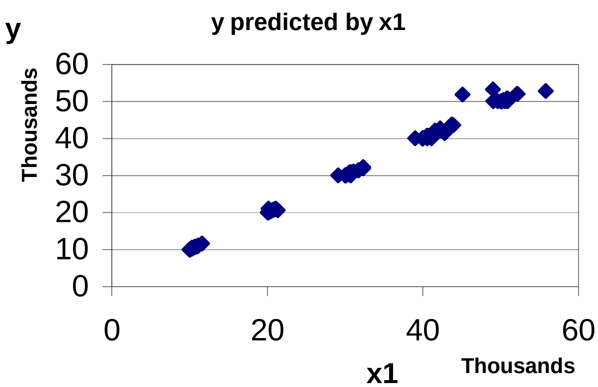
No Fuel Switching

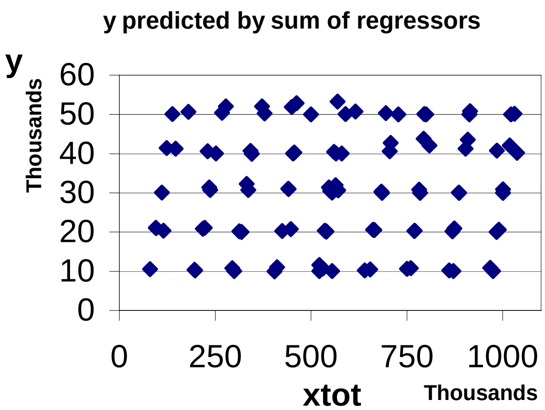
(artificial data illustrated in graphs to follow)

Regressors	Sum(e_i^2)	var due to coefficients	total variance	RSESP
XF1, XF2, XF3, XF4	83,476,608	112,104,923	196,581,531	0.46%
XF1 and TOTO	83,751,000	111,916,000	195,667,000	0.46%
XF1 only	82,695,000	56,826,000	139,521,000	0.38%
TOTO only	43,801,000,000	87,084,000,000	130,885,000,000	11.78%

*See "RSESP" in Knaub(2002) JSM CD & ***InterStat***. On page 4 of Knaub(1999) in ***InterStat***, the 'exact' variance estimate shown for totals is used here, except that summations above are over N rather than N-n.

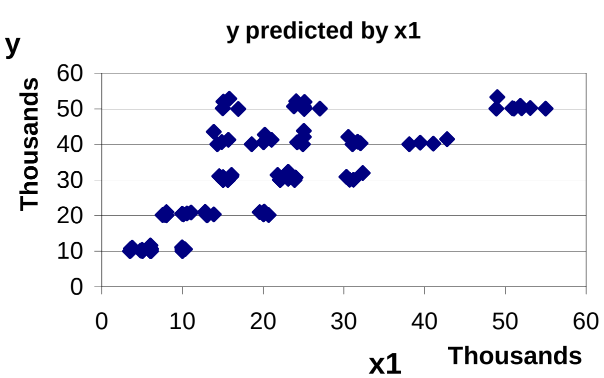
Artificial Test Data: No Fuel Switching Occurs

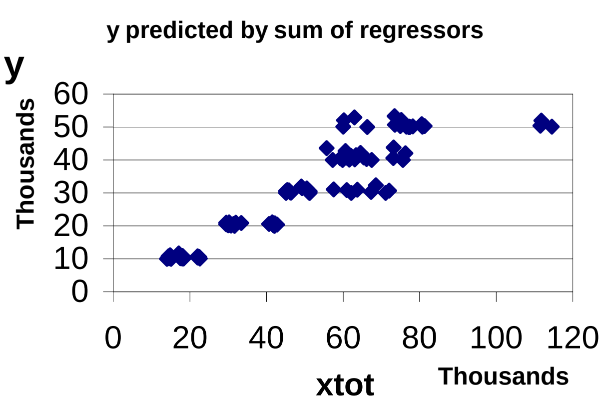




Note: $xtot = x_1 + x_2 + x_3 + x_4$

Artificial Test Data: Some Fuel Switching Occurs





Note: $xtot = x_1 + x_2 + x_3 + x_4$

Assuming one has appropriate models, the RSESP is an indicator of total survey error:

Hydroelectric generation data discussed in Knaub(1999) and Knaub(2002) are shown in a table and graphs to follow. In an example of imputation, annual census data were used for regressor data, and random sets of data were deleted from another annual census of hydroelectric generation data to simulate missing data among these real data. The imputed data compared well with the data they replaced; variance was low. These same data were used in the table below to compare relative standard errors (RSEs) with the RSESP indicators of total survey error.

Variance estimation here followed the approximation found in Knaub(1999), with δ=0.25, γ=0.5, and the regression weight $w=(x_1+2c)^{-2\gamma}$, where x_1 is the annual generation for a given respondent in the previous year's census, and c is the nameplate capacity. Here, x_1+2c is used as a measure of "size," in this case, a preliminary estimate of y.

Let "NCDC" represent the National Oceanic and Atmospheric Administration, National Climatic Data Center's U.S. Standard Regions for Temperature and Precipitation.

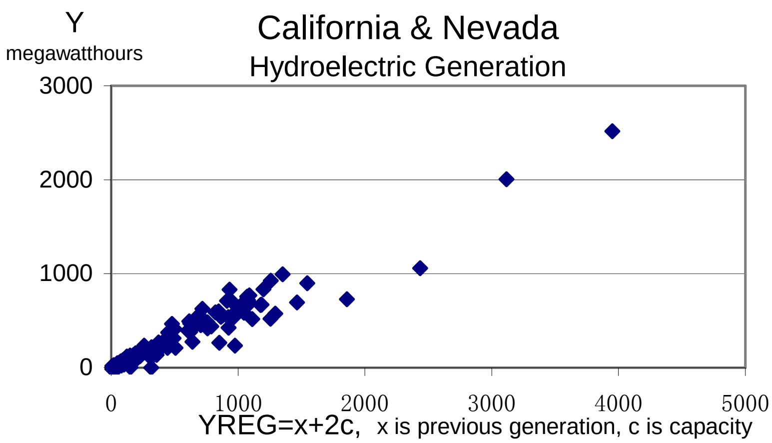
Imputation Examples

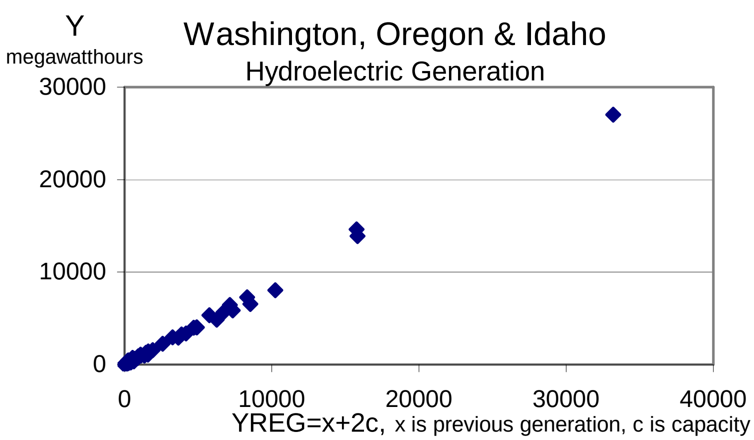
Region	Population Size, N	Num. Obs., n	RSE	RSESP
NCDC West (CA & NV)	233	206	0.7%	3.3%
NCDC Northwest (WA, OR & ID)	147	134	0.2%	2.4%
Pacific Contiguous Census Division	331	296	0.2%	1.9%
NCDC West (CA & NV)	233	233	0%	3.0%
NCDC Northwest (WA, OR & ID)	147	147	0%	2.4%
Pacific Contiguous Census Division	331	331	0%	1.8%

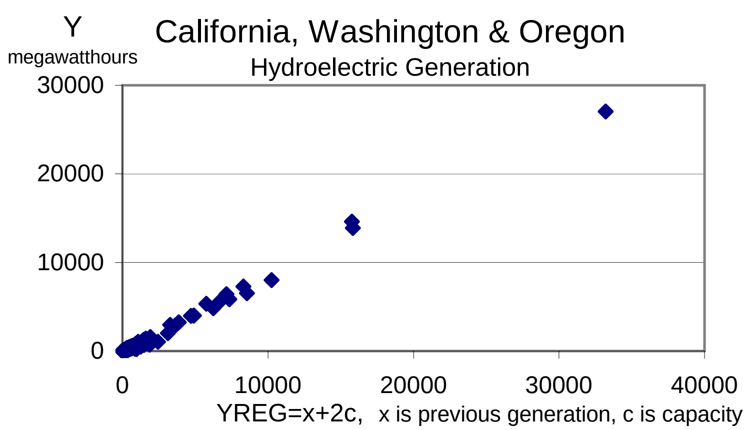
Although all data appear of reasonable data quality, note that the NCDC West results are perhaps the least reliable, but that would not be evident from the estimated RSE values when n = N. Even in extreme cases, those RSE estimates would still all tie, at zero.

Data for Imputation Examples from Hydroelectric Generation

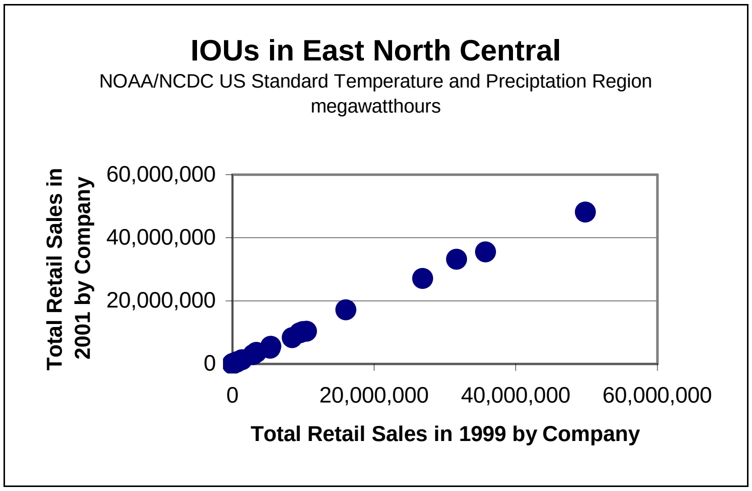
(for table above)



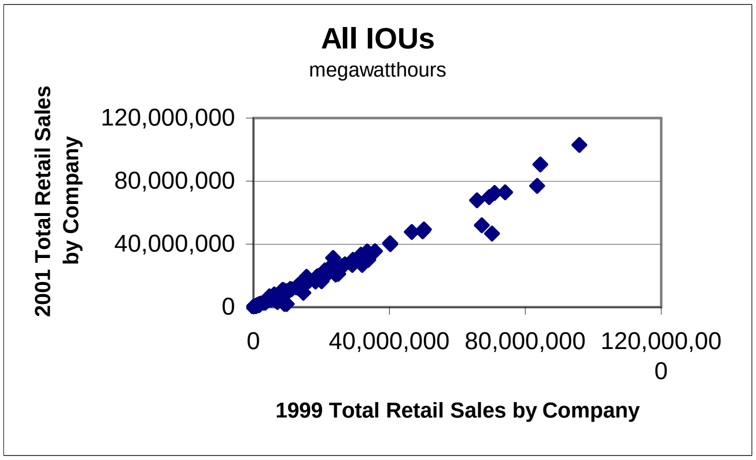




As another example, investor owned utilities (IOUs) are collected with certainty (a certainty stratum) on the EIA-826 form, a monthly sample of retail sales and revenue data. Because there is a census of IOU data, these data do not contribute to relative standard errors, and thus the IOU data appear to be collected with complete accuracy. However, estimating model-based RSESPs may yield some idea of the relative contribution to inaccuracy when compared to model-based RSESPs for data that are sampled from utilities of other ownership types. Further, such modeling will be useful in case of imputation for nonresponse, and associated scatterplot editing. The author has had an interest in the study of estimation groups for these data. Below are two cases: one where IOU total retail sales data from a specific geographic region, East North Central as defined by the National Climatic Data Center, are compared over two annual surveys, and a second where all US IOU data are treated in the same estimation group (not recommended but done for illustration). Scatterplots for those two cases are shown below, along with estimated RSESPs using the model-based ratio estimate the author has described as 'robust'. (Note that the estimated RSESPs here used 'exact' variance calculations, but results were not very different from those obtained using the methodology in Knaub (1999), and using delta = 0.3. "Exact" variance calculation has been an issue with the ASA Energy Committee.)



single regression coefficient: 1.007 with standard error 0.007, n = 24
estimated RSESP = 1.05%



single regression coefficient: 0.984 with standard error 0.010, n = 183
estimated RSESP = 1.45% - Note 2 of 183 data points seem particularly odd, but points near the origin may have a greater impact on estimated variance for weighted least square regression.

Estimation groups:
Estimation groups may be compared to see which ones should be combined or subdivided, based upon practical and physical considerations (*e.g*., geographical considerations based on weather), and upon test results estimating regression coefficients and their standard errors. In the case of multiple regression, for purposes of these tests, to make direct comparisons, a function of those regressors should be used as a single regressor so that a single coefficient will be the source of comparison. Attention should also be paid to total estimated RSESP values for differing estimation groups.

A note in closing these examples:
Finally, note that a preliminary investigation into cost and quality of fuels data for electric plants indicates that a much larger inherent variance component is to be expected for model-based (and probably design-based, if it were feasible) RSESP estimates for these data than for sales and revenue, or even generation and consumption of fuels data. Perhaps the situation will improve substantially if access to better regressor data becomes possible. More on this may be forthcoming. As for the cases above, it would appear that 'processes' are 'under control.'

Restating the Premise:
Standard errors are estimated for sample surveys, but they are only designed to estimate the error in estimating a total, when not all data are observed. That is, the question being addressed is "How much error can I expect there will be when estimating a given total, due to the fact that I did not collect all the data?" Variability in the collected data is projected upon the uncollected data. That variability is measured in design-based sampling, based upon the sampling design and the randomization principle. For regression model-based sampling, which may be thought of as 'mass imputation,' or when just using regression modeling to account for nonresponse, the variability is measured based on uncertainty in the estimated regression coefficients, and uncertainty in the residuals. In any case, variability found in the collected data is used to estimate variability in an estimate of the unobserved data. This means that nonsampling/measurement error is not directly considered. However, large nonsampling error will result in a large estimate of variability that will be projected upon the uncollected data.

What if we took the variability estimated from the observed data and applied it to the entire population? For some design-based sampling, that would mean removing the finite population correction factor, fpc, so that fpc=1. (This is often more complicated, however. See Korn and Graubard (1998).) For regression modeling, that would mean calculating variance of a total by using predicted values for all members of the population, replacing observed with imputed values. In regression modeling, the ratio estimate, variance proportionate to the measure of 'size,' which appears to be fairly robust, would result in predicted values whose subtotal for the units observed is equal to the total of the corresponding observed values. (See "Fun Fact" number 4 near the end of this article.) Summations in variance calculations would be over predicted data for all members of the population, N, not just over the unobserved units, N-n. Thus, in any case, for design-based inference or for model-based inference or imputation, we take the variability estimated in the data observed, and apply it to the entire population, as if the population was drawn from a superpopulation, and we are interested in all variability involved in this process. The model-based case relates to editing in that a scatterplot edit indicates suspicious data that contribute heavily to this variance. In the case of a census, a high variability here would indicate possible nonsampling error, which would similarly be indicated by a scatterplot edit.

In addition to nonsampling error, any strata or 'estimation groups' used would impact on this indicator. Various considerations may apply to design-based sampling (see Korn and Graubard(1998)). For model-based sampling and estimation, with regard to the establishment of estimation groups for which one model per group is applied, the ratio estimate, variance proportionate to the measure of 'size,' has good properties. The largest subsets of the population are found for a given data element or elements such that each subset is reasonably homogeneous. This, in general, means finding the lowest overall estimate of variance. Applying this to the entire population, as indicated above, if the smallest variance found for an estimated total is too large, there could be unacceptably large nonsampling error, or the sample size may not have been adequate, or else we do not know enough about how to adequately predict these numbers, so even publishing a census could be dangerous as we would not know if the data quality is very low. Further, there is always a trade-off between data quality and sample size (data quantity). The lowest obtainable variance, calculated as described above for the entire population, would be an indicator of the usefulness of the data.

When editing data, attention is drawn to data that deviate substantially from the norm. This would result in a large RSESP. A large RSE has often been used as an indicator of substantial nonsampling error in a sample, highlighting a few data points to be investigated. However, the RSESP can be used regardless of sample size, and even for a census. A scatterplot edit would indicate possibly substantial nonsampling error for the same members of the population whose data would likely substantially increase the numerical value of the estimated RSESP. Regression models are often used for imputation for nonresponse, and in the case of imputation, may even be used in combination with design-based sampling and inference, as suggested in Lee, Rancourt and Sarndal (1999). Thus it may be important to have a model-based form of the RSESP, even when dealing only with design-based sampling and estimation.

Philosophy of Total Survey Error:

Whether doing a design-based sample, a model-based sample, or a census, with or without nonresponse, there is always going to be nonsampling error, and that is hard to measure. Dobbs, Gibbins, Martin, Davies and Dodd(1998) state that "Non-sampling error can … have considerable impact on the accuracy of estimates, resulting in bias and/or an increase in the variance of estimates." They consider the usefulness of "special studies," and subjective descriptions of the whole range of data quality problems, rather than concentrate only on publishing the few quantifiable measures we have. So, even if one can estimate standard errors quite well, that does not mean that a census, with a total standard error thus defined as zero, would have no error. They say that "A much wider view of quality needs to be taken, with quality reporting including not just quantifiable information …" but also an examination of the processes involved. However, they also say "Users may also wish to compare information about the quality of different data sources so it would be desirable to develop more common or standardised methods of quality reporting." So, perhaps, like many others, they recognize the shortcomings of the 'quantifiable indicators,' such as standard error, that we have. However, might they also be inclined to see an advantage in having something more direct for comparison than a lengthy and subjective quality profile, although such a profile is the only means for an in-depth understanding of data quality for a given survey? In Federal Committee on Statistical Methodology(2001), there is a chapter on total survey error that also suggests the usefulness of such studies. Still, it would certainly appear to this author to be helpful to be able to pack as much information into one indicator as possible, for comparison purposes, or to use this as a performance measure to indicate whether survey results may be publication worthy.

Standard error is a very widely used measure, and one with which the public has some familiarity. One standard error can at least be described to laymen as the range, above and below a point estimate, within which we believe there are about two chances in three that the correct 'answer' resides. But what about bias? What about nonsampling error? Often in practice, variance is of more importance to overall accuracy than bias. If there is a substantial systematic bias, there may be a reason for this that could be virtually eliminated by better practice. (See Knaub (2001) for a study of bias.) So, we will concentrate on variance. As for nonsampling error, that will impact on variance estimates. In fact, a high variance for a data element for which a number is to be published may indicate the presence of substantial nonsampling error. At the Energy

Information Administration (EIA), large relative standard errors (RSEs) have often acted as an edit-indicator for serious data collection errors. However, *high nonsampling error due to low quality observed data results in large estimates of standard error that are applied only to the estimation/imputation of missing data.* Thus RSE estimates may not alert us to some substantial data collection problems.

As a measure of total survey error, like the design-based analog, the choice of strata, if used, and other mechanics will impact upon results for the model-based RSESP. Perhaps the minimum obtainable value for the RSESP is the most meaningful in terms of determining data quality. If after accommodating as much as possible for the natural division of a population into "estimation groups" (see Knaub(1999) and Knaub(2001)), there is still a large RSESP, this may be a much better indicator of what estimated (sub)totals should not be published. This may be applied to census surveys. It has long been apparent that electric power census survey data that were published for years with high but unrecognized nonsampling error were quickly identified when using a model to relate such data to currently collected data. By using the RSESP on a continuous basis for data collected periodically, like a Shewhart Chart for industrial quality control, one may be able to see quickly when a process is out of control. (Note the comment by John Vetter reported early in this article.) In the case of a census, using the model-based approach may be more practical than a design-based one, as software such as SAS PROC REG can be used to quickly adjust model parameters for each "estimation group" (*i.e.*, relatively homogeneous subdivision of the population). However, a design-based approach could be used as described above. A model-based approach requires correlated regressor data.

Examples of model-based applications are found in Knaub(2002) and Knaub(2003), as shown above. In the longer, ***InterStat*** version of Knaub(2002), found at http://interstat.stat.vt.edu/InterStat/index/JUL02.html, on pages 21 – 23, there is a discussion of the RSESP, and on page 23 some 'typical' examples show RSESP values of 2 to 3 percent. The shorter version of Knaub(2002) for the JSM 2002 CD had less than a page on this topic, with no numerical examples. Knaub(2003) has a page of results, comparing different models, which is a different application from indicating total survey error, as discussed above.

Note that for measurement error studies involving repeated measurements, the reader is referred to work by Paul Biemer and others, such as that found in Biemer, Groves, Lyberg, Mathiowetz and Sudman(1991). Another good reference, for nonsampling errors, is Lessler and Kalsbeek(1992).

Acknowledgement: We make progress when we listen. Thank you to all who have made the effort to discuss in one way or another.

Appendix Comparing Variance Computations:

When using only one regressor and a zero intercept, the "exact" variance with regard to the model, $y_i = \overset{*}{\beta} x_i + e_{0i} x_i^{\gamma}$, with $\overset{*}{\beta} \equiv$ a WLS estimate of β, alternatively written below as "b," is

(1) $V_L^*(T^* - T) = \sum_{N-n} \sigma_e^{*2}/w_i + \left(\sum_{N-n} x_i\right)^2 V^*(b)$, $w_i = \left[x_i^{\gamma}\right]^{-2}$. See Knaub(1999).

(If γ = 0.5, then one has the model-based ratio estimate with variance proportionate to x.)

More generally, $y_i = \sum_j \overset{*}{\beta}_j x_{j,i} + e_{0i} z_i^{\gamma}$, or $y_i = \mathbf{x}_i \overset{*}{\boldsymbol{\beta}} + e_{0i} z_i^{\gamma}$, where *z* may be a preliminary estimate of *y*, say approximately *y**, or any other measure of 'size,' and the intercept is zero. The general approximation for the variance of the total, using any number of regressors, easily adaptable to changing survey requirements, also found in Knaub(1999), is

(2) $V_L^*(T^* - T) = \delta(N-n) \sum_{N-n} \left\{ V_L^*(y_i^* - y_i) - \frac{\sigma_e^{*2}}{w_i} \right\} + \sum_{N-n} \frac{\sigma_e^{*2}}{w_i}$, where δ is not known exactly. For establishment surveys, $\delta \approx 0.3$ appears to do well. Variance estimation is accomplished by aggregating over the N-n cases where there are no observed data. Here we have $w_i = \left[z_i^{\gamma}\right]^{-2}$.

If we sum over the predictions for all the members of the finite population, even for predictions where observed data are available, the relative standard error of that total, the model-based RSESP, is the square-root of the estimated variance of the estimated total, $V_{SP}^*(T^* - T)$, divided by that estimated total. These estimated variances are the $V_L^*(T^* - T)$ above when we sum over N rather than N-n. Thus, in general:

$$V_{SP}^*(T^* - T) = \delta N \sum_N \left\{ V_L^*(y_i^* - y_i) - \frac{\sigma_e^{*2}}{w_i} \right\} + \sum_N \frac{\sigma_e^{*2}}{w_i}.$$

Fun Facts:

1) When the data are grouped carefully for modeling purposes ("estimation groups" in Knaub (1999)), then otherwise "nonignorable" nonresponse can sometimes be made "ignorable."

2) For many of the smallest electric power establishments, it appears that data collected frequently may often be imputed more accurately than they may be observed.

3) Sometimes an obviously incorrect value for γ will result in a lower estimate of RSE, so the lowest estimated RSE is not necessarily a good test for deciding which γ to use.

Regarding the Model-Based Ratio Estimator (γ=0.5, zero intercept; the Classical Ratio Estimator, CRE, in October 2005 *InterStat*)**:**

4) $\sum y_i = \sum \overset{*}{y}_i$ when γ = 0.5, for the model-based ratio estimator. That is, if all observed values are replaced by predicted values, using a model-based ratio estimate, γ = 0.5, then the sum of the observed values will be equal to the sum of the corresponding predicted replacement values.

5) Only when γ = 0.5 can a lump sum of the regressor data for the unobserved cases be enough information for estimation of the variance of the estimated total, as shown on pages 4 and 5 of Knaub, in Proceedings of the Survey Research Methods Section, ASA 1991, found at the following URL:
http://www.amstat.org/sections//srms/proceedings/papers/1991_133.pdf.

References

Biemer, Groves, Lyberg, Mathiowetz and Sudman (1991), Measurement Errors in Surveys, John Wiley & Sons, Inc.

Brewer, KRW (2002), Combining survey sampling inferences: The weighing of Basu's elephants, Arnold: London.

Dobbs, Gibbins, Martin, Davies and Dodd(1998), "Reporting on Data Quality and Process Quality," ASA Survey Research Methods Section proceedings, 1998. Also found at http://www.amstat.org/sections/srms/Proceedings/papers/1998_005.pdf.

Federal Committee on Statistical Methodology, Subcommittee on Measuring and Reporting the Quality of Survey Data (2001), Measuring and Reporting Sources of Errors in Surveys. Washington, DC: U.S. Office of Management and Budget (Statistical Policy Working Paper 31), NTIS PB2001-104329, http://www.fcsm.gov/01papers/SPWP31_final.pdf.

Knaub, J.R., Jr. (1999), "Using Prediction-Oriented Software for Survey Estimation," InterStat, August 1999, http://interstat.stat.vt.edu, partially covered in "Using Prediction-Oriented Software for Model-Based and Small Area Estimation," in ASA Survey Research Methods Section proceedings, 1999, and partially covered in "Using Prediction-Oriented Software for Estimation in the Presence of Nonresponse," presented at the International Conference on Survey Nonresponse, 1999.

Knaub, J.R., Jr. (2001), "Using Prediction-Oriented Software for Survey Estimation - Part III: Full-Scale Study of Variance and Bias," InterStat, June 2001, http://interstat.stat.vt.edu. (Note shorter version in ASA Survey Research Methods Section proceedings, 2001, JSM CD.)

Knaub, J.R., Jr. (2002), "Practical Methods for Electric Power Survey Data," detailed version, InterStat, July 2002, http://interstat.stat.vt.edu. (Note shorter version in ASA Section on Government Statistics proceedings, 2002, JSM CD.)

Knaub, J.R., Jr. (2003), "Applied Multiple Regression for Surveys with Regressors of Changing Relevance: Fuel Switching by Electric Power Producers," InterStat, May 2003, http://interstat.stat.vt.edu. (Note, also to be found in ASA Section on Government Statistics proceedings, 2003, JSM CD.)

Korn, Edward L., and Graubard, Barry I. (1998), "Variance Estimation for Superpopulation Parameters," Statistica Sinica, Vol. 8, No. 4, co-sponsored by the Institute of Statistical Science, Academia Sinica and the International Chinese Statistical Association, found at URL http://www.stat.sinica.edu.tw/statistica/.

Lee, H., Rancourt, E., and Sarndal, C.-E. (1999), "Variance Estimation from Survey Data Under Single Value Imputation," presented at the International Conference on Survey Nonresponse, Oct. 1999, published in Survey Nonresponse, ed. by Groves, Dillman, Eltinge and Little, 2002, John Wiley & Sons, Inc., pp 315-328.

Lessler, Judith T., and Kalsbeek, William D. (1992), Nonsampling Error in Surveys, John Wiley & Sons, Inc.

Maddala, G.S. (1992), Introduction to Econometrics, 2nd ed., Macmillan Pub. Co.

Royall, R.M. (1970), "On Finite Population Sampling Theory Under Certain Linear Regression Models," Biometrika, 57, pp. 377-387.

Waugh, S., Norman, K. and Knaub, J. (2003) "Proposed EIA Guidance on Relative Standard Errors (RSEs)," Presentation to the ASA Committee on Energy Statistics, October 17, 2003, http://eia.doe.gov/smg/asa_meeting_2003/fall/files/rseguidance.pdf.