Centre for Statistical and Survey Methodology

The University of Wollongong

Working Paper

24-10

The impact of complex survey design on prevalence estimates of intakes of food groups in the Australian National Children's Nutrition and Physical Activity Survey

Sandy Burden, Yasmine Probst, David Steel and Linda Tapsell

Centre for Statistical & Survey Methodology, University of Wollongong, Wollongong NSW 2522. Phone +61 2 4221 5435, Fax +61 2 4221 4845. Email: anica@uow.edu.au

The impact of complex survey design on prevalence estimates of intakes of food groups in the Australian National Children's Nutrition and Physical Activity Survey

Sandy Burden (corresponding author)[1], Yasmine Probst[2], David Steel[1], Linda Tapsell[2]

Abstract

Objective: To assess the impact of the complex survey design used in the 2007 Australian National Children's Nutrition and Physical Activity Survey (ANCNPAS) on prevalence estimates for intakes of groups of foods in the population of children.

Design: The impacts on prevalence estimates were determined by calculating design effects for values for food group consumption. The implications of ignoring elements of the sample design including stratification, clustering and weighting were discussed.

Setting: The 2007 ANCNPAS used a complex sample design involving stratification, a high degree of clustering and estimation weights.

Subjects: Australian children aged 2-16 years.

Results: Design effects ranging from <1 to 5 were found for the values for means and proportions of food groups consumed. When survey weights were ignored, prevalence estimates were also biased.

Conclusions: Ignoring complex survey design used in the ANCNPAS could result in underestimating the width of confidence intervals, higher mean square errors and biased estimators. The magnitude of these effects depends on both the parameter under consideration and the chosen estimator.

Introduction

The degree of complexity in survey design depends on the nature of the research question, just as the method of data collection influences the choice of sampling technique. For example, straightforward telephone interviews allow relatively simple sample designs, but comprehensive nutrition surveys tend to be longer and more complex. Moreover, collection of reliable anthropometric data involves face-to-face interviewing. In this case, nutrition and physical activity surveys often use complicated sample designs, involving stratified multistage sampling techniques. To improve efficiency and reduce costs, these designs can include the use of stratification,

[1] Centre for Statistical and Survey Methodology, University of Wollongong, Northfields Avenue, Wollongong, NSW 2522, Australia
[2] Smart Foods Centre, University of Wollongong, Northfields Avenue, Wollongong, NSW 2522, Australia

clustering and unequal probabilities of selection for different individuals. The resulting sample is not spread evenly throughout the population, but occurs in groups or clusters.

The 2007 Australian National Children's Nutrition and Physical Activity Survey (ANCNPAS07) (1) was undertaken to obtain food, nutrient, physical activity and anthropometric data on a national sample of children aged 2-16 years. The purpose of the survey was to enable food, beverage, supplement, and nutrient intakes and physical activity levels among children to be assessed against relevant national guidelines.

The survey was conducted using a sampling scheme stratified by state/territory and by capital city statistical division/rest of state into 13 strata. The number of children included from each state was proportional to the population of children in that state or territory. To collect physical activity data, anthropometric measurements and a 24 hour diet history, an initial face-to-face interview was used. To facilitate the face-to-face interviewing and to help meet budget and time restrictions, the sample was obtained from 246 postcodes clustered in 54 locations which were effectively Primary Sampling Units (PSUs). Initial selection and contact was made using Random Digit Dialing (RDD) and data was collected using Computer Assisted Personal Interviewing (CAPI) and subsequent Computer Assisted Telephone Interviewing (CATI). One child was selected per household leading to 4837 selected children and complete data for 4487 children. Within selected clusters, the probability of selection of a child depended on their location (stratum), age, gender and household composition. To account for the non-proportional sampling, weights were created based on age (divided into 4 groups), gender and stratum. A single weight, called the initial weight and denoted w_1, was produced for each child in the survey on the basis of a sample size of 4837.

For these data, the use of clustering, unequal selection probabilities, stratification and sample weighting lead to estimates having a sampling variance different from that which would have been obtained using a Simple Random Sample (SRS). An SRS gives each possible sample the same chance of selection and means that the sample is spread approximately evenly through the population. For an SRS the calculation of estimates and associated standard errors is relatively straightforward. Standard methods of statistical analysis assume that an SRS has been obtained.

If the analysis of a nutrition survey ignores the complex design, the results will be methodologically unsound and subject to serious dispute. Typically confidence intervals will be too small, leading to inflation of type I error rates. That is, statistical significance is found when there is no real effect. The problem is not solved merely by using the sample weights, which account for differences in selection probabilities, although this is often incorrectly assumed. In fact this view is implicitly encouraged if the survey data are released with, for example, no cluster information. Even when an analysis uses sample weights and the contribution of the clustering to the overall design effect is low, use of standard analysis will not reflect the impact of the weights on variances. The use of sampling weights and the impact of complex sampling methods on survey analysis has received considerable attention over the past two decades, see (2) (3) (4) and (5).

The importance of properly accounting for sampling weights and the sample design is strongly emphasized in well established surveys in the USA, for example the National Health and Nutrition Examination Survey (NHANES). Information on the NHANES website (6) states “For NHANES datasets, the use of sampling weights and sample design variables is recommended for all analyses because the sample design is a clustered design and incorporates differential probabilities of selection. If you fail to account for the sampling parameters, you may obtain biased estimates and overstate significance levels.” Moreover the National Centre for Health Statistics (NCHS) Analytic and Reporting Guidelines state that "Sample weights and the stratification and clustering of the design must be incorporated into an analysis to get proper estimates and standard errors of estimates” and that proper variance estimation procedures be used (7 p. 7).

Complex sample design also needs to be taken into account in meta analysis in which the results of two or more surveys are combined or survey data is combined with data from clinical trials. This may be done, for example, in establishing an evidence base for the effects of food consumption patterns on health. In undertaking meta analysis the results or data from each study is weighted according to its quality and this leads to the use of effective sample size, which depends on the design. Thus it is important that the design features are considered if appropriate conclusions regarding food intake patterns are to be made. The aim of the study reported here was to assess the impact of complex survey design used in the ANCNPAS07 on prevalence estimates for intakes of groups of foods in the population of Australian children.

Methods

The study used the concept of design effects to quantify the effect of the sample design on prevalence estimates. For each estimate, design effects were used to measure the impact of stratification, clustering, unequal inclusion probabilities and other features of the sampling used. The design effect (deff), is the ratio of the sampling variance obtained using a complex survey design relative to the variance that would have been obtained from a simple random sample without replacement (SRSWOR) with the same expected sample size (8). The deff for a parameter θ is calculated using the relationship $deff = \frac{\hat{V}(\hat{\theta})}{\hat{V}_{srswor}(\hat{\theta}_{srs})}$ where $\hat{V}(\hat{\theta})$ is the design based estimate of the variance for the parameter estimate $\hat{\theta}$ from a complex survey of size n , and $\hat{V}_{srswor}(\hat{\theta}_{srs})$ is the variance estimate of the parameter θ estimated from a similar hypothetical survey using SRSWOR and a sample size of n .

A design effect greater than one increases the width of confidence intervals, reduces the amount of disaggregation that is possible and reduces the power of analyses that are properly carried out. This limits the strength and value of the results. For example, suppose a survey has been designed using standard methods which assume a SRS to give a power of detecting important effects of 80%. With a deff of 1.5 the power reduces to 65%; for a deff of 2 it becomes 45%; and for deff of 4 it is 35%. Tests of statistical significance are also affected and a deff of 4 increases the conventional 5% false positive rate used in hypothesis testing to 33%.

A deff can also be expressed as the effective sample size, $n_{eff} = \frac{n}{deff}$. For example, a sample of 4000 respondents has an effective sample size of 1000 if the deff is 4. So selecting several respondents within a cluster will be less efficient in terms of variance than using SRS. This has substantial implications for the way in which the survey data may be acceptable to the wider community and used in policy development.

For ANCNPAS07, the use of RDD led to the inclusion of individuals from 481 postcodes, due to both the overlap between telephone number prefixes and postcodes and telephone number portability. In the available data files, locations for the sample were recoded after selection in such a way that participants in a close geographic proximity, based on their postcodes, were given the same location. To enable variance calculations, locations with single observations were grouped within strata, reducing the total number of locations from 210 to 194. Locations had an average size of 24.8 responding children (for the CAPI), varying from 2 to 177 (standard deviation (SD) of 29). The publically available dataset initially included only these locations, along with state and region variables, not the original 54 PSUs (which have subsequently been released), so deffs were calculated using both the locations and the 54 PSUs. One additional PSU was also created to enable variance calculations, resulting in 55 clusters. We found the responding sample to be highly clustered with an average of 89 children per cluster, varying from 19 to 161 (SD of 28.6).

Weights were created to account for non-proportional sampling based on age (divided into 4 groups), gender and location. A single weight was produced for each child in the survey on the basis of a sample size of $n = 4837$. The initial weight for child $i, i = 1, \ldots, n$ equals $w_{1i} = N_h / n_h$, where n_h is the number of respondents in stratum h and N_h is the corresponding population. Separate weights were not included for respondents who did not complete all components of the survey and household size and family structure were not included in the weights. The probability of selection was therefore only partially accounted for in the weights. Furthermore, as there was only complete nutrient data from the CAPI for 4826 of the 4837 participants, the population totals using the weights did not correspond to those for Australia available from the 2006 Census.

Due to the limitations in the weighting process, a final weight (w_2) was created by adjusting w_1 to fit population benchmarks and accounting for the probability of selection of each child in a household. Assuming that all children in a household had an equal probability of selection, the probability of selection for each subject was calculated as π_i=1/(no. children aged 2-16 in household). The total effective sample size in stratum h using the final weight was $\hat{n}_h = \sum_{i \in h} \pi_i^{-1}$ and using initial weights was n_h. The final weight for each subject in stratum h was obtained using $w_{2i} = w_{1i} \pi_i^{-1} \frac{n_h}{\hat{n}_h}$.

Sample weights are used to produce an estimate which is less biased than its unweighted counterpart. However, the increased accuracy must be balanced against the increased design effect (9). One approach to choosing the most efficient estimator is to examine the mean square error ($MSE = bias^2 + Var$) for each parameter (10 p. 176). For multiple variables, the relative mean square error (RMSE) can be used. Assuming the final weights produced unbiased estimates, the

RMSE for the estimate of a mean ($\overline{Y}$) is $RMSE_{w_2} = {[V(\overline{Y}_{w_2})]}/{\overline{Y}_{w_2}}$. Where $V(\overline{Y}_{w_2})$ is the variance of the estimate, calculated using the final weight w_2. Otherwise, $RMSE_{w_1} = {[(\overline{Y}_{w_1} - \overline{Y}_{w_2})^2 + V(\overline{Y}_{w_1})]}/{\overline{Y}_{w_1}}$. The estimator with the smallest RMSE is preferred.

The coefficient of variation of the weights is given by $C_w = \frac{s_w}{\overline{w}}$, where s_w is the standard deviation of the weights and $\overline{w}$ is the mean. It measures the increased variance of the estimate due to the use of weights. When selection probabilities are not correlated with a variable, the deff due to weighting is given by $1 + C_w^2$ (8) (11). When correlation is present, approximations can be made (12).

Under some mild assumptions the contribution of sample clustering for the estimation of prevalence of a condition or risk factor is reflected in the relationship $deff = 1 + (\overline{n} - 1)\delta$, where $\overline{n}$ is the average number of respondents per cluster and δ is a measure of the within cluster homogeneity or intraclass correlation (ICC). Values of δ around 0.05 are common, which with $\overline{n} = 81$ gives *deff*=5 and with $\overline{n} = 11$ gives *deff*=1.5. Hence the more clustered the design the higher the deff. If the size of the clusters varies considerably, more complicated formulas apply. For applications where the clustering and weighting effects are multiplicative, the deff is given by $deff = [1 + (\overline{n} - 1)\delta] \cdot (1 + C_w^2)$ (11).

Statistical Analysis

Deffs were estimated for the prevalence estimates of food consumption for ANCNPAS07 using STATA (StataCorp. 2007. Stata Statistical Software: Release 10. College Station, TX: StataCorp LP). The variables chosen for analysis were the 120 three digit sub-major groups used in the food categorization. The parameters chosen for analysis were mean consumption of each sub-major food group in grams and the proportion of the population consuming each food group. The CAPI 24 hour recall diet history was used for all analyses.

Estimates and estimates of sampling variances were produced under a number of options for treating the weights and sample design features including:

1. Unweighted analysis assuming SRS
2. Weighted analysis assuming SRS
3. Weighted analysis incorporating stratification (13 strata) and clustering using the 210 locations in the data file.
4. Weighted analysis incorporating stratification (13 strata) and clustering using 55 PSUs.

Analysis under 1 was the naive analysis. The estimates and estimated variances were compared with analyses under 4, which properly reflected the weighting and complex design. Option 2 accounts for the weighting but ignores the sample design and option 3 uses incorrect clusters.

Results

Of the 120 sub-major food groups, 36 had less than 55 non-zero observations, fewer than or equal to the number of clusters. For these groups, the observed deffs averaged only 1.11 compared with 2.35 for the other groups (Table 1) and 44% of the groups had a final deff of less than one (Figure 1). The lower deffs occurred because the average number of observations per cluster was one or less, so there was effectively no clustering. The results for this group are presented separately, and for notational convenience the mean and proportion estimators are denoted mean>55; mean≤55; prop>55 and prop≤55.

The effect of complex survey design

When consumption of 3 digit food groups was estimated using the correct design, the average deff was 1.1(1.4) for mean<=55(prop<=55) (Table 1) and for mean>55(prop>55) was 2.1(2.3). The effect of the survey design was highly variable (Figure 1), with deffs ranging from 0.3 to 5.1 for different food groups and estimators. These results are important for the analysis of nutrition surveys because an increase in the deff affects the significance of the results. For example, a deff of 2 increases the width of the confidence interval of an estimator by 1.4 and a deff of 4 increases it by 2.0.

A common error is to regard the estimate with the lowest estimated standard error as the best. However, the standard error is only correct when all aspects of the weighting and design are accounted for. For mean>55 (prop>55), 84 (89) groups had greater than 55 observations. Of these, the 45(52) groups with deff > 2.0 are listed in Table 3 (Table 4). Most estimates were biased when a SRS was assumed and in all cases the confidence intervals were substantially wider when the correct sample design was used for estimation.

In the following sections, the impacts of elements of the design are considered separately.

Weighting

The design and final weights both had a similar right skewed distribution with the same mean. The final weights had a higher standard deviation and a wider range due to the inclusion of the additional weighting component (Table 2).

The theoretical deff due to the initial weights was $1 + C^2_{W_1}$ = 1.33 and the final weights was $1 + C^2_{W_2}$ = 1.63. The 0.3 increase equals the increase in the average observed deff for groups with greater than 55 observations (Figure 3 and Figure 4 and Table 1) as neither set of weights is highly correlated with the response variables. Most food groups were similarly affected, with deffs generally below 2 and slightly higher for the final weights. The exceptions were two groups with deffs >3 and four groups with deffs between 2 and 3.

Weighted estimation increased the deff, but it also reduced the bias of estimators in the survey. Assuming the estimate obtained using the final weights was unbiased, when an SRS was incorrectly assumed the percentage bias for the mean>55 [prop>55] estimator was between (-15%, 16%)[(-10%, 22%)] for 95% of groups. Using the initial weights the percentage bias was (-9%, 11%)[(-4%,9%)] (Figure 2). For mean≤55 [prop≤55], the percentage bias had a much wider range for both an SRS (-23%, 221%)[(-43%, 235%)] and the initial weights (-23%, 83%)[(-39%, 91%)] and was generally positive.

The RMSE of the estimates assuming SRS, initial weights and final weights all followed a similar distribution (Figure 5). On average, the final estimates had the lowest average RMSE (0.05), followed by the initial weighted estimates (0.07) and SRS estimates (0.12).

Stratification

For this survey stratification had very little impact on the deffs. All of the estimates showed no change in the deffs when stratification was included (Figure 3 and Figure 4 and Table 1).

Clustering

Clustering had a much greater effect, increasing the average deff for mean>55 (prop>55) by 0.4 (0.7) (Table 1). The change in deff due to clustering was highly variable for different parameters and different estimators of the same parameter (Figure 3 and Figure 4). It depended on the pattern of responses for the variable and the location and size of clusters. When the correct clusters were not used – for example if location is incorrectly treated as the sampling unit – the variance was underestimated, decreasing the average deff for mean>55 (prop>55) by 0.14 (0.27) since the full cluster effect and the variation of locations within clusters was effectively ignored.

The deff associated with clustering was also estimated using the relationship $deff = 1 + (\bar{n} - 1)\delta$. Values of δ calculated from the data varied between 0 and 0.039. For many food groups, assuming $\bar{n} = {}^{4826}/_{54} = 89.4$ overestimated the deff when groups were not consumed by all respondents. To predict the deff for a food group, the proportion of the population consuming the food(s) must also be estimated to obtain a measure of $\bar{n}$ for the food group (for example using the values of N in Table 3 and Table 4).

Discussion

For the groups with less than 55 observations, most of the design effect arises due to weighting, with no appreciable change due to stratification and clustering. This occurs because the average number of non-zero observations per cluster is close to one. There is effectively no clustering, so the deff is also close to one (13). The variability around one is most probably due to estimation of the sample variance. As the effective degrees of freedom may be significantly less than the nominal degrees of freedom (= number of sampled PSUs – number of strata = 41) (10) the stability of the variance estimator may be questionable. Sampling error can then cause the observed deff to vary randomly above and below one (14). However, further investigation of this possibility is beyond the scope of this paper.

Considering the food groups with greater than 55 observations per cluster, the effect of weighting is generally similar to that expected by the theory. Deffs due to weighting depend on the coefficient of variation of the weights and the correlation between the weights and the survey variables (9). The inclusion of a component of weight due to the number of children per household increased the design effect by a relatively small and consistent amount for most food groups in the survey. Those with larger changes were food groups which have a different consumption pattern for households with small or large numbers of children. For example the groups '*dishes other than confectionery where sugar is the main ingredient*' and '*jam and lemon spreads, chocolate spreads, sauces*'.

The results illustrate that choosing an appropriate estimator entails a tradeoff between bias and deff.

Ignoring sample weights, or using the wrong weight can result in a biased estimator. However, the use of weights may increase the deff which affects the potential significance of the results. The effect of weighting also depends on the coefficient of variation of the weights and the correlation between the weights and the survey variables (9). For the 3 digit food groups, if the clustering effect is ignored, the relatively small deff from weighting means that in some cases an un-weighted estimator has lower RMSE and may be preferred. However, it is not always possible to quantify all sources of bias. An alternative to a weighted estimator is to include survey design variables in a model for the variable of interest with un-weighted regression estimation (10).

The effect of stratification is generally to decrease the deff, because stratification removes one component of variance from the estimator. However, unless there is a large difference between strata the impact on the deff is small, as it is in this case. The effect of clustering is determined by the number of sample units per cluster and the ICC within each cluster. As the ICC varies between 0 and 0.04 for different variables, the deff due to clustering is highly variable and estimator specific.

Overall, deffs can be large and they depend on both the chosen parameter and estimator. The main outputs from a survey frequently consist of prevalence estimates, such a means, proportions and population totals and the deff for each of these will be different (15), (13). Similarly, complex sample design also has an impact on the estimates of parameters of statistical analysis, such as regression parameters from a linear or logistic regression and associated odds ratios, but they differ from the deff on prevalence estimates and are not considered here. Furthermore, care must be taken when accounting for ratios, post-stratification and how deff is calculated both for the population and for estimates for subgroups.

Deffs arise due to the interaction of the sample design and the population structure, so they will be low for universal items which do not vary geographically or by cluster such as the groups '*milk*' or '*savoury biscuits*'. They are higher when consumption varies by, for example, geographic location, age and/or gender. Deffs arise through both unequal selection probabilities and other elements of the sample design such as stratification and clustering. Hence, to obtain accurate standard errors, all of these elements need to be taken into consideration during analysis of complex survey data.

Many statistical computing packages use a designed-based or pseudo likelihood approach which uses sample weights to estimate what would have resulted had population data been available. The complex design is then accounted for in the variance estimation. To properly account for the design, the data file needs to include stratum and cluster indicator variables.

Knowing the approximate magnitude of design effects for a particular estimator is useful when designing future surveys. The variables chosen here were particular food groups and the estimators were means and proportions, but design effects can be calculated for any variable, including macro or micro nutrient intakes for the population or for sub-populations. Also for any estimator, including means, totals, proportions or more complex estimators such as regression estimators – for which cluster effects are often lower (13). Being able to estimate the design effects allows required sample sizes for future surveys to be estimated. The use of weighting and choice of weights and also the degree of clustering and stratification in the survey can be tailored to achieve the desired standard error or power. Developing an appropriate design for a nutrition survey is difficult because there is considerable uncertainty about the values of relevant population characteristics such as δ. Also these parameters vary between variables and the type of analysis. A high degree of clustering can lead to large deffs, but reducing the clustering when it is not necessary increases costs.

References

1. **Commonwealth Department of Health and Ageing, et al.** *The 2007 National Children's Nutrition and Physical Activity Survey. [Computer file].* Canberra : Australian Social Science Data Archive, The Australian National University, 2009.

2. **Clark, R and Steel, DG.** The effect of using household as a sampling unit. *International Statistical Review.* 2002, Vol. 70, 2, pp. 289-314.

3. **Skinner, CJ, Holt, D and Smith, TMF.** *Analysis of Complex Surveys.* Chichester : Wiley and Sons, 1989.

4. **Chambers, RL and Skinner, CJ.** *Analysis of Survey Data.* Chichester : Wiley and Sons, 2003.

5. **Pfeffermann, D.** The use of sampling weights for survey data analysis. *Statistical methods in Medical Research.* 1996, Vol. 5, pp. 239-261.

6. **National Center for Health Statistics.** Continuous NHANES Web Tutorial: Sample Design. [Online] [Cited: August 16, 2010.] http://www.cdc.gov/nchs/tutorials/Nhanes/SurveyDesign/SampleDesign/intro.htm.

7. **National Centre for Health Statistics.** Analytic and Reporting Guidelines for the National Health and Nutrition Examination Survey. [Online] [Cited: August 16, 2010.] http://www.cdc.gov/nchs/data/nhanes/nhanes_03_04/nhanes_analytic_guidelines_dec_2005.pdf.

8. **Kish, L.** *Survey Sampling.* New York : John Wiley and Sons, 1965.

9. —. Weighting for unequal Pi. *Journal of Official Statistics.* 1992, Vol. 8, 2, pp. 183-200.

10. **Korn, EL and Graubard, BI.** *Analysis of Health Surveys.* New York : Wiley and Sons, 1999.

11. **Gabler, S, Haeder, S and Lahiri, P.** A model based justificationof Kish's formula for design effects for weighting and clustering. *Survey Methodology.* 1999, Vol. 25, 1, pp. 105-106.

12. **Spencer, BD.** An approximate design effect for unequal weighting when measurements may correlated with selection probabilities. *Survey Methodology.* 2000, Vol. 26, 2, pp. 137-138.

13. **Gambino, J.** Design effect caveats. *The American Statistician.* 2009, Vol. 63, 2, pp. 141-146.

14. **Curtin, LR, et al.** Estimation and analytic issues for rare events in NHANES. *Proceedings of the Survey Research Methods Section, American Statistical Association.* 2006.

15. **Park, I and Lee, H.** Design effects for the weighted mean and total estimators under complex survey sampling. *Survey Methodology.* 2004, Vol. 30, 2, pp. 183-193.

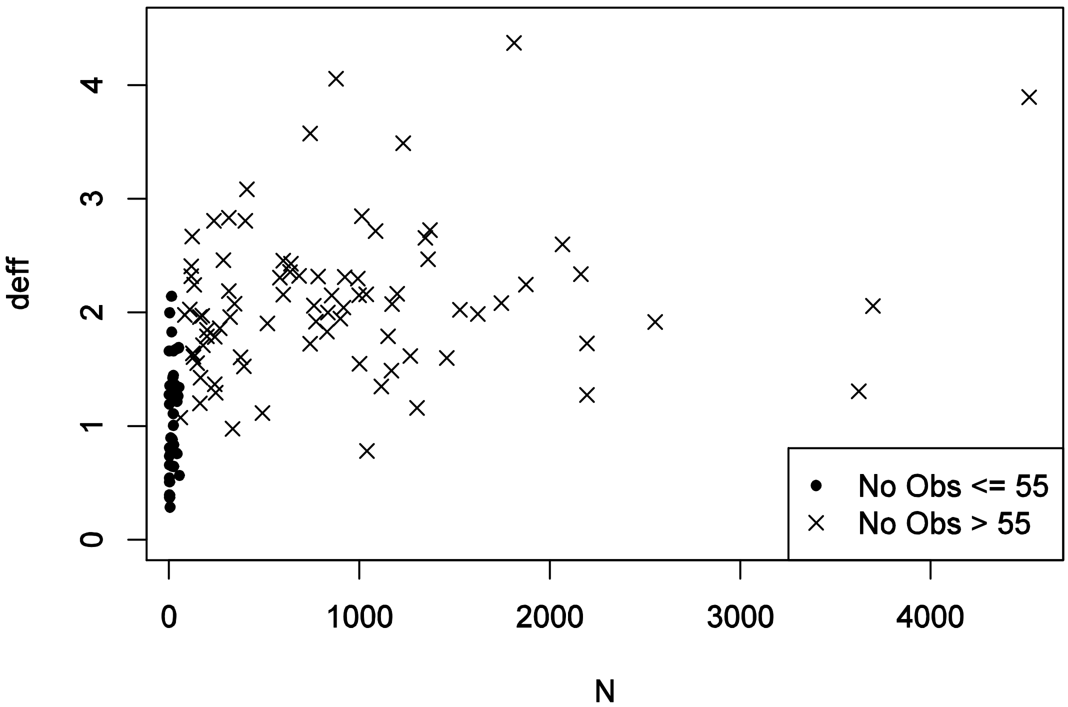

Figure 1 Final design effects for mean consumption of the three digit food groups by number of observations.

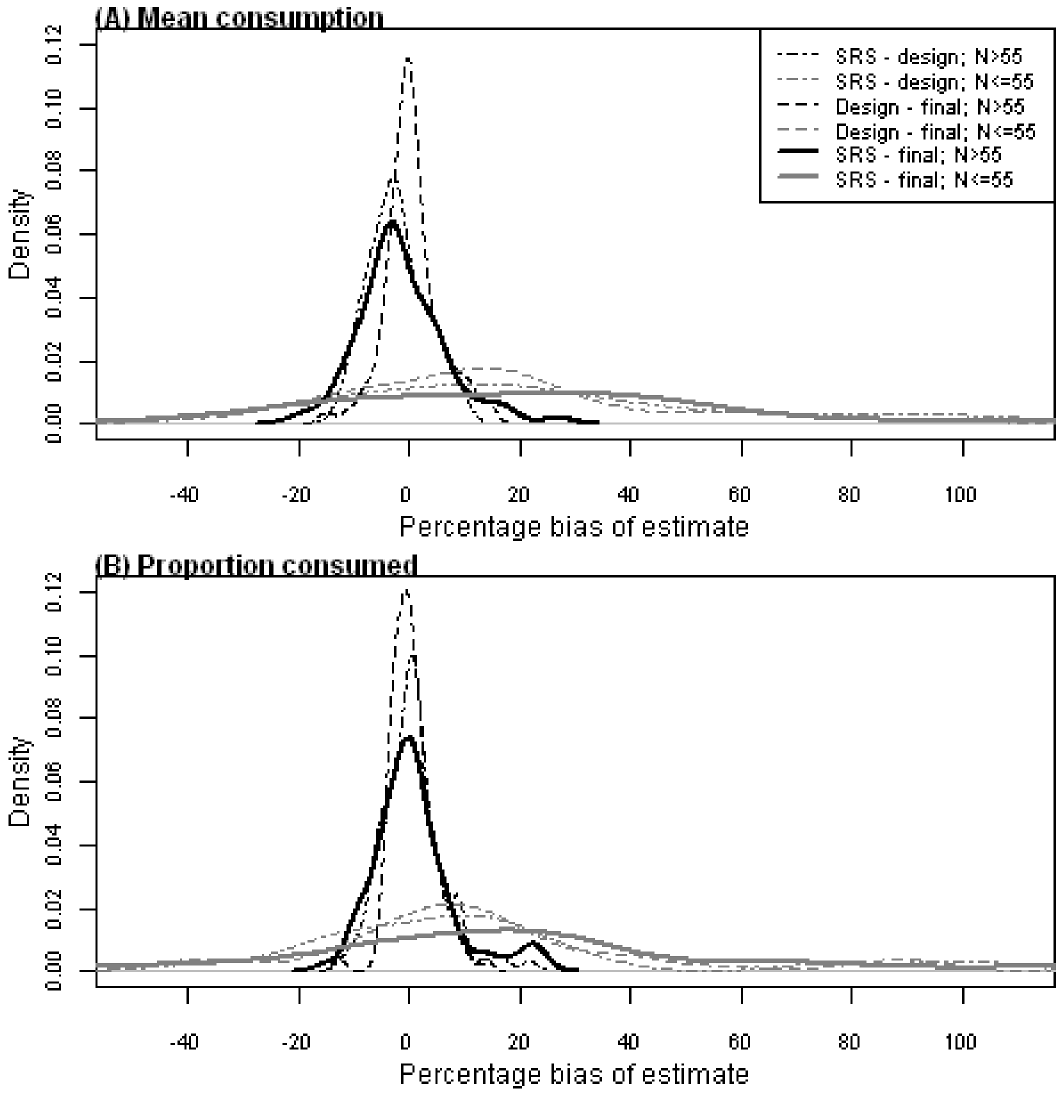

Figure 2 Density plot of the percentage bias of estimates of the mean and proportional consumption of the three digit food groups. The solid lines shows SRS estimation compared with initial weighted estimation and the dashed line shows initial weights compared with final weights.

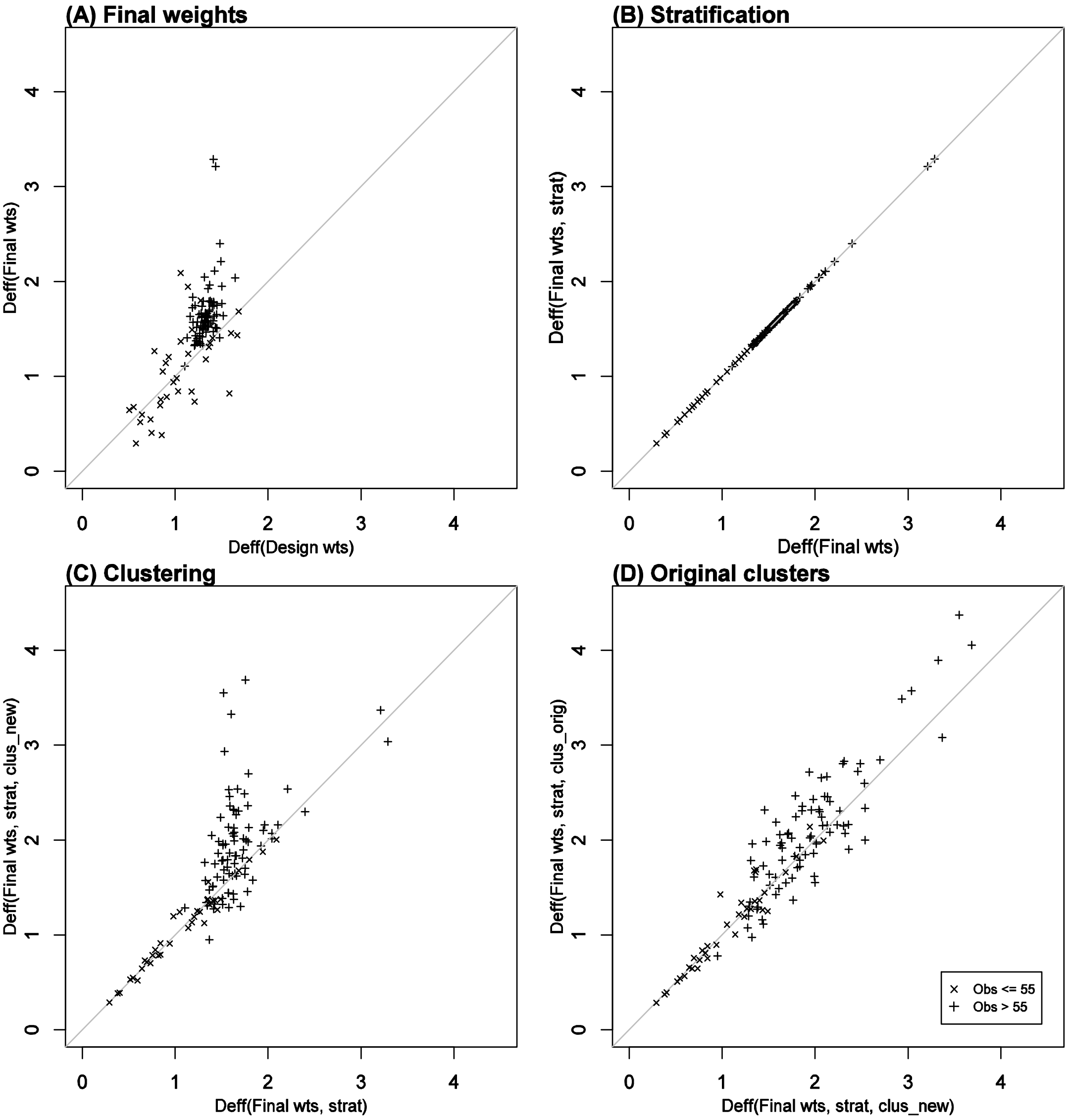

Figure 3 Deffs for mean population consumption (g) of each 3 digit sub-major food group from ANCNPAS 2007. The contribution of each sample design feature to the deff is illustrated by incremental addition of (A) final weights; (B) stratification; (C) clustering using provided locations; and (D) clustering using original 54 clusters.

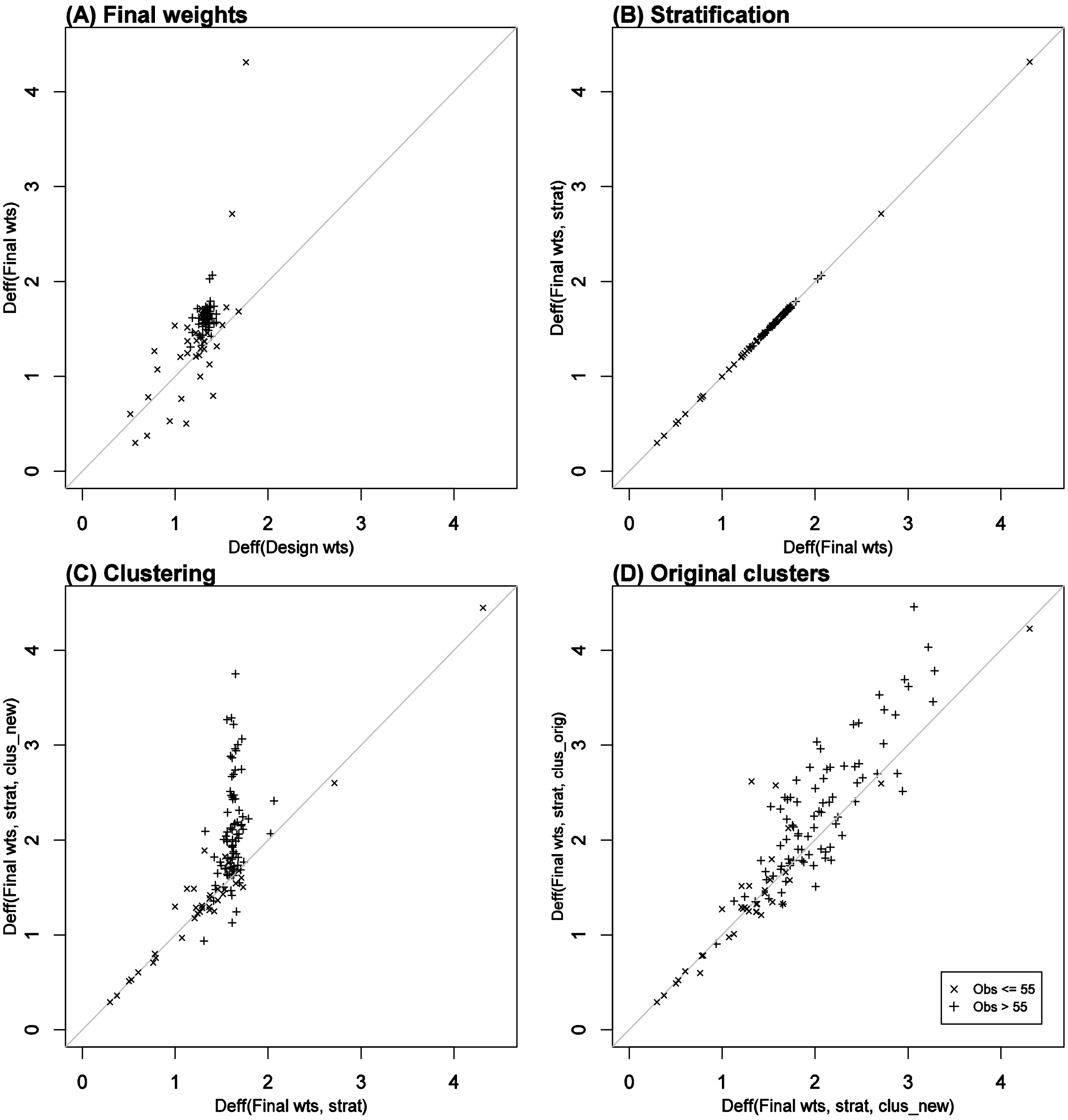

Figure 4 Deffs for proportion of population who consumed each 3 digit sub-major food group from ANCNPAS 2007. The contribution of each sample design feature to the deff is illustrated by incremental addition of (A) final weights; (B) stratification; (C) clustering using provided locations; and (D) clustering using original 54 clusters.

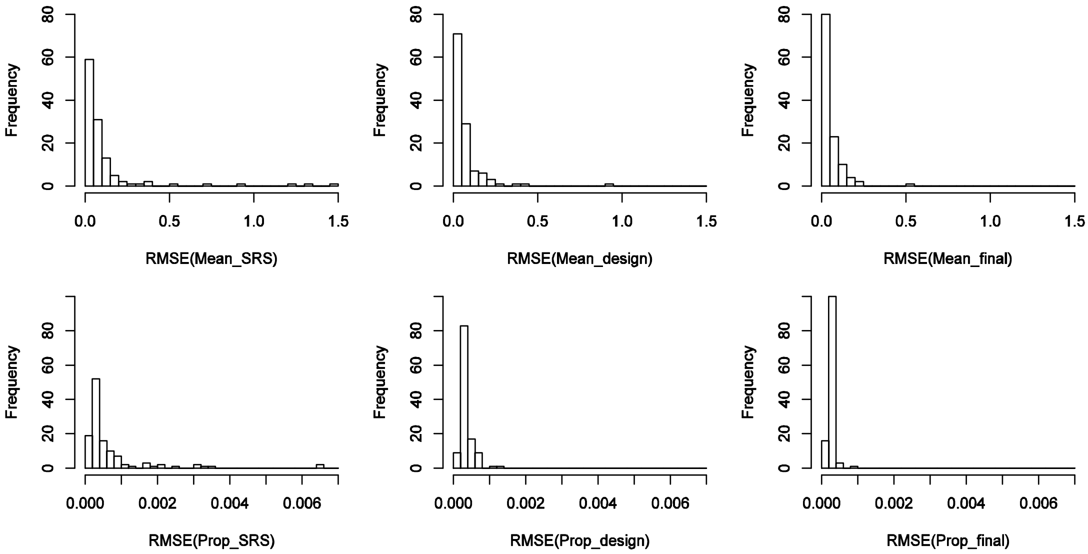

Figure 5 difference between RMSE of SRS and initial weighted estimates and the RMSE of the final weighted estimate for mean and proportion parameters. Positive values indicate that the final weighted estimate has a lower RMSE.

Table 1 Average ((SD), [min,max]) design effects for the three digit food groups for mean consumption (g) (mean) and proportion of population (prop) who consumed each food group. The results are split by the number of non-zero observations.

	Initial weight	Final weight	Stratification	Strat/clustering 210 locations	Strat/clustering 54 clusters
Mean≤55	1.1 (0.32)	1.1 (0.45)	1.1 (0.45)	1.1 (0.44)	1.1 (0.48)
	[0.5, 1.7]	[0.3, 2.1]	[0.3, 2.1]	[0.3, 2.0]	[0.3, 2.1]
Prop≤55	1.2 (0.29)	1.3 (0.69)	1.3 (0.69)	1.4 (0.71)	1.4 (0.75)
	[0.5, 1.8]	[0.3, 4.3]	[0.3, 4.3]	[0.3, 4.4]	[0.3, 4.2]
Mean>55	1.3 (0.09)	1.7 (0.32)	1.7 (0.32)	2.0 (0.53)	2.1 (0.66)
	[1.1, 1.6]	[1.1, 3.3]	[1.1, 3.3]	[0.9, 3.7]	[0.8, 4.4]
Prop>55	1.3 (0.05)	1.6 (0.11)	1.6 (0.11)	2.1 (0.53)	2.3 (0.75)
	[1.2, 1.4]	[1.3, 2.1]	[1.3, 2.1]	[0.9, 3.7]	[0.9, 5.1]

Table 2 Distributional information for the initial weights and final weights and their correlation with the survey variables.

	Mean	**Median**	**SD**	**Min**	**Max**	C_w
Initial weights w_1	727	596	419	114	1722	0.575
Final weights w_2	727	541	577	62.1	5845	0.792
Corr[3](w_1, Y_i)	0.0014	0.0028	0.0269	-0.0926	0.0769	-
Corr(w_2, Y_i)	-0.0005	-0.0006	0.0252	-0.1032	0.0660	-

[3] Corr, correlation

Table 3 Food groups with deffs > 2.0 for mean consumption in grams of each three-digit food group. For each group the deff, the number of observations (N), and for both a simple random sample (SRS) and the clustered design (Clustered) the estimated mean consumption and the 95% confidence interval limits (CI) are included in the table.

		SRS[4]		Clustered		
deff	N	Mn[5]	CI	Mn	CI	Name
2.02	1529	13.6	(12.6, 14.6)	12.9	(11.6, 14.2)	Other Vegetables And Vegetable Combinations
2.02	109	12.2	(9.68, 14.7)	11.9	(8.3, 15.6)	Electolyte, Energy and Fortified Drinks
2.04	916	21.0	(19.3, 22.6)	21.3	(18.9, 23.7)	Poultry And Feathered Game
2.06	3697	67.1	(65.4, 68.9)	69.4	(66.7, 71.8)	Regular Breads, And Bread Rolls (Plain/Unfilled/Untopped Varieties)
2.06	760	5.71	(5.18, 6.23)	6.44	(5.63, 7.27)	Potato Snacks
2.07	1172	13.7	(12.7, 14.8)	13.6	(12.1, 15.2)	Other Fruiting Vegetables
2.07	344	13.6	(11.9, 15.4)	13.8	(11.2, 16.4)	Mixed Dishes Where Beef, Veal Or Lamb Is The Major Component
2.08	1746	5.47	(5.08, 5.87)	5.89	(5.25, 6.52)	Sugar, Honey And Syrups
2.15	855	5.47	(4.92, 6.02)	5.70	(4.89, 6.51)	Leaf And Stalk Vegetables
2.15	999	17.5	(16.2, 18.7)	18.3	(16.4, 20.3)	Cakes, Buns, Muffins, Scones, Cake-Type Desserts
2.16	601	9.07	(8.25, 9.9)	9.51	(8.22, 10.8)	Sausages, Frankfurts And Saveloys
2.16	1036	48.8	(45.3, 52.3)	48.9	(43.8, 54)	Mixed Dishes Where Cereal Is The Major Ingredient
2.16	1199	9.42	(8.68, 10.2)	9.7	(8.6, 10.8)	Chocolate And Chocolate-Based Confectionery
2.19	314	6.61	(5.65, 7.56)	7.27	(5.85, 8.71)	Dishes Where Vegetable Is The Major Component
2.24	132	3.61	(2.91, 4.31)	4.18	(3.05, 5.31)	Other Dishes Where Milk Or A Milk Product Is The Major Component
2.24	1875	63.1	(60.2, 66.1)	65.6	(61.2, 70.1)	Pome Fruit
2.30	992	14.2	(13, 15.4)	14.1	(12.3, 15.8)	Cordials
2.31	583	1.27	(1.14, 1.4)	1.22	(1.02, 1.41)	Dairy Blends
2.31	924	7.83	(7.19, 8.47)	7.71	(6.71, 8.71)	Peas And Beans
2.31	784	25.2	(22.9, 27.6)	26.1	(22.4, 29.8)	Other Fruit
2.32	117	2.59	(2, 3.17)	2.60	(1.72, 3.49)	Fin Fish (Excluding Commerically Sterile)
2.32	683	4.77	(4.38, 5.15)	4.90	(4.31, 5.49)	Cereal-, Fruit-, Nut- And Seed-Bars
2.34	2163	20.2	(18.9, 21.6)	21.2	(18.8, 23.6)	Gravies And Savoury Sauces
2.35	635	2.87	(2.52, 3.22)	2.94	(2.39, 3.5)	Nuts And Nut Products
2.41	117	0.67	(0.52, 0.82)	0.84	(0.48, 0.95)	Extruded Or Reformed Snacks
2.43	640	17.3	(15.5, 19)	18.3	(15.4, 21.3)	Mixed Dishes Where Poultry Or Game Is The Major Component
2.46	601	8.52	(7.63, 9.41)	8.70	(7.29, 10.1)	Batter-Based Products
2.46	287	0.12	(0.09, 0.14)	0.13	(0.06, 0.14)	Multivitamin and/or Mineral
2.47	1361	13.1	(12.3, 14)	13.0	(11.7, 14.3)	Processed Meat

[4] SRS, simple random sample
[5] Mn, mean

2.60	2067	53.5	(50.9, 56)	54.4	(50.3, 58.6)	Potatoes
2.66	1347	26.2	(24.6, 27.9)	27.2	(24.4, 29.8)	Muscle Meat
2.67	122	7.24	(5.63, 8.84)	6.61	(4.15, 9.06)	Dairy Milk Substitutes, Unflavoured
2.71	1085	2.70	(2.01, 3.38)	2.80	(1.47, 4.12)	Herbs, Spices, Seasonings And Stock Cubes
2.72	1371	124	(116, 131)	129	(117, 142)	Soft Drinks, And Flavoured Mineral Waters
2.80	400	0.51	(0.43, 0.60)	0.58	(0.41, 0.74)	Vegetable/Nut Oil
2.80	237	1.26	(1.03, 1.49)	1.21	(0.834, 1.59)	Cream
2.83	315	22.7	(19.7, 25.6)	23.0	(17.8, 28.2)	Soup (Prepared, Ready to Eat)
2.84	1013	15.8	(14.6, 17)	16.5	(14.4, 18.6)	Tomato And Tomato Products
3.08	410	9.90	(8.68, 11.1)	11.9	(9.25, 14.6)	Dishes And Products Other Than Confectionery Where Sugar Is the main component
3.49	1230	30.5	(28.7, 32.4)	28.9	(25.5, 32.3)	Tropical Fruit
3.57	742	2.57	(2.33, 2.82)	2.88	(2.31, 3.45)	Jam And Lemon Spreads, Chocolate Spreads, Sauces
3.89	4518	792	(774, 810)	826	(789, 862)	Mineral Waters And Water
4.05	878	29.6	(27, 32.3)	30.9	(25.5, 36.2)	Flours And Other Cereal Grains And Starches
4.37	1813	3.87	(3.67, 4.07)	3.76	(3.34, 4.17)	Margarine and Table Spreads

Table 4 Food groups with deffs > 2.0 for proportion of population consuming the food group. For each group the deff, the number of observations (N), and for both a simple random sample (SRS) and the clustered design (Clustered) the estimated proportion of the population consuming the food group and the 95% confidence interval limits (CI) are included in the table.

		SRS		Clustered		
deff	N	Mn	CI	Mn	CI	Name
2.01	1268	0.3	(0.25, 0.28)	0.3	(0.24, 0.28)	Pasta And Pasta Products
2.04	1746	0.4	(0.35, 0.38)	0.4	(0.34, 0.38)	Sugar, Honey And Syrups
2.04	635	0.1	(0.12, 0.14)	0.1	(0.12, 0.15)	Nuts And Nut Products
2.05	583	0.1	(0.11, 0.13)	0.1	(0.1, 0.13)	Dairy Blends
2.07	126	0.03	(0.02, 0.03)	0.03	(0.02, 0.03)	Single vitamin
2.13	601	0.1	(0.12, 0.13)	0.1	(0.11, 0.14)	Sausages, Frankfurts And Saveloys
2.14	175	0.0	(0.03, 0.04)	0.0	(0.03, 0.05)	Packed (Commercially Sterile) Fish And Seafood
2.16	640	0.13	(0.12, 0.14)	0.13	(0.12, 0.15)	Mixed Dishes Where Poultry Or Game Is The Major Component
2.17	410	0.08	(0.08, 0.09)	0.09	(0.08, 0.11)	Dishes And Products Other Than Confectionery Where Sugar Is the major component
2.22	344	0.1	(0.06, 0.08)	0.1	(0.06, 0.08)	Mixed Dishes Where Beef, Veal Or Lamb Is The Major Component
2.24	518	0.11	(0.1, 0.12)	0.11	(0.1, 0.12)	Eggs
2.25	117	0.0	(0.02, 0.03)	0.0	(0.02, 0.03)	Fin Fish (Excluding Commerically Sterile)
2.29	132	0.03	(0.02, 0.03)	0.0	(0.02, 0.04)	Other Dishes Where Milk Or A Milk Product Is The Major Component
2.30	315	0.07	(0.06, 0.07)	0.06	(0.05, 0.07)	Soup (Prepared, Ready to Eat)
2.33	1875	0.39	(0.37, 0.4)	0.39	(0.37, 0.41)	Pome Fruit
2.35	129	0.0	(0.02, 0.03)	0.0	(0.02, 0.03)	Dishes Where Egg Is The Major Ingredient
2.39	829	0.2	(0.16, 0.18)	0.2	(0.14, 0.18)	Cabbage, Cauliflower And Similar Brassica Vegetables
2.40	683	0.14	(0.13, 0.15)	0.15	(0.13, 0.16)	Cereal-, Fruit-, Nut- And Seed-Bars
2.40	287	0.06	(0.05, 0.07)	0.06	(0.05, 0.07)	Multivitamin and/or Mineral
2.40	1172	0.2	(0.23, 0.25)	0.2	(0.22, 0.26)	Other Fruiting Vegetables
2.43	4518	0.94	(0.93, 0.94)	0.94	(0.93, 0.95)	Mineral Waters And Water
2.45	1361	0.28	(0.27, 0.29)	0.28	(0.26, 0.3)	Processed Meat
2.45	900	0.2	(0.18, 0.2)	0.2	(0.17, 0.21)	Citrus Fruit
2.45	916	0.19	(0.18, 0.2)	0.19	(0.17, 0.2)	Poultry And Feathered Game
2.51	835	0.17	(0.16, 0.18)	0.19	(0.17, 0.2)	English-Style Muffins, Flat Breads, And Savoury and Sweet Breads
2.54	122	0.0	(0.02, 0.03)	0.0	(0.02, 0.03)	Dairy Milk Substitutes, Unflavoured
2.60	601	0.12	(0.12, 0.13)	0.13	(0.11, 0.14)	Batter-Based Products
2.63	322	0.07	(0.06, 0.07)	0.07	(0.05, 0.08)	Tea
2.65	266	0.1	(0.05, 0.06)	0.0	(0.04, 0.05)	Berry Fruit

2.65	237	0.0	(0.04, 0.06)	0.0	(0.04, 0.06)	Cream
2.70	1036	0.2	(0.2, 0.23)	0.2	(0.2, 0.24)	Mixed Dishes Where Cereal Is The Major Ingredient
2.70	1813	0.38	(0.36, 0.39)	0.36	(0.34, 0.38)	Margarine and Table Spreads
2.74	1622	0.34	(0.32, 0.35)	0.32	(0.3, 0.34)	Carrot And Similar Root Vegetables
2.76	178	0	(0.03, 0.04)	0	(0.03, 0.04)	Pickles, Chutneys And Relishes
2.77	1347	0.28	(0.27, 0.29)	0.28	(0.26, 0.3)	Muscle Meat
2.77	999	0.21	(0.2, 0.22)	0.21	(0.19, 0.23)	Cakes, Buns, Muffins, Scones, Cake-Type Desserts
2.78	760	0.2	(0.15, 0.17)	0.2	(0.16, 0.19)	Potato Snacks
2.81	1001	0.2	(0.2, 0.22)	0.2	(0.2, 0.24)	Other Confectionery
2.96	314	0.07	(0.06, 0.07)	0.1	(0.06, 0.08)	Dishes Where Vegetable Is The Major Component
3.01	992	0.2	(0.19, 0.22)	0.2	(0.19, 0.23)	Cordials
3.03	784	0.16	(0.15, 0.17)	0.16	(0.14, 0.18)	Other Fruit
3.22	117	0	(0.02, 0.03)	0	(0.02, 0.03)	Extruded Or Reformed Snacks
3.23	2195	0.5	(0.44, 0.47)	0.4	(0.4, 0.45)	Cheese
3.32	2067	0.4	(0.41, 0.44)	0.4	(0.4, 0.45)	Potatoes
3.37	855	0.2	(0.17, 0.19)	0.2	(0.17, 0.21)	Leaf And Stalk Vegetables
3.46	924	0.2	(0.18, 0.2)	0.2	(0.16, 0.2)	Peas And Beans
3.53	400	0.1	(0.08, 0.09)	0.1	(0.06, 0.09)	Vegetable/Nut Oil
3.62	1013	0.2	(0.2, 0.22)	0.2	(0.19, 0.24)	Tomato And Tomato Products
3.69	1371	0.3	(0.27, 0.3)	0.3	(0.27, 0.32)	Soft Drinks, And Flavoured Mineral Waters
3.78	1230	0.3	(0.24, 0.27)	0.2	(0.21, 0.26)	Tropical Fruit
4.03	2163	0.4	(0.43, 0.46)	0.5	(0.42, 0.48)	Gravies And Savoury Sauces
4.46	878	0.2	(0.17, 0.19)	0.2	(0.16, 0.21)	Flours And Other Cereal Grains And Starches
5.09	1085	0.22	(0.21, 0.24)	0.22	(0.2, 0.25)	Herbs, Spices, Seasonings And Stock Cubes