DECOMPOSING DESIGN EFFECTS

Thanh Lê, J. Michael Brick, and Graham Kalton, Westat
Thanh Lê, Westat, 1650 Research Boulevard, Rockville, Maryland 20850

Key Words: Intraclass correlation, Stratification, Clustering, Unequal probability sampling

1. Introduction

Complex sample designs may include multistage sampling, stratification, and unequal selection probabilities. A consequence of the use of a complex sample design is that the sampling errors of the survey estimates cannot be computed assuming that the observed variables are independently and identically distributed (*iid*) random variables. Variances of survey estimates from complex sample designs may be estimated correctly by replication or by a Taylor Series linearization method (Wolter, 1985). When variances are computed properly accounting for the complex sample design, they typically are larger than those obtained from standard *iid* formulas.

Kish coined the term "design effect" to denote the ratio of the variance of any estimate *z* obtained from a complex design to the variance of *z* that would apply with a SRS or unrestricted sample of the same size (Kish, 1965). The denominator of the *Deff* is the SRS variance. An alternative to the *Deff* is the $Deft^2$, which has the same numerator but the denominator is the variance of an unrestricted sample (Kish, 1982). The only difference between *Deff* and $Deft^2$ is whether the fpc term is included. In our discussion, we assume that the fpc term can be ignored and thus use $Deft^2$. Since the design effect depends both on the sample design and on the survey estimate under consideration, we employ the notation $D^2(z)$ for the design effect ($Deft^2$) of the estimate *z* (mean, proportion, total, or regression coefficient), i.e.,

$$D^2(z) = \frac{\text{Variance of } z \text{ with complex design}}{\text{Variance of } z \text{ with unrestricted sample of the same size}} . \qquad (1)$$

The concept of a design effect has proven to be a valuable tool in the design of complex samples. It has also been used in the analysis of data from such samples, but with modern computing facilities the role of the design effect in analysis should diminish. Here we restrict our attention to uses in the design of surveys, particularly household surveys. This paper is based on a more extensive examination of these issues contained in a chapter prepared for a United Nations technical report (United Nations Statistical Division, forthcoming). The inclusion of a chapter in that report on using design effects in designing household surveys in developing countries attests to the practical importance of the topic.

Complex designs involve a combination of a number of design components, such as stratification, multi-stage sampling, and selection with unequal probabilities. The analysis of the design effects for each of these components individually sheds useful light on their effects on the precision of survey estimates, and thus helps guide the development of efficient sample designs. We review methods of estimating the design effects for individual components in Section 2, with special emphasis on the assumptions that are needed to compute design effects using these formulas. Since modeling the overall design effects that might arise from a specific combination of components is an important function, we briefly review these models in Section 3. We conclude with some general observations about the uses of design effects in the design of surveys in Section 4.

2. Components of Design Effects

In this section, we examine separately the design effects resulting from the following components of a complex sample design: stratification, unequal selection probabilities, clustering, sample weighting adjustments for nonresponse, and population weighting adjustments for noncoverage and for improved precision. The joint effects are discussed in Section 3. The main statistic considered is an estimate of a population mean $\bar{Y}$. Since a population proportion is a special case of a mean, the treatment covers proportions also. The effects of weighting and clustering on the design effects of subgroup estimates is also discussed.

2.1 Stratification

In a stratified simple random sample, the sample mean is

$$\bar{y}_{st} = \Sigma_h \frac{N_h}{N} \Sigma_i \frac{y_{hi}}{n_h} = \Sigma_h A_h \bar{y}_h , \qquad (2)$$

where n_h is the size of the sample selected from the N_h units in stratum *h*, $N = \Sigma N_h$ is the population size, $A_h = N_h / N$ is the proportion of the population in stratum *h*, y_{hi} is the value for sampled unit *i* in stratum *h*, and $\bar{y}_h = \Sigma_i y_{hi} / n_h$ is the sample mean in stratum *h*. In practice, $\bar{y}_{st}$ is computed as a weighted estimate, where each sampled unit is assigned a base weight that is the inverse of its selection probability (ignoring for the moment sample and population weighting adjustments).

Here each unit in stratum *h* has a selection probability of n_h / N_h and hence a base weight of $w_h = N_h / n_h$. Thus $\bar{y}_{st}$ may be expressed as

$$\bar{y}_{st} = \frac{\Sigma_h \Sigma_i w_{hi} y_{hi}}{\Sigma_h \Sigma_i w_{hi}} = \frac{\Sigma_h \Sigma_i w_h y_{hi}}{\Sigma_h n_h w_h} . \qquad (3)$$

Assuming that the fpc can be ignored, the variance of the stratified mean is given by

$$V(\bar{y}_{st}) = \Sigma_h \frac{A_h^2 S_h^2}{n_h} , \qquad (4)$$

where $S_h^2 = \Sigma_i (Y_{hi} - \bar{Y}_h)^2 / (N_h - 1)$ is the population unit variance within stratum *h*.

With proportionate allocation, the weights for all sampled units are the same and the stratified mean reduces to the simple unweighted mean $\bar{y}_{prop} = \Sigma\Sigma y_{hi} / n$, where $n = \Sigma n_h$ is the overall sample size. Its variance reduces to

$$V(\bar{y}_{prop}) = \frac{\Sigma A_h S_h^2}{n} = \frac{S_w^2}{n} , \qquad (5)$$

where S_w^2 denotes the average within-stratum unit variance. The design effect for $\bar{y}_{prop}$ for a proportionate stratified simple random sample is

$$D^2(\bar{y}_{prop}) = \frac{S_w^2}{S^2} . \qquad (6)$$

Since the average within-stratum unit variance is no larger than the overall unit variance (provided that N_h is large), the design effect for the mean of a proportionate sample is no greater than 1. Thus, proportionate stratification cannot lead to a loss in precision, and generally leads to some gain in precision. A gain in precision occurs when the strata means $\bar{Y}_h$ differ: the larger the variation between the means, the greater the gain.

In many surveys a disproportionate stratified sample is needed to meet the needs of the survey to provide estimates for particular domains. The gain in precision from proportionate stratification does not necessarily apply with a disproportionate allocation of the sample. To simplify the discussion for this case, we *assume that the within-stratum population variances are constant*, i.e., $S_h^2 = S_c^2$ for all strata. Under this assumption equation (4) simplifies to

$$V(\bar{y}_{st}) = S_c^2 \Sigma_h \frac{A_h^2}{n_h} = \frac{S_c^2}{N} \Sigma_h A_h w_h . \qquad (7)$$

The design effect in this case is

$$D^2(\bar{y}_{st}) = \frac{S_c^2}{S^2} \frac{n}{N} \Sigma_h A_h w_h . \qquad (8)$$

In addition to the assumption of constant within-stratum variances used in deriving equation (8), it may be reasonable to *assume approximately equal stratum means*, that is to assume that $\bar{Y}_h = \bar{Y}$ for all strata. This type of assumption is often useful in multipurpose surveys and when strata are defined to provide estimates for domains such as regions. With this further assumption, $S_c^2 = S^2$ and the design effect reduces to

$$D^2(\bar{y}_{st}) = \frac{n}{N} \Sigma_h A_h w_h = \frac{n}{N^2} \Sigma_h \frac{N_h^2}{n_h} . \qquad (9)$$

Kish (1992) presents a formula for design effect due to disproportionate allocation under the above assumptions in a form that is very useful in practice:

$$D^2(\bar{y}_{st}) = (\Sigma_h A_h w_h)(\Sigma_h A_h / w_h) . \qquad (10)$$

Equation (10) is widely used in sample design to assess the effect of the use of disproportionate allocation on national estimates. In employing it, however, the assumptions of equal within-stratum means and variances are often ignored. When the assumptions are ignored, the predicted design effects may be inappropriate. For example, consider the situation where the means are different but the variances are not. In this case, the design effect from disproportionate stratification is given by equation (8), which is equation (10) with the additional factor S_c^2 / S^2. Since this factor is less than 1, the design effect is overestimated using equation (10). If we wish to isolate the effect of the disproportionate allocation without the stratification effect, the appropriate comparison is between the disproportionate stratified sample and a proportionate stratified sample of the same size. The ratio of the variance of $\bar{y}_{st}$ for the disproportionate design to that of $\bar{y}_{prop}$ is derived from equations (5) and (7) with $S_w^2 = S_c^2$,

$$R = V(\bar{y}_{st}) / V(\bar{y}_{prop}) = (\Sigma_h A_h w_h)(\Sigma_h A_h) / w_h .$$

Thus, in this case, formula in equation (10) can be interpreted as the effect when the weights varying across the strata.

The second assumption of equal within-stratum unit variances is more critical than the equal means assumption. While disproportionate allocation leads to a loss of precision in overall estimates when within-stratum unit variances are equal, this result does not necessarily hold when the within-stratum unit variances are unequal. Indeed, when within-stratum variances are unequal, the optimum sampling fractions to use are proportional to the standard deviations in the strata (Cochran, 1977). This type of disproportionate allocation is widely used in business surveys. It can lead to substantial gains in precision over proportionate allocation when the within-stratum standard deviations differ markedly.

2.2 Clustering

Another source of design effects is selecting clusters of units. Household samples are clustered to reduce data collection costs. Typically two or more stages of sampling are employed, where the first-stage units or primary sampling units (PSUs) are geographical areas sampled with probabilities proportional to the estimated numbers of households or persons. Within the selected PSUs, one or more additional stages of area sampling may be conducted, but for the purposes of this discussion we assume a household survey with only two stages of sampling (PSUs and households). The extension to multiple stages is direct.

In most household surveys, PSUs vary in size and are sampled with probability proportional to estimated size (PPES) to reduce the variance of the estimates. In addition, the sample sizes selected from selected PSUs may vary between PSUs. However, for simplicity, we start with the assumption that the population consists of A PSUs, each of which contains B units. *An SRS of a* of the *A PSUs is selected and an SRS of b of the B units is selected in each selected PSU. We assume that the fpc can be ignored*. Since the sample is epsem, the population mean can be estimated by the simple unweighted sample mean $\bar{y}_{cl} = \sum_{\alpha}^{a} \sum_{\beta}^{b} y_{\alpha\beta} \big/ n$, where $n = ab$ and the subscript *cl* is for cluster. The variance of $\bar{y}_{cl}$ can be written as

$$V(\bar{y}_{cl}) = \frac{S^2}{n}\left[1 + (b-1)\rho\right], \tag{11}$$

where S^2 is the unit variance in the population and ρ is the intra-class correlation coefficient that measures the homogeneity of the *y*-variable in the PSUs. In practice, units within a PSU tend to be somewhat similar to each other for most variables, although the degree of similarity is usually low. Hence, typically ρ is positive and small. The design effect in this simple situation is

$$D^2(\bar{y}_{cl}) = 1 + (b-1)\rho. \tag{12}$$

Since ρ is generally positive, the design effect tends to be greater than 1 with clustered samples. Frequently ρ is inversely related to the size of the PSU because larger clusters tend to be more diverse, especially when PSUs are geographic areas. These types of relationships are exploited in the optimal design of surveys, where PSUs that are large and more diverse are used when there is an option. Estimates of ρ for key survey variables are needed for planning sample designs. These estimates are usually based on estimates from previous surveys for the same or similar variables and PSUs, and the belief in the portability of the values of ρ across similar variables and PSUs.

In more practical settings, PSUs are not of equal size and they are selected using PPES sampling. As a result, equation (12) does not directly apply. However, the approximation may still be a useful model of the design effect due to clustering for epsem sample designs, provided suitable modifications are made in the interpretation of ρ. These situations are covered below.

We now consider an unstratified sample of PSUs selected with probability proportional to size (PPS), where the exact measures of size are known. The *assumption of an unstratified PPS sample of a of the A PSUs and an epsem sample of b of the B units from each sampled PSU* results in an overall epsem sample. With such a design, equation (12) still holds, but with ρ now interpreted as a synthetic measure of homogeneity within the ultimate clusters created by the subsample design (Kalton, 1979). The value of ρ is, for instance, different for a subsample design that selects *b* units by systematic sampling from one in which the subsample design divides each sampled PSU into subareas containing *b* units each and selects one subarea (the value of ρ is likely to be larger in the latter case). This extension thus deals with both PPS sampling and with various alternative forms of subsample design.

Now we relax the design assumption to allow stratification of the PSUs. Kalton (1979) shows that the design effect due to clustering *in an overall epsem design in which a stratified sample of a of A PSUs is selected and b of B elementary units are sampled with equal probability within each of the selected PSUs* can be approximated by

$$D^2(\bar{y}_{cl}) \approx 1 + (b-1)\bar{\rho}, \tag{13}$$

where $\bar{\rho}$ is the average within-stratum measure of homogeneity, *provided that the homogeneity within each stratum is roughly of the same magnitude*. The gain from effective stratification of PSUs can be substantial when *b* is sizable because the overall measure of homogeneity in (12) is replaced by a smaller within-stratum measure of homogeneity in (13). In other words, the reduction in the design effect of $(b-1)(\rho - \bar{\rho})$ can be large when *b* is sizable.

In practice, unequal-sized PSUs are sampled by PPES, with estimated measures of size that are inaccurate to some degree. In this case, the application of the subsampling rates in the sampled PSUs to give an overall epsem design results in some variation in subsample size. *Provided that the variation in the subsample sizes is not large*, then equation (13) may still be an adequate approximation with *b* being replaced by the average subsample size, that is

$$D^2(\bar{y}_{cl}) \cong 1+(\bar{b}-1)\bar{\rho} \quad (14)$$

where $\bar{b}=\sum b_\alpha / a$ and b_α is the number of elementary units in PSU α. Equation (14) has proven to be of great practical utility for situations in which the number of sampled units in each of the PSUs is relatively constant.

When the variation in the subsample sizes per PSU is substantial, however, the approximation involved in equation (14) becomes inadequate. Holt (1980) extends the above approximation to deal with unequal subsample sizes by replacing $\bar{b}$ in (14) by a weighted average subsample size. The design effect due to *clustering with unequal cluster sizes* can be written as

$$D^2(\bar{y}_{cl}) \cong 1+(b'-1)\bar{\rho}, \quad (15)$$

where $b'=\sum b_\alpha^2 / \sum b_\alpha$. (The quantity b' can be thought of as the weighted average $b'=\Sigma k_\alpha b_\alpha / \Sigma k_\alpha$, where $(k_\alpha = b_\alpha)$.) This approximation, as well as all of the others considered for clustering thus far, *assumes an overall epsem sample design*.

The results given above can also be applied to subgroup estimates, but careful attention to the underlying assumptions is required. For handling subgroup estimates we introduce a three-fold classification of types of subgroups according to their distributions across the PSUs. At one end, there are subgroups that are evenly spread across the PSUs, known as crossclasses. For example, age/sex subgroups are generally crossclasses. At the other end of the spectrum, there are subgroups, that are concentrated in a subset of PSUs, called segregated classes. Urban and rural subgroups are examples of this type. In between, there are subgroups that are somewhat concentrated by PSU, called mixed classes.

As noted above, crossclasses have, by definition, the same distribution as the total sample across the PSUs. *If the total sample is fairly evenly distributed across the PSUs*, then equation (14) may be used to compute an approximate design effect from clustering for a crossclass. However, in this case $\bar{b}$ is the average crossclass subsample size per PSU rather than the full sample size. Because the crossclass estimate has a smaller cluster sample size, design effects for crossclass estimates are smaller than those for total sample estimates.

Segregated classes comprise all or most of the units in a subset of the PSUs in the full sample. Since the subclass sample size for a segregated class is the same as that for the total sample in that subset of PSUs, the design effect for an estimate for a segregated class are approximately equal to the design effects for a total sample estimate. The design effect for an estimate for a segregated class will differ from that for a total sample estimate only if the average subsample size per PSU in the segregated class differs from that in the total sample, or if the homogeneity differs. The homogeneity might differ if, for example, the subsample design in the segregated class is not the same as the subsample design for other sampled groups. If the total sample is evenly spread across the PSUs, equation (14) may again be applied, with $\bar{b}$ and ρ being values for the set of PSUs in the segregated class.

The uneven distribution across the PSUs associated with a mixed class implies that equation (14) is not applicable. For estimating the design effect from clustering for an estimate from a mixed class, equation (15) may be a reasonable approximation, with b_α being the number of sampled members of the mixed class in PSU α.

2.3 Weighting Adjustments

As discussed earlier, designs with unequal selection probabilities between strata with disproportionate stratification require the use of weights in the analysis of the survey data. Equations (9) and (10) give the design effect arising from the disproportionate stratification and resulting unequal weights under the assumptions that the strata means and unit variances are all equal. Variable weights are needed in the analysis to compensate for unequal selection probabilities, but these do not arise only from disproportionate stratification. Even without oversampling of certain domains, sample designs usually deviate from equal probability of selection method (epsem) because of frame problems. For example, if households are selected with equal probability from a frame of households and then one household member is selected at random in each selected household, household members are sampled with unequal probabilities. In addition, the base weights, the inverses of the probabilities of selection, are often then adjusted to compensate for nonresponse and to make weighted sample totals conform to known population totals. As a result, final weights are almost always have more variability than implied by disproportionate stratification.

It is important to consider the effects these later adjustments of the weights may have on the precision of the estimates when the sample is being designed, because otherwise the estimates may not have the appropriate precision. The same formulas may also be used after the data are collected in the analysis stage, but that is not the focus of this review.

Virtually all surveys encounter some amount of nonresponse. A common approach used to reduce possible nonresponse bias involves differentially adjusting the base weights of the respondents. These weighting adjustments cause the weights to vary from the base weights and the effect is often an increase in the design effect of an estimate.

When related population information is available from some other source, the nonresponse-adjusted weights may be further adjusted to make the weighted sample estimates conform to the population information. For example, if good estimates of regional population sizes are available from an external source, the sample estimates of these regional populations can be made to coincide with the external estimates. A poststratification or some other calibration adjustment of the weights is often used for this population adjustment. The adjustment can help to compensate for noncoverage and can improve the precision of some survey estimates. However, it adds further variability to the weights and this can adversely affect the precision of survey estimates that are unrelated to the population variables employed in the adjustment.

Kish (1992) presents another way of expressing the design effect for a stratified mean that is a different way of expressing equations (9) and (10) . This expression is based on the *same assumptions given earlier* for those equations. We denote it with a small *d* since it is often computed after the sample is observed.

$$d^2\left(\bar{y}_{st}\right) = \frac{n \sum_h \sum_i w_{hi}^2}{\left(\sum_h \sum_i w_{hi}\right)^2} = 1 + rv\left(w_{hi}\right), \qquad (16)$$

where $rv\left(w_{hi}\right)$ is the relvariance of the weights and $\bar{w}$ is the mean of the weights. A more widely used form of this equation is

$$d^2\left(\bar{y}_{st}\right) = \frac{n \sum_j w_j^2}{\left(\sum_j w_j\right)^2} = 1 + cv^2\left(w_j\right), \qquad (17)$$

where each of the n units in the sample has its own weight w_j $\left(j = 1, 2, ..., n\right)$.

As noted above, the design effect due to unequal weighting given by equation (17) depends on the *assumptions of equal strata means and unit variances*, particularly the latter. This equation can provide a good measure of the effect of weighting if these assumptions hold at least approximately.

Nonresponse adjustments are generally made within classes defined by auxiliary variables known for both respondents and nonrespondents. To be effective in reducing nonresponse bias, the variables measured in the survey do need to vary across these weighting classes. However, generally the variation is not great, particularly in the unit variance. As a result, equation (17) is widely used to examine the effect of nonresponse weighting adjustments on the precision of survey estimates. At the planning stage, the expected response rate and experience with previous nonresponse adjustments may be used to predict the effect of the nonresponse weighting adjustments on the design effect.

While equation (17) is reasonable for most nonresponse sample weighting adjustments, it is often not a good approximation for the effect of population weighting adjustments. In particular, when the weights are poststratified or calibrated to known control totals from an external source, then the design effect for the mean of *y* is poorly approximated by equation (17) when *y* is highly correlated with the one or more of the control totals. For example, assume the weights are poststratified to control totals of the numbers of persons in the population by sex. Consider the extreme case that the survey data are used to estimate the proportion of women in the population. In this case of perfect correlation between the *y* variable and the control variable, the estimated proportion is not subject to sampling error and hence has zero variance. In practice, the correlation will not be perfect, but it may be sizable for some of the survey variables.

All the results presented in this section and Section 2.1 can be applied straightforwardly to give the design effects for subgroup estimates simply by restricting the calculations to subgroup members. However, care must be taken in trying to infer the design effects from weighting for subgroup estimates from results for the full sample. For this inference to be valid, the distribution of weights in the subgroup must be similar to that in the full sample.

3. Models for Design Effects

The previous section presented some results for design effects associated with weighting and clustering separately, with a primary focus on design effects for means and proportions. Those results can be extended to the design effects from a combination of weighting and clustering and the design effects for some other types of estimates. A number of models have been used to represent the design effects for these extensions, both in the design and analysis of complex sample designs (Kalton, 1977; Wolter, 1985). We briefly mention here that the calculation of the design effect $d^2\left(\bar{y}\right) = v_c\left(\bar{y}\right) / v_u\left(\bar{y}\right)$ encompasses the combined effects of weighting and clustering. However, in using the data from the current survey to plan a future survey, the two components of the design effect need to be separated. For example, the future survey may be planned to be an epsem one whereas the current survey may have oversampled certain domains. Also, although using the same PSUs and stratification, the future survey may want to change the subsample size per PSU.

Here we give a model that may be used *when the weights are random or approximately so*. In this case, the overall design effect can be decomposed approximately into a product of the design effects of weighting and clustering:

$$d^2(\bar{y}) = d_w^2(\bar{y}).d_{cl}^2(\bar{y}) \quad (18)$$

where $d_w^2(\bar{y})$ is the design effect from weighting as given by equation (17) and $d_{cl}^2(\bar{y})$ is the design effect from clustering given by equation (14) or (15). Using this model with equation (14), $\bar{\rho}$ is thus estimated by

$$\hat{\rho} = \frac{\left[d^2(\bar{y}) / d_w^2(\bar{y}) - 1\right]}{\bar{b} - 1}. \quad (19)$$

Generally, $\bar{\rho}$ is a more important parameter to estimate for planning purposes than the design effect from clustering because it is more portable across different designs. The design effect for one survey can be directly applied in planning another only if the subsample size per PSU remains unchanged.

4. Conclusion

An understanding of design effects and their components is valuable in developing sample designs for new surveys. For example, the magnitudes of the overall design effects for key survey estimates may be used in determining the required sample size. The sample size needed to give the specified level of precision for each key estimate may be computed for an unrestricted sample, and this sample size may then be multiplied by the estimate's design effect to give the required sample size for that estimate with the complex sample design. When a disproportionate stratified sample design is to be used to provide domain estimates of required levels of precision, the resultant loss of precision for estimates for the total sample and for subgroups that cut across the domains can be assessed by computing the design effect due to variable weights. If the loss is found to be too great, then a change in the domain requirements that leads to less variable weights may be indicated. If the design effect from clustering is very large for some key survey estimates, then the possibility of increasing the number of sampled PSUs (*a*) with a smaller subsample size (*b*) may be considered.

Estimating design effects from clustering requires estimates of ρ values for the key survey variables. These estimates are inevitably imperfect, but reasonable estimates may suffice. To err towards the use of a value of ρ larger than predicted leads to the specification of a larger required sample size, and hence this is a conservative strategy.

Finally, it should be noted that the purpose of using these design effect models is to produce an efficient sample design. A failure of the models to hold exactly will result in some loss of efficiency. However, the use of inappropriate models to develop the sample design does not affect the validity of the survey estimates. With probability sampling, the survey estimates remain valid estimates of the population parameters.

Examples of design effects, more discussions of models of design effects, and a complete list of references can be found in the UN report on the operating characteristics of household surveys in developing and transition countries (United Nations Statistical Division, forthcoming).

5. References

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

Holt, D. H. (1980). Discussion of the paper by Verma, V., Scott, C. and O'Muircheartaigh, C. Sample designs and sampling errors for the World Fertility Survey. *Journal of the Royal Statistical Society A*, 143, 468-469.

Kalton, G. (1977). Practical methods for estimating survey sampling errors. *Bulletin of the International Statistical Institute*, 47, 495-514.

Kalton, G. (1979). Ultimate cluster sampling. *Journal of the Royal Statistical Society A*, 142, 210-222.

Kish, L. (1965). *Survey Sampling*. Wiley, New York.

Kish, L. (1982). Design effect. In *Encyclopedia of Statistical Sciences*, S. Kotz and N.L. Johnson eds., Volume 2, 347-348. Wiley, New York.

Kish, L. (1992). Weighting for unequal P_i. *Journal of Official Statistics,* 8, 183-200.

Kish, L. (1995). Methods for design effects. *Journal of Official Statistics*, 11, 55-77.

United Nations Statistical Division. (Forthcoming). *An Analysis of Operating Characteristics of Household Surveys in Developing and Transition Countries: Survey Costs, Design Effects, and Non-sampling Error*. (Provisional title)

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