DESIGN EFFECTS FOR LINEAR CONTRASTS OF PROPORTIONS AND LOGITS

James M. Lepkowski and J. Richard Landis
Survey Research Center and University of Michigan

SUMMARY

The design effects of two measures of interaction, linear contrasts of subclass proportions (L) and logits (L_g), are examined analytically and empirically. The effects of clustered selection are expressed in terms of measures of intracluster homogeneity and measures reflecting the degree of covariance of cluster means for different characteristics across clusters, or cross-homogeneity. The design effects L and L_g are shown to be nearly identical. Three different models are proposed for the design effects of these linear contrasts in terms of a weighted average of the design effects of the subclass populations, the intracluster homogeneity and the cross-homogeneity quantities. The design effects are estimated for fifty tables generated from data from the Health Examination Survey. They are shown to be smaller than the average subclass proportion design effects from the tables and to depend on the size of the average subclass proportion design effect.

1. INTRODUCTION

Increasingly, design effects and functions of design effects are being used to adjust analytic statistics computed under independence assumptions for complexities in the sample design. As proposed by Kish (1962), design effects are summary measures of the effects of clustered sample selection and other features of complex survey sample designs on the precision of estimates. They are useful in the planning of survey samples, especially when models for design effects are available which utilize measures that are somewhat "portable" across survey designs (see e.g., Kish, Groves and Krotki (1976)).

The analysis of contingency table data arising from complex sample surveys has recently been the focus of developments concerning such design effect adjustments. Since many of the variables measured in such surveys are reported on categorical data scales, computational techniques such as iterative proportional fitting or weighted least squares are frequently used to fit log-linear or linear models to the corresponding multidimensional contingency tables. Consequently, the effects of complexities in the sample design on the resulting statistics are of considerable interest to analysts.

For instance, Fellegi (1978) suggests that the ordinary chi-square statistic for independence developed under simple random sample assumptions be divided by an average design effect for the proportions in the table to adjust for the effects of a complex sample design. Rao and Scott (1979) find such an adjustment overly conservative. For a contingency table with r cells, and $\{p_j\}$ and $\{\text{deff }(p_j)\}$ denoting the set of proportions and design effects of the proportions in the table, respectively, they suggest dividing the ordinary chi-square test statistic by a factor $[1/(r-1)]\sum_j(1-p_j)\text{ deff }(p_j)$. This factor will be larger than an average design effect since $(1-p_j) \leq 1$; hence the adjustment is less conservative.

In the subsequent discussion, design effects for two statistics useful in the analysis of categorical data will be examined. These will suggest the size and nature of factors appropriate for adjustment of ordinary chi-square test statistics in some analytic problems. Two approaches to the examination of these design effects will be developed. Analytically, the design effects will be modeled in terms of synthetic measures of homogeneity following Kish (1965, chapter 5). Empirically, design effects and various parameters suggested in the analytic development will be estimated from two survey data sets. These developments will be utilized to examine the nature of design effects and of adjustments to ordinary chi-square test statistics applied to data from complex sample designs.

2. NOTATION AND BACKGROUND

Consider the factor-response framework for the analysis of multidimensional contingency tables (Bhapkar and Koch, 1968) illustrated in Table 1. Variables used to classify the sample elements into the cells of the table are of two types, factors and responses. The factors are variables that are used to classify sample elements into subpopulations or subclasses, and are often referred to as independent variables. The response variables are used to classify elements into categories of a response profile and are often referred to as dependent variables. The terminology of experimental design is thus borrowed to clarify features of the analysis of categorical data from observational studies.

Table 1: An (s x r) Sample Contingency Table

	Response categories				
Subclass	1	2	...	r	Total
1	n_{11}	n_{12}	...	n_{1r}	n_1
2	n_{21}	n_{22}	...	n_{2r}	n_2
⋮	⋮	⋮		⋮	⋮
s	n_{s1}	n_{s2}	...	n_{sr}	n_s

Let π_{ij} denote the proportion of elements in a finite population from which a sample is to be selected that are in the ith subclass and are classified into the jth category of the response profile corresponding to Table 1. Let $p_{ij} = n_{ij}/n_i$ denote the corresponding sample proportion. If the selection procedure is simple random sampling, denote the variance of p_{ij} as $\sigma^2_{ij} = \pi_{ij}(1-\pi_{ij})/n_i$.

Suppose, however, that the sample selection is clustered where A primary selections or clusters each of k_α elements are selected such that $\sum_\alpha k_\alpha = n$, the sample size. Let d^2_{ij} denote the design effect of p_{ij}, $\bar{b}_i = n_i/A$ the average cluster size for the ith subclass, and Roh_{ij} a measure of the homogeneity within clusters for elements in the (i,j)th cell. Then following the proposed model suggested in Kish, Groves and Krotki (1976), the variance of p_{ij} is

$$\text{Var}(p_{ij}) = \sigma^2_{ij}\, d^2_{ij} / n_i$$
$$= \sigma^2_{ij}\, [1 + (\bar{b}_i - 1)\, Roh_{ij}]/n_i \ . \quad (1)$$

If a model for the variance of p_{ij} in (1) is applied to data from a complex sample, the homogeneity term Roh_{ij} reflects more than a clustering effect, in particular the effects of stratification, multi-stage selection, weighting, and other design features. Empirical investigations of Roh_{ij} (e.g., Kish, Groves and Krotki, 1976) indicate that $Roh_{ij} > 0$, that $d^2_{ij} > 1$, and that Roh_{ij} is useful not only for estimating d^2_{ij} for one survey but is portable to other surveys and new sample designs.

For contingency tables derived from data obtained by clustered sample selection for which subclass proportions are to be compared, the covariance among subclass proportions also must be considered. In simple random sampling such covariances are zero; in clustered sampling the covariance is usually greater than zero. Let $\text{Cov}(p_{ij}, p_{i'j})$ denote the covariance of the two subclass proportions p_{ij} and $p_{i'j}$, $i \neq i'$, and let $R_{ii'j} = \text{Cov}(p_{ij}, p_{i'j}) / [\text{Var}(p_{ij})\ \text{Var}(p_{i'j})]^{\frac{1}{2}}$ denote the subclass correlation. The subclass covariance and correlation are induced by the homogeneity arising from the clustering in the sample design. Just as homogeneity is a useful concept for examining the nature of design effects for proportions, it is possible to propose a subclass cross-homogeneity, denoted $Rohs_{ii'j}$, to examine the effects of subclass covariance and correlation on comparisons of subclass proportions. Subclass cross-homogeneity is a measure of the tendency of elements classified into the same response category but different subclasses to be in the same clusters. In particular, let $R_{ii'j}$ be expressed as

$$R_{ii'j} = (\bar{b}_{ii'} - 1)\ Rohs_{ii'j}/d_{ij}\, d_{i'j} \ , \quad (2)$$

where $\bar{b}_{ii'} = 2/[(n_i^{-1} + n_{i'}^{-1})A]$ denotes an average harmonic mean of subclass sizes n_i and $n_{i'}$ across clusters and $d_{ij} = \sqrt{d^2_{ij}}$, $i \neq i'$. That is, expression (2) suggests that subclass correlation depends on average subclass sizes, on subclass proportion design effects, and on some measure of cross-homogeneity between the two subclass proportions.

Subclass cross-homogeneity is important because it can affect the variance of comparisons and contrasts of subclass proportions. If two subclass proportions with similar subclass homogeneities, Roh_{ij}, are based on elements distributed in different clusters, $Rohs_{ii'j}$ can be smaller than Roh_{ij}; in fact it can be negative. However, the empirical rule of thumb for data from complex samples,

$$(\sigma^2_{ij}/n_i) + (\sigma^2_{i'j}/n_{i'}) < \text{Var}(p_{ij} - p_{i'j})$$
$$< \text{Var}(p_{ij}) + \text{Var}(p_{i'j}) \quad (3)$$

(Kish 1965, section 14.1), suggests that, in practice at least, $0 < Rohs_{ii'j} < Roh_{ij}$. As a result, design effects for differences of subclass proportions are smaller on the average than those for the subclass proportions themselves.

3. LINEAR CONTRASTS FOR SUBCLASS PROPORTIONS

Consider the design effects of linear contrasts of subclass proportions and subclass logits for a fixed level of the response category j (hence the subscript j is ignored). For both contrasts let $\{c_i\}$ be a set of fixed, known constants such that $\sum_i c_i = 0$. Furthermore, for the ith subclass let $\ell_i = \ell n\,[p_i\ (1-p_i)]$ be the observed sample logit. Then let

$$\hat{L} = \sum_i c_i p_i \quad (4)$$

be the linear contrast of subclass proportions and let

$$\hat{L}_g = \sum_i c_i \ell_i \quad (5)$$

be the linear contrast of sample logits under investigation.

Models for the design effects of (4) and (5) can be proposed under various sets of assumptions regarding parameters such as the subclass proportion design effects d^2_i and the subclass correlations, $R_{ii'}$. For this purpose, the variance of the linear contrast $\hat{L}$ in (4) can be expressed as

$$\text{Var}(\hat{L}) = \sum_i c^2_i\, \text{Var}(p_i) + 2\sum_{i<i'} c_i\, c_{i'}\, \text{Cov}(p_i, p_{i'})$$
$$= \sum_i c_i\, \sigma^2_i\, d^2_i / n_i$$
$$+ 2\sum_{i<i'} c_i c_{i'}\, (\bar{b}_{ii'} - 1)$$
$$\cdot\ Rohs_{ii'}\sigma_i\sigma_{i'}/\sqrt{(n_i n_{i'})} \ . \quad (6)$$

The first terms on the right hand side of (6) are a function of subclass homogeneity, Roh_i, through the design effects d^2_i, while the other terms are a function of the cross-homogeneities, $Rohs_{ii'}$. On the other hand, the simple random sampling variance of $\hat{L}$ can be expressed as

$$\text{Var}_{SRS}(\hat{L}) = \sum_i c^2_i\, \sigma^2_i/n_i \ , \quad (7)$$

since the covariances among subclass proportions are zero. Consequently, the design effect of $\hat{L}$ can be expressed in terms of (6) and (7) as

$$\text{Deff}(\hat{L}) = \text{Var}(\hat{L}) / \text{Var}_{SRS}(\hat{L}) \ . \quad (8)$$

For the linear contrast of subclass logits, $\hat{L}_g$, it is useful to examine first the design effect of the subclass logit, ℓ_i. Since ℓ_i is a nonlinear function of the subclass proportion p_i, its complex sample and simple random sample variances can be approximated by a first order Taylor series expansion as

$$\text{Var}(\ell_i) \doteq \text{Var}(p_i)/[\pi_i(1-\pi_i)]^2 \quad (9)$$

and

$$\text{Var}_{SRS}(\ell_i) \doteq \text{Var}_{SRS}(p_i)/[\pi_i(1-\pi_i)]^2 \ . \quad (10)$$

As a result the design effect of ℓ_i, the ratio of (9) to (10), is identical to d^2_i, the design for the subclass proportion p_i. Because of this identity, the design effect of $\hat{L}_g$ is similar, although not identical to $\text{Deff}(\hat{L})$. For this reason, the subsequent development focuses on $\text{Deff}(\hat{L})$ in examining models for design effects of linear contrasts.

4. MODELS FOR Deff($\hat{L}$)

Several different models for the design effects of linear contrasts of subclass proportions can be posited under various sets of assumptions about the relative values of d_i^2, σ_{ii}^2, $R_{ii'}$, Roh_i, and $Rohs_{ii'}$. These models can be utilized to explain the nature of Deff($\hat{L}$) and provide estimates directly from the model. The following subsections enumerate some specific models of particular interest to this investigation.

4.1 Ordinal Model

The empirical model of Kish and Frankel (1974) for design effects of analytic statistics and the empirical rule of thumb summarized in (3) suggest that Deff($\hat{L}$) is greater than 1 but less than an average of subclass proportion design effects denoted as $\bar{d}^2$. That is, the *ordinal model* for Deff($\hat{L}$) can be given as

$$1 < \text{Deff}(\hat{L}) < \bar{d}^2 \qquad (11)$$

The relationship in (11) describes the range of values expected for Deff($\hat{L}$), but it does not provide an estimate of Deff($\hat{L}$) directly. The upper limit, $\bar{d}^2$, is a conservative estimate of Deff($\hat{L}$), as suggested by Rao and Scott (1979), for instance, while unity is an uninteresting lower bound.

4.2 Proportional Reduction Model

Let $\bar{d}_w^2$ denote the weighted average of subclass proportion design effects

$$\bar{d}_w^2 = \textstyle\sum_i (c_i^2\sigma_i^2/n_i)\, d_i^2 \,/\, \sum_i (c_i^2\sigma_i^2/n_i) \ . \qquad (12)$$

Considering the relation in (11), define the proportional reduction of Deff($\hat{L}$) relative to the weighted average $\bar{d}_w^2$ to be

$$\begin{aligned}\text{Pref}(\hat{L}) &= [\text{Deff}(\hat{L}) - \bar{d}_w^2]/\bar{d}_w^2 \\ &= [2\textstyle\sum_{i<i'} c_i\, c_{i'}\, (\bar{b}_{ii'}-1)\text{Rohs}_{ii'}\,\sigma_i\,\sigma_{i'}/ \\ &\qquad \sqrt{(n_i n_{i'})}]/\textstyle\sum_i c_i^2\sigma_i^2\, d_i^2/n_i \ . \qquad (13)\end{aligned}$$

From this definition the *proportional reduction model* for Deff($\hat{L}$) can be expressed as

$$\text{Deff}(\hat{L}) = \bar{d}_w^2\,[1 + \text{Pref}(\hat{L})] \ . \qquad (14)$$

Obviously Deff($\hat{L}$) < $\bar{d}_w^2$ whenever Pref($\hat{L}$) < 0. But the nature of Pref($\hat{L}$) is not obvious from expression (13). A detailed discussion of the conditions for which Pref($\hat{L}$) < 0 is given in Lepkowski (1980) and is not elaborated here.

4.3 Attenuation Model

Consider the subclass difference $\hat{L} = p_i - p_{i'}$ which is a special case of $\hat{L}$ with $c_i = -c_{i'} = 1$ and $c_{i''} = 0$ for all $i'' \neq i, i'$. Suppose that the homogeneities for the two proportions are the same (i.e., $Roh_i = Roh_{i'} = Roh$), and let $Rohs_{ii'} = \alpha Roh$, for some constant α. Also, suppose that the population subclass proportions corresponding to p_i and $p_{i'}$, namely π_i and $\pi_{i'}$, are in the same general range of values so that the element variances σ_i^2 and $\sigma_{i'}^2$ are approximately equal, i.e., $\sigma_i^2 = \sigma_{i'}^2 = \sigma^2$. Let $\bar{b}_w = (W_i^2 n_i + W_{i'}^2 n_{i'}) / [(W_i^2 + W_{i'}^2) A]$ be a weighted average of the subclass sizes n_i and $n_{i'}$ where as before $W_i = c_i\sigma_i/\sqrt{n_i}$, $i=1,\ldots,s$. Within this context the *attenuation model* for the design effect of $\hat{L}$ can be given as

$$\text{Deff}(\hat{L}) = 1 + (\bar{b}_{ii'}-1)(1-\alpha)\text{Roh} . \qquad (15)$$

Under these assumptions the design effect for a subclass difference can be expressed in a form similar to expression (1) for means, but with an attenuated homogeneity for the subclass difference $(1-\alpha)$Roh. For Roh > 0 and $0 < \alpha < 1$, this effective homogeneity for the subclass difference will be smaller than Roh, and hence Deff($\hat{L}$) < d^2. The attenuation effect in (15) is a direct effect of the cross-homogeneity between p_i and $p_{i'}$.

Extending the attenuation model in (15) for subclass differences to an attenuation model for linear contrasts in general requires stronger assumptions, especially about subclass sizes. Such a model for linear contrasts of subclass proportions in general is given by

$$\text{Deff}(\hat{L}) = 1 + (\bar{b}_H-1)\,[(1-\alpha)\text{Roh}] \qquad (16)$$

where $\bar{b}_H = (\sum_i c_i^2)/[(\sum_i c_i^2 n_i^{-1})A]$ is a harmonic mean of subclass sizes across clusters. Expression (16) is useful for modeling Deff($\hat{L}$) provided values of α are available (values for Roh and $(\bar{b}_H-1)$ should be available already). The value of α for linear contrasts with more nonzero values of the constants c_i will be larger than α for subclass differences and therefore Deff($\hat{L}$) is less than the design effect for a subclass difference.

5. EMPIRICAL INVESTIGATIONS

Estimates of the design effects and various other related quantities were made for contingency tables from two survey datasets. The first dataset was obtained from the National Center for Health Statistics Health Examination Survey Cycle II (HES). The HES is a large multi-stage national probability sample of 7119 children ages 6-11 conducted from 1963 to 1965. The sample design is highly clustered ($\bar{b}$ = n/A = 178) and the homogeneities for the dental variables from that dataset considered here are large (some as large as 0.1-0.2). As a result the design effects for the HES are large (most larger than 2.0) facilitating the examination of small subclasses. The second dataset is the Fall, 1974 University of Michigan Survey Research Center Omnibus Survey (OMNI) of 1519 adults. The variables concern attitudes about social, economic, and political issues and have smaller homogeneities than those for the HES dataset. With small average cluster size ($\bar{b}$ = n/A = 15), the design effects for the OMNI are smaller than for the HES being mostly between 1.0 and 2.0 for the subclass proportions. The OMNI results contrast with the HES results and illustrate the limits and utility of the models suggested earlier.

Fifty 2^3 tables in which two of the three variables can be considered factors were generated from each dataset. Specifically five subclass configurations were created and applied to ten different dichotomous response variables to generate 50 (4 x 2) tables in each dataset. Two linear contrasts for interaction among the two factors, one of subclass proportions ($\hat{L} = p_1-p_2-p_3+p_4$) and the other of subclass logits ($\hat{L}_g = \ell_1-\ell_2-\ell_3+\ell_4$), were estimated together with their design effects, deff($\hat{L}$) and deff($\hat{L}_g$), respectively, for each table. In addition, several other related quantities were estimated for each table. In particular, the six possible subclass difference design effects were estimated for each table and averaged, the average being denoted as diff for a table. The average of the four subclass proportion design effects (denoted as $\bar{d}^2$) and the proportional reduction in design effect of Deff ($\hat{L}$) relative to $\bar{d}^2$ (denoted pref ($\hat{L}$)) were also computed. The subclass homogeneities were estimated as $roh_i = (d_i^2-1)/(\bar{b}_i-1)$; while the subclass cross-homogeneities were estimated as $rohs_{ii'} = cov(p_i,p_{i'})/(\bar{b}_{ii'}-1)\cdot[var_{SRS}(p_i)var_{SRS}(p_{i'})]^{\frac{1}{2}}$. Finally, the value of the attenuation factor α was estimated for the linear contrast $\hat{L}$ as

$$\hat{\alpha} = [\bar{d}^2 - deff(\hat{L})]/(\bar{d}^2 - 1) \quad . \qquad (17)$$

These results are examined first for a single HES table and then summarized by dependent variable for all 100 tables from the two datasets.

5.1 An Example from the HES Dataset

Consider the (4 x 2) table in Table 2. The question to be considered and summarized in terms of the linear contrasts $\hat{L}$ and $\hat{L}_g$ is the extent to which the proportion of children with no Periodontal Index varies within a family income group for different ages. That is, the contrasts measure the interaction of family income and age for Periodontal Index. As family income increases, the proportion of children with no Periodontal Index increases for both age groups. However, the age groups effect seems to be stronger for low family income children.

Table 2: Number and Proportion of Children with None and Some Periodontal Index by Age and Family Income Subclasses

Subclass		Periodontal Index		
Age	Family Income	None	Some	Total
6-8	< \$7000	1307 (0.61)	825 (0.39)	2132
	≥ \$7000	882 (0.70)	382 (0.30)	1264
9-11	< \$7000	1061 (0.53)	953 (0.47)	2014
	≥ \$7000	887 (0.67)	441 (0.33)	1328
Total		4137 (0.61)	2601 (0.39)	6738

The estimated linear contrasts of subclass proportions for interaction in this table is $\hat{L} = 0.056$ with estimated design effect deff($\hat{L}$) = 1.66, a 66% increase in variance due to the complexities of the sample design. The average subclass proportion design effect is $\bar{d}^2 = 5.70$, indicating that for this table there is a 70% proportional reduction in the design effect of $\hat{L}$ over an average subclass design effect.

The four homogeneities and six cross-homogeneities for the four subclass proportions were also estimated. They are given in the upper triangular matrix of homogeneities, $\underset{\sim}{H}$, as

$$\underset{\sim}{H} = \begin{bmatrix} 0.117 & 0.077 & 0.115 & 0.065 \\ & 0.075 & 0.090 & 0.072 \\ \text{(symmetric)} & & 0.140 & 0.078 \\ & & & 0.108 \end{bmatrix}$$

The homogeneities along the diagonal are fairly similar across the subclass with an average of 0.110. The cross-homogeneities are the six off-diagonal values with an average of 0.083. Although there is some variation in the cross-homogeneity values, they are the same order of magnitude and are approximately 75% the value of the homogeneity values.

Interpretation of the relationship between the homogeneities and cross-homogeneities is difficult without examining the proportions in each subclass cluster by cluster. Given the close relationship evident here, one interpretation is that children without any Periodontal Index tend to occur in the same cluster within a subclass to almost the same degree that they do among the different subclasses. Such a phenomenon is more likely a reflection of the nature of the clusters themselves as substantive units (in this instance, counties or groups of counties) than it is a substantive quality associated with Periodontal Index. Perhaps for smaller, more homogeneous clusters the intra-cluster homogeneities would be large while the cross-homogeneities would tend to be much smaller, if not negative in value. In such an instance, deff($\hat{L}$) might be larger than the average subclass design effect, $\bar{d}^2$.

The results for the linear contrast of subclass logits $\hat{L}_g$ are virtually identical. The estimated logit contrast is $\hat{L}_g = 0.213$ with estimated design effect of 1.63. Again, the proportional reduction is 70% over average subclass proportion design effect. In the subsequent developments, therefore, references to design effects for $\hat{L}$ also apply to those for $\hat{L}_g$.

5.2 Results from 100: (4 x 2) Tables

The full set of fifty 4 x 2 tables and results from the HES and OMNI datasets are described elsewhere (Lepkowski 1980). The results are summarized here by dependent variable and by dataset. Table 3 presents the average deff($\hat{L}$), average subclass proportion design effect, $\bar{d}^2$, average estimated pref($\hat{L}$), and average attenuation factor, $\hat{\alpha}$, by dependent variable for the one hundred tables examined.

Examining the average subclass proportion design effects $\bar{d}^2$ across the two surveys, it is apparent that larger design effects occur for the HES with an average of 3.91 over the 200 proportions in the 50 tables. However, there still remains considerable variability among design effects by dependent variable within surveys.

Similar remarks apply to the average design effects for $\hat{L}$. For the HES tables, overall there is an average deff($\hat{L}$) = 1.17, but average values for dependent variables (5 tables each) range from 0.80 to 1.54. The OMNI tables show an average deff($\hat{L}$) = 0.99 with a range of 0.75 to 1.19. Since the subclass sizes of the subclass configurations are identical across dependent variables, the variability in average design effects can be attributed to different levels of homogeneity among the dependent variables.

Table 3: Average deff($\hat{L}$), $\bar{d}^2$, pref ($\hat{L}$), and $\hat{\alpha}$ by Dependent Variable, HES and OMNI 4 x 2 Tables

Variable	deff($\hat{L}$)	$\bar{d}^2$	pref($\hat{L}$)	$\hat{\alpha}$
	a. HES Tables			
Periodontal Index	1.32	6.42	-0.80	0.95
DEF Score	1.39	1.81	-0.23	0.63
Gingivitis	1.26	6.69	-0.82	0.97
Oral Hygiene	1.07	7.53	-0.86	0.01
Malocclusion	1.10	2.35	-0.54	0.94
Last Dental Visit	1.17	5.66	-0.79	0.96
Overbite	1.03	1.66	-0.39	1.14
Decayed Primary	0.80	1.67	-0.52	1.30
Extraction	1.54	3.43	-0.56	0.78
Filled Primary	1.00	1.89	-0.45	1.03
Total	1.17	3.91	-0.60	0.97
	b. OMNI Tables			
Fuel Shortages	1.07	1.02	0.05	1.86
Marijuana Use	0.75	1.03	-0.27	6.71
Anti-democracy	1.02	0.91	0.12	-1.10
Physician Attitudes	0.93	0.91	0.02	-0.65
Unions	0.83	1.39	-0.41	1.53
Communist Books	1.12	1.22	-0.09	0.81
Male/Female Politicians	1.06	0.98	0.08	-4.27
Atheists	1.19	1.34	-0.12	0.67
Liberal/Conservative	0.97	1.13	-0.14	1.76
Vietnam War	1.00	1.11	-0.10	2.02
Total	0.99	1.10	-0.09	0.93

Nonetheless, in both surveys deff($\hat{L}$) < $\bar{d}^2$ on the average. The size of pref($\hat{L}$) is consistent with this observation, although there is considerable variability not only across surveys but also across dependent variables. While the HES survey has an average of 60% proportional reduction from $\bar{d}^2$ to deff($\hat{L}$), the OMNI tables have only a 9% average reduction.

The proportional reduction is related to the size of the initial average subclass design effect, $\bar{d}^2$, for each table as illustrated in Figure A for the HES tables. As the average subclass proportion design effect increases, so does the proportional reduction in deff($\hat{L}$) from $\bar{d}^2$. A similar although somewhat weaker relationship can be demonstrated for the OMNI tables. Since the average subclass design effect depends on the homogeneities, roh_i, and the proportional reduction depends on the cross-homogeneities, $rohs_{ii'}$, there is a likely relationship between the homogeneities and cross-homogeneities.

The average estimated subclass proportion homogeneity of the (4)(5)=20 subclass proportions in the five tables for Periodontal Index (PI) is 0.131. The average estimated cross-homogeneity [(6)(5)=30 values] is 0.083 for those five tables. The average estimated $rohs_{ii'}$ value for PI is thus 0.63 times the average roh_i value. The relationship between roh_i and $rohs_{ii'}$ is fairly consistent across dependent variables with an overall $rohs_{ii'}/roh_i$ ratio of 0.64 for the 400 roh_i and 600 $rohs_{ii'}$ values. The relationship is illustrated for the HES tables in Figure B where the fifty average roh_i and $roh_{ii'}$ values from each table are plotted. Although there is some variability present, there is a strong linear relationship indicated with $rohs_{ii'} < roh_i$ consistently across the tables. The results are not presented for the OMNI tables because the relationship is not as clearly demonstrated by dependent variables or by individual tables, although the overall results are consistent with those of the HES data.

Figure A: pref($\hat{L}$) by $\bar{d}^2$ for Fifty HES 4 x 2 Tables

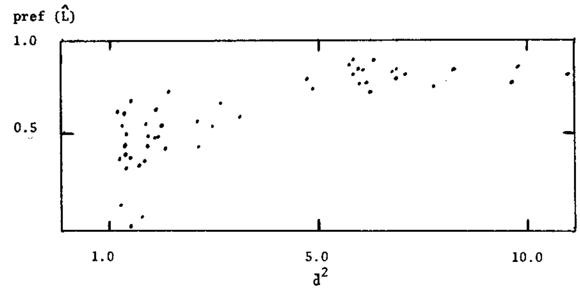
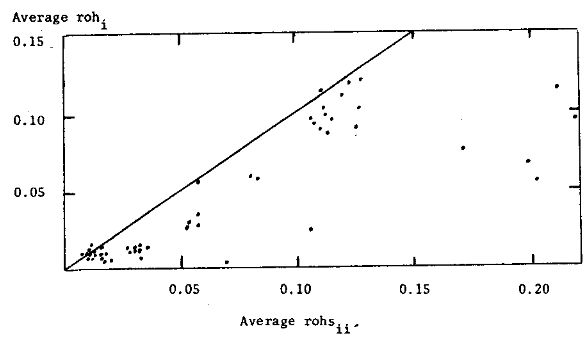


Figure B: Average roh_i by average $rohs_{ii'}$ for Fifty HES 4 x 2 Tables



Finally, Table 3 also presents the average estimated attenuation factor $\hat{\alpha}$ for the fifty deff($\hat{L}$) values from the two surveys as estimated by (17). There is considerable variability among the average $\hat{\alpha}$ across dependent variables. Overall the average $\hat{\alpha}$ for the HES is 0.97 and for the OMNI 0.93. Thus the effective homogeneity for the HES and OMNI are (0.03)roh and (0.07)roh, respectively. Since the OMNI has small average cluster sizes and homogeneities, the effective homogeneity for the linear contrast $\hat{L}$ is negligible and Deff($\hat{L}$) is essentially 1. For the HES, however, the average cluster sizes and homogeneities are larger; the effective homogeneity for $\hat{L}$, although small, yields an average Deff($\hat{L}$) across fifty tables of 1.17.

6. DISCUSSION

The empirical results demonstrate the validity and the utility of the models discussed previously. The ordinal model in (11) is correct for nearly every dependent variable across both survey datasets. The model can be extended to include the design effect for subclass differences diff, as $1 < \text{deff}(\hat{L}) < \text{diff} < \bar{d}^2$. Even this refinement in the ordinal model is not sufficient for estimation of Deff($\hat{L}$) in instances when $\bar{d}^2$ is large.

The proportional reduction model in (13) and (14) would provide estimates of Deff($\hat{L}$)using the empirical values of pref($\hat{L}$) and $\bar{d}^2$ given for HES and OMNI. However, since pref($\hat{L}$) and $\bar{d}^2$ are related, it is difficult to predict a useful value for pref($\hat{L}$) for other dependent variables than those given here.

The attenuation model in (15) and (16), however, offers estimates of Deff($\hat{L}$) provided values of α are available from the survey. It is based on an intuitively appealing model for design effects of proportions which demonstrates the nature of the relationship between Deff($\hat{L}$) and cross-homogeneity through the attenuation factor α. Thus, for a particular response variable for which roh_i is known or can be estimated, and for a particular set of subclass sizes, values of Deff($\hat{L}$) and Deff($\hat{L}_g$) can be estimated for subclass differences and interaction contrasts as given here. Additional empirical results on other survey datasets should indicate whether these values of $\hat{\alpha}$ can be applied to other surveys.

Adjustment of an ordinary chi-square statistic for the effects of the complex sample design using design effects estimated from the attenuation model is a straightforward division of the ordinary chi-square statistic by the design effect to obtain the adjusted chi-square statistic. The effect would be the same for a linear or a log-linear modeling situation in which the contrasts $\hat{L}$ or $\hat{L}_g$ could be used to test the hypothesis of no second-order interaction for the contingency table. It is important to note that for health survey data, where large average cluster sizes and large homogeneities are to be expected, the adjustment will not necessarily be negligible. Further investigation of the problem is warranted on the basis of these findings.

7. SUMMARY

Three models for the design effects of linear contrasts of subclass proportions (Deff($\hat{L}$)) and subclass logits (Deff ($\hat{L}_g$)) have been suggested analytically and investigated empirically. The ordinal model is an adequate description of the general nature of Deff($\hat{L}$) and Deff($\hat{L}_g$) but a useful estimate based on the model cannot be obtained. The proportional reduction model is more useful for estimating the design effects provided estimates of certain parameters are available from the survey. However, those parameter estimates are usually not available. The attentuation model also depends on the estimation of a parameter not usually available for all surveys, the attenuation factor. However, the model is intuitively appealing and the attenuation factor may be portable across survey datasets facilitating the application of the model.

REFERENCES

Bhapkar, V.P., and Koch, Gary G. (1968), "Hypothesis of 'No Interaction' in Multidimensional Contingency Tables," Technometrics, 10, 107-123.

Fellegi, I.P. (1978), "Approximate Tests of Independence and Goodness of Fit Based on Stratified Multi-Stage Samples," Survey Methodology 4 (1): 29-56.

Kish, Leslie (1962), "A Measure of Homogeneity in Areal Units," Bulletin of the International Statistical Institute, 34(4), 201-210.

_____ (1965), Survey Sampling, New York: J. W. Wiley and Sons.

_____, and Frankel, Martin R. (1974), "Inference from Complex Samples," Journal of the Royal Statistical Society, Series B, 36, 1-38.

_____, Groves, Robert M., and Krotki, Karol J. (1976), Sampling Errors for Fertility Surveys," World Fertility Survey Occasional Paper Number 17, Voorburg, Netherlands: International Statistical Institute.

Lepkowski, James M. (1980), "Design Effects for Multivariate Categorical Interactions," unpublished Ph.D. Dissertation, Department of Biostatistics, University of Michigan.

Rao, J.N.K., and Scott, Alistair (1979), "Chi-Square Tests for Analysis of Categorical Data from Complex Surveys," Proceedings of the Survey Research Methods Section of the American Statistical Association, 58-66.