WEIGHTING PROCEDURES FOR USDA'S CONTINUING SURVEY OF FOOD INTAKES BY INDIVIDUALS 1994-96

Adam Chu, Westat, Joseph Goldman, USDA Agricultural Research Service
Adam Chu, Westat, 1650 Research Boulevard, Rockville, Maryland 20850

Key Words: ratio raking estimation, calibration, nonresponse adjustments

1. Introduction

This paper discusses the weighting procedures employed for the USDA's Continuing Survey of Food Intakes by Individuals (CSFII) 1994-96. The CSFII 1994-96, which surveyed over 15,000 individuals over a three-year data collection period, was designed to provide separate annual estimates of mean intakes of food groups and nutrients, as well as combined three-year averages. For this purpose, sampling weights were constructed for each reporting year and for all three years combined. An important feature of the weighting methodology was the use of raking adjustments to calibrate the weighted sample counts to population totals obtained from the Current Population Survey (CPS) on up to 14 different classification variables. Although the raking algorithm provides a straightforward way of making the required adjustments, it also tends to increase the variation in the sampling weights when the number of raking variables is large. Some limited evaluation of the impact of the raking adjustments on the precision of estimates developed from the CSFII 1994-96 is provided.

2. Components of the Sampling Weight

In general, the analysis of survey data from complex sample designs requires the use of weights to compensate for variable probabilities of selection, differential response rates, and deficiencies in the sampling frame (e.g., undercoverage of certain groups of individuals). For the CSFII 1994-96, the overall probabilities of selection were designed to vary by sex, age, and income level in order to meet specified sample size requirements. For example, in the 1994 survey, there was a more than fivefold variation in the probabilities of selection, ranging from about 1 in 60,000 of nonlow-income females between the ages of 30-39 years to about 1 in 11,000 of low-income males between the ages of 50-59 years and children between the ages of 1 and 2 years. The need to modify the sampling rates periodically and the occasional use of special subsampling procedures introduced additional variation in the overall selection probabilities (Goldman et al., 1997).

Base Weight. The base weight associated with a sampled person (SP) is equal to the reciprocal of the probability of including that person in the sample. As described in Goldman et al. (1997), individuals were selected for the CSFII through a multistage area probability sampling design in which 62 primary sampling units (PSUs) consisting of metropolitan statistical areas (MSAs) or counties were selected at the first stage, area segments consisting of Census-defined blocks or groups of blocks were selected within PSUs at the second stage, households were selected within segments at the third stage, and SPs were selected within households at the fourth and final stage. Thus, the following probabilities were required for the calculation of the base weights: (1) the probability of selecting PSU h, P_h; (2) the probability of selecting area segment i within the PSU, S_{hi}; (3) the probability of selecting household j within the segment, H_{hij}; and (4) the probability of selecting a person (SP) k in a specified sex-age-income domain a within the household for the intake interviews, I_{ahijk}.

Since a goal of the sample design was to select selfweighting samples of SPs within each sex-age-income domain, the overall probability of selecting person k in sex-age-income domain a is generally:

$$P_{ahijk} = P_h S_{hi} H_{hij} I_{ahijk} = 1/K_a, \qquad (1)$$

where K_a depends only on the sex-age-income domain to which the SP was assigned at the time of screening. Note that the within-segment sampling rate, H_{hij}, was designed to yield a selfweighting sample of households in each survey year. The term K_a in formula (1) is the reciprocal of the desired overall rate of selecting persons in domain a and can be thought of as the "desired" base weight. The values of K_a specified for each of the three years of the CSFII 1994-96 are documented in Westat (1997). The actual base weight differs from the desired base weight because of the subsampling used to select some new or missed dwelling units. The actual base weight for an SP in domain a was computed as

$$w_{ahijk}^{base} = K_a / H_{hij}^R \qquad (2)$$

where H_{hij}^R is the subsampling interval used to select new or missed dwelling units (DUs) in designated area segments. In most cases, $H_{hij}^R = 1$. Values of $H_{hij}^R > 1$ applied to a small number of DUs found through the "missed structure" procedure described in Goldman et al. (1997).

The weights given by formula (2) applied to SPs 1 year of age or older. Infants under 1 year of age were included in the sample if another eligible SP (1 year old or older) in the household had been selected for the CSFII. Therefore, the probability of selecting an infant for the CSFII was the same as the probability of retaining that household for the CSFII. The base weight for infants is the inverse of the household selection probability.

Nonresponse Adjustments. Unit nonresponse (i.e., whole questionnaire nonresponse) occurs when an eligible SP fails to respond to the survey for any reason. As described below, two stages of nonresponse adjustments were made. The first-stage adjustment was designed to compensate for nonresponse to the Screener Interview. The second-stage adjustment was designed to compensate for nonresponse to the Day 1 Intake Interview, given that the Screener had been completed. The general approach for both stages of adjustments was to divide the sample into a number of homogeneous weighting classes, within which nonresponse-adjusted weights were calculated by multiplying the base weights by the corresponding inverse of the weighted response rate for the class.

To develop the first-stage adjustment, each sampled dwelling unit (DU) was assigned a Screener base weight, w_i^{scr}, defined as:

$$w_i^{scr} = I_t / H_i^{ret} \tag{3}$$

where I_t is the reciprocal of the overall probability of selecting a DU in year t of the study, and H_i^{ret} is the probability of retaining new and missed DUs identified by the "missed structure" procedure. Screener nonresponse adjustment classes were then defined by crossing the following four segment-level variables: Census region; MSA status; minority status of segment (segments with "high" percentage of black or Hispanic residents versus all others); and quarter of data collection. Within each cell defined by this cross-classification, an adjustment factor, F_g^{scr}, was computed as the ratio of the sum of the weights of the eligible DUs in the sample to the corresponding sum of weights of the responding DUs. The factor F_g^{scr} is a DU-level adjustment that was used to inflate the base weights of the responding DUs.

An analogous adjustment was made to compensate for nonresponse to the Day 1 intake interview. Separately by year of study, each SP selected for the intake interview was assigned an initial weight, $w_{gi}^{I1} = F_g^{scr} w_{gi}^{base}$, where w_{gi}^{base} is the base weight for SP i in screener adjustment class g, and F_g^{scr} is the corresponding screener nonresponse adjustment factor. Weighting classes for adjusting the initial weights were then identified by a CHAID analysis (Magidson, 1993). The CHAID algorithm (which stands for "chi-square automatic interaction detector") was used to identify subsets of the sample that were internally homogeneous with respect to response rates. In addition to the four classification variables used to calculate the Screener nonresponse adjustment factors, sex, age, and income status (below/above 130 percent of Federal poverty guidelines as reported in the Screener or Household questionnaires) were specified as independent variables in the CHAID analysis. The number and structure of the subsets (adjustment classes) that were determined by the CHAID analysis differed by year. Within adjustment class h, a Day 1 Intake nonresponse adjustment factor, F_h^{I1}, was computed using formulas analogous to those used to compute F_g^{scr}. The corresponding nonresponse-adjusted Day 1 Intake weight for SP i in class h was then computed as: $w_{hi}^{A1} = F_h^{I1} w_{hi}^{I1}$.

Poststratification Adjustments. In addition to compensating for unequal selection probabilities and nonresponse, another important function of weighting is to adjust for sampling variability and possible undercoverage in the sampling frame. Therefore, the final step of the weighting process was to poststratify or "calibrate" the nonresponse-adjusted weights so that the sum of the final weights equaled the corresponding Current Population Survey (CPS) population estimates within cells defined by the following variables: sex, age group (seven categories based on intake interview), Census region, MSA status (MSA vs. nonMSA), household income level (defined in terms of percent of Federal poverty guidelines in four broad categories: 0-75%; 76-130%; 131-300%; 301%+), whether or not the household received food stamps in the last 12 months, home ownership status, presence of children under six years of age in household, presence of children 6-17 years of age in household, number of adults in household, presence of female head of household 40 years old or younger and no one under 18 years of age, employment status of the female head of household, employment status of the SP, race of SP (black vs. nonblack), and Hispanic origin of SP (Hispanic vs. nonHispanic). The weights were also balanced by season of intake (winter, spring, summer, fall) and day of week of intake.

The poststratification was implemented by an iterative process known as raking (e.g., see Oh and Scheuren, 1978). The raking process was carried out separately for each of the following four major subsets: (1) males, 20 years of age or older; (2) females, 20 years of age or older; (3) children 0-5 years of age; and (4) persons 6-19 years of age. The variables used in the raking process are documented in Westat (1997). Not all of the variables applied to all subsets. For example, the variable on employment status of the SP was not applicable to the subsets of younger persons. Most of the raking variables were intrinsically univariate, but a few were bivariate in structure.

The raking algorithm used to calculate the final (poststratified or "raked") weights for the Day 1 Intake respondents was essentially as follows. First, for each level defined by the first raking variable DIM1, an adjustment factor, $F_{DIM1}^{(1)}$, was computed as: $F_{DIM1}^{(1)} = N_{DIM1} / \sum_{i=1}^{n_1} w_i^{NR}$, where N_{DIM1} is the CPS control total for the given level of DIM1, w_i^{NR} is the nonresponse-adjusted Day 1 weight, and where the denominator of $F_{DIM1}^{(1)}$ extends over the responding SPs in the given cell (level) of DIM1. An intermediate DIM1-adjusted weight was then calculated as: $w_i^{DIM1} = w_i^{NR} F_{DIM1}^{(1)}$. Next, the w_i^{DIM1}'s calculated above were used to calculate an adjustment within each level of DIM2 in an analogous manner: $F_{DIM2}^{(1)} = N_{DIM2} / \sum_{i=1}^{n_2} w_i^{DIM1}$, where N_{DIM2} is the CPS control total for the given level of DIM2 and where the denominator of $F_{DIM2}^{(1)}$ extends over the responding SPs in the given cell (level) of DIM2. An intermediate DIM2-adjusted weight was then calculated as: $w_i^{DIM2} = w_i^{DIM1} F_{DIM2}^{(1)}$.

The w_i^{DIM2}'s were then used to calculate an intermediate DIM3-adjusted weight, w_i^{DIM3}, using procedures analogous to those described above. This procedure continued up to the last raking variable, DIM13 (or DIM14 depending on the subset), using the previously adjusted intermediate weights. Using the weights developed from the previous iteration, the whole process was then repeated starting with DIM1 and continuing through DIM13 (or DIM14). The iteration process continued until the difference between the calculated weighted sums and the corresponding CPS totals was acceptably small for *all* levels of each raking variable (i.e., within 0.005% of the corresponding CPS totals for each level of each raking variable). Generally, convergence was achieved within 15 iterations.

The effect of the raking adjustments on the variation of weights is illustrated in Table A. The table summarizes the coefficient of variation (CV) of the nonresponse-adjusted (pre-raked) weights and final raked weights for male respondents in the 1994 and 1996 surveys. The coefficient of variation of the weights (expressed as a percentage) is $c_w = 100 s_w / \bar{w}$, where $\bar{w}$ and s_w are the (unweighted) mean and standard deviation of the weights, respectively. The term $1+(c_w/100)^2$ is a variance inflation factor (VIF) or "design effect" due to unequal weighting. As a rough guide, VIF−1 (expressed as a percentage) is the expected amount by which the variance of an estimate is increased as a result of the variation in weights (e.g., see Kish, 1992). The average VIF in Table A (averaged over the 10 age groups and all three years of the study) corresponding to the nonresponse-adjusted weights is about 1.12. The average VIF corresponding to the final raked weights is about 1.21. As can be seen in the last column of Table A, the average extra increase in variance (averaged over the 3 years of the study) varies by age group from 4 to 11 percent. For females (results not shown), the average extra increase in variance ranged from 3 to 8 percent.

Table A. CV (%) of weights for male respondents by age and survey year

Age group (males)	1994 CV (%) of NR-adjusted weights	1994 CV (%) of final (raked) weights	1996 CV (%) of NR-adjusted weights	1996 CV (%) of final (raked) weights	Average extra increase in variance
1 to 2	11.9	25.6	16.2	34.7	4%
3 to 5	23.3	39.3	23.0	32.8	8%
6 to 11	35.8	40.7	42.2	52.1	5%
12 to 19	42.4	46.6	16.3	37.4	8%
20 to 29	38.6	50.9	16.2	43.6	9%
30 to 39	54.2	59.0	38.2	49.9	9%
40 to 49	41.5	49.8	38.3	50.7	7%
50 to 59	36.6	48.3	41.5	56.1	10%
60 to 69	40.0	47.1	35.6	60.0	11%
70+	18.4	38.2	18.5	43.6	11%

3. Development of Weights for Analysis of Combined CSFII Data Set

The appropriate weights for analysis of the combined 3-year CSFII data set depend on the types of analyses to be conducted and on assumptions about changes in food consumption over the three years of data collection. If year-to-year changes are negligible, pooling the annual samples will provide a better basis for estimating the distribution of nutrient intakes, as well provide more precise estimates of population parameters such as means and percentiles. In this case, it is possible to simply pool the three annual samples, attach base weights that reflect the probabilities of selection over the three years of the study, and to apply nonresponse and poststratification adjustments to the combined sample. The resulting weights may be thought of as "pooled" weights. Although such weights may be more efficient under certain circumstances, they have not been computed for the CSFII 1994-96 for reasons given below.

Since the selection probabilities of SPs varied from year to year (and, occasionally, within the year), a combined estimate using the pooled weights will be biased if the expected value of the survey item is not constant over the three-year period. Under these conditions, an overall three-year mean may be of less interest analytically than the individual annual means. However, an approximately unbiased estimate of an overall three-year mean can be constructed by re-calibrating the nonresponse-adjusted weights previously computed for each annual sample to the corresponding three-year average CPS population counts. The resulting weights (referred to as the "combined annual weights") are applicable whether or not there are changes in the items of interest over time. However, because SPs in different survey years have different probabilities of selection, this set of weights will be less efficient (result in larger variance) than the pooled weights for survey items that do not vary in expectation over time.

4. Discussion

Total Design Effects. The design effect of an estimated mean, $\bar{y}$, based on a complex sample design is defined as $D_T(\bar{y}) = var(\bar{y})/var(\bar{y}_0)$, where $var(\bar{y})$ is the design-based variance of the estimate, and $var(\bar{y}_0)$ is the corresponding variance that would have been obtained from a simple random sample of the same size. The unequal weighting effect is one component of the total design effect. For the CSFII sample design, the other major component of the design effect is the clustering design effect, $D_c(\bar{y}) \approx 1+(b-1)\rho$, where b is the average cluster size and ρ is the intraclass correlation for the given survey item. Under the assumption that the two components are multiplicative i.e., $D_T(\bar{y}) = D_w D_c(\bar{y})$, where D_w is the unequal weighting effect (using VIF as a rough indicator of D_w), $D_c(\bar{y})$ can be estimated by $D_T(\bar{y})/D_w$ (e.g., Verma, 1993, Chapter 6).

Table B summarizes the average coefficients of variation and design effects for selected nutrient intake statistics for the total (all-income) sample. The CVs and design effects presented in the table were computed using jackknife replication, and are simple averages over the 10 age groups listed in Table A (excluding the under 1 year-old age group). The results are intended to give a rough overall indication of the magnitude of the CVs and design effects for the selected statistics. In particular, it should be noted that the three-year precision goal established for the CSFII 1994-96 was that the estimated mean iron and saturated fat intake for each of 20 all-income sex-age domains be subject to a coefficient of variation of no more than 3 percent. The average CVs shown in Table B for iron and saturated fat intake are based on one year of CSFII data (i.e., about 1/3 of the total sample size). Since the average CVs are less than $3\sqrt{3}$ = 5.2 percent, they are consistent with the goals established for the all-income sample. It can also be noted that the average clustering design effect, D_c, is generally close to 1, suggesting that the effects of clustering are negligible for age-group-specific estimates. This is not unexpected in view of the relatively small average domain sample size per PSU. (This would not be true for overall estimates involving all age groups, where D_c has been estimated to be as high as 1.5 to 2.0.)

Table B. CVs and design effects for selected nutrient intake statistics (1994 CSFII data)

Item	Sex	Survey est.*	Avg.† CV (%) of estimate	Avg.† total design effect (D_T)	Avg.† clust. design effect (D_c)
Total energy (kcal)	Male	2,356	2.78	1.20	0.98
	Female	1,631	2.74	1.37	1.18
Iron (mg)	Male	17.55	4.24	1.24	1.02
	Female	12.69	3.92	1.08	0.93
Sat. fat (g)	Male	30.50	3.61	1.03	0.84
	Female	20.63	3.82	1.31	1.13

*Survey estimate using final raked weights.

†Average over the 10 1+ year age groups specified for the CSFII.

Effect of Number of Raking Dimensions on Variation of Weights. As indicated in Table A, a consequence of the raking procedures employed for the CSFII was to increase the variation of the sampling weights. Moreover, as illustrated in Figure 1, the raking algorithm has a tendency to generate some extremely large weights. Both the number and choice of variables to be included in the raking process will influence the variation in weights. In particular, a variable with a (weighted) distribution that differs substantially from the corresponding CPS distribution will typically result in much a higher VIF than one whose distribution closely matches the CPS distribution.

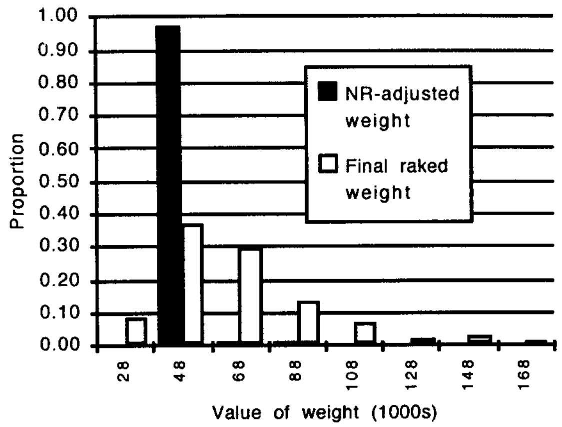

Figure 1. Distribution of 1996 CSFII weights for nonlow-income males (20+ years)

To investigate the variance impact of the large number of raking variables employed, alternative sets of weights were generated for a subset of the 1996 CSFII data (males, 20 years and over) by varying the choice and number of variables included in the raking algorithm. In addition to the complete set of 13 raking variables (dimensions) that were used to create the final CSFII weights for this subset, three different sets of raking variables were specified for generating alternative versions of final weights: (1) a set of 12 raking variables obtained by deleting the day-of-week-of-intake variable (this variable was deleted because its distribution in the sample differed considerably from the intended uniform distribution); (2) a subset consisting of the 7 "most important" variables with respect to food consumption; and (3) a smaller subset consisting of the 4 "most important" variables with respect to food consumption. Some results of this analysis are summarized in Table C. In addition to the CVs corresponding to the four versions of the raked weights, the CVs of the nonresponse-adjusted (pre-raked) weights are included for comparison. Note that the relatively large CVs of the nonresponse-adjusted weights are due primarily to the oversampling of low-income persons in each age group. As can be seen in the bottom half of the table, the percentage increase in variance more than doubles on average as the number of raking dimensions goes from 4 to 13. Also, as expected, the day-of-week-of-intake variable appears to have a significant impact on the increased variation of the weights. By omitting this one variable, the average increase in variance can be reduced from 14 percent (Version 0) to 10 percent (Version 1).

It is not sufficient to simply look at the variation in the final weights when assessing the performance of a set of weights. It is also important to examine the impact the weights have on the survey estimates themselves. Large differences in the estimates under alternative sets of weights may suggest that some versions of the weights are less effective in removing biases than others. In Table D, we have calculated estimates of total energy intake using the four alternative sets of weights described previously. The estimates are for males 20 years of age or older in the 1996 CSFII data set. As can be seen in this table, there are virtually no differences among the alternative estimates.

Table C. CV of weights and increase in variance under alternative weighting schemes (males, 1996)

Age group (males)	Version 0 (13 raking dimensions)	Version 1 (12 raking dimensions)	Version 2 (7 raking dimensions)	Version 3 (4 raking dimensions)	NR-adjusted weights (0 raking dimensions)
	CV of weights (%):				
20 to 29	43.6	38.7	31.7	30.0	16.2
30 to 39	49.9	47.3	48.6	46.9	38.2
40 to 49	50.7	47.9	45.0	43.4	38.3
50 to 59	56.1	50.6	49.9	48.6	41.5
60 to 69	60.0	55.0	48.8	43.4	35.6
70+	43.6	39.3	32.3	31.8	18.5
	% Increase in variance relative to NR wts.				
20 to 29	16%	12%	7%	6%	0%
30 to 39	9%	7%	8%	6%	0%
40 to 49	10%	7%	5%	4%	0%
50 to 59	12%	7%	6%	5%	0%
60 to 69	21%	16%	10%	5%	0%
70+	15%	12%	7%	6%	0%
Average	*14%*	*10%*	*7 %*	*6 %*	*0 %*

On the other hand, the calculated CVs of the estimates tend to be somewhat larger with the sets of weights derived from the larger numbers of raking dimensions (i.e., Versions 0 or 1). However, this pattern does not hold for all age groups. Moreover, the calculated CVs of estimates of mean iron and saturated fat intakes under the various weighting schemes exhibit the opposite pattern for some age groups; i.e., for some age groups, the CVs tend be slightly smaller with the weights derived using the larger numbers of raking variables. One possible reason why the calculated CVs do not exhibit the trends noted earlier in Table C is that the various raking schemes may have had a differential effect on the precision of the survey estimates. In general, poststratification (using raking or standard methods) tends to improve precision. The resulting increased precision may thus offset the theoretical increase in sampling variance arising from the variability of the weights. As a result, the calculated CVs shown in Table D reflect the combined effects of the extra variation in weights and poststratification, and not simply the effect of the variation in weights alone. Further study of the impact of the variation in weights is warranted. In particular, it would be interesting to see how the calculated CVs vary for other CSFII survey items (including both continuous and categorical variables), and whether the effect of the variation in weights can be isolated from the calculated results.

Table D. Comparison of estimates of total energy intake under alternative weighting schemes for males, 1996

Age group (males)	Version 0 (13 raking dimensions)	Version 1 (12 raking dimensions)	Version 2 (7 raking dimensions)	Version 3 (4 raking dimensions)
	Estimated mean (1996):			
20 to 29	2,781	2,784	2,797	2,817
30 to 39	2,605	2,596	2,613	2,619
40 to 49	2,530	2,508	2,489	2,481
50 to 59	2,291	2,292	2,300	2,304
60 to 69	2,030	2,019	2,055	2,059
70+	1,790	1,785	1,772	1,775
	Coefficient of variation (%):			
20 to 29	3.67	3.61	2.91	3.15
30 to 39	2.42	2.47	2.22	2.27
40 to 49	2.89	2.91	2.90	2.76
50 to 59	2.70	2.82	2.80	2.87
60 to 69	2.15	2.15	2.18	2.16
70+	2.40	2.46	2.22	2.24

Effect of Choice of Raking Variables on Combined Three-Year Estimates. In developing the weighting procedures for the combined three-year sample, consideration was given to using "month of intake" as a raking dimension instead of the 12-level year-specific "season of intake." For the annual samples, season rather than month of intake was included as a raking variable because the monthly sample sizes were considered to be too small and unequally distributed. However, it was hoped that the much larger size of the combined three-year sample would mitigate these concerns. To examine this issue, two sets of combined three-year weights were initially developed. The first of these (referred to as the "Q-weights") was generated using year and quarter (12 levels) as a raking variable in addition to those specified for the annual samples. The second option (referred to as the "M-weights") was generated using month of intake (12 levels) as a raking variable in place of quarter. Finally, a third set (referred to as "average annual weights") was created for comparison purposes by simply dividing the previously calculated annual weights by 3.

Some results of this analysis are given in Tables E and F. Table E summarizes the coefficient of variation of the weights and corresponding variance inflation factors under the alternative weighting schemes. Not unexpectedly, the M-weights generally have much larger VIFs than the either the Q- or average annual weights. The average annual weights have slightly higher VIFs on average than the Q-weights. Table F summarizes estimates of mean total energy intake and their CVs using the alternative sets of weights. The computed estimates differ trivially among the three sets of weights. The corresponding CVs also vary modestly. For the 30-39, 40-49, 50-59, and 60-69 year age groups for males, the CVs are largest when the M-weights are used. Similar patterns (not shown) appear to obtain for females, and for the low-income groups.

Table E. Comparison of alternative three year weights

Age group (males)	Avg. annual weights		Q-weights		M-weights	
	CV(%)	VIF	CV(%)	VIF	CV(%)	VIF
20 to 29	54.8	1.30	53.9	1.29	63.7	1.41
30 to 39	61.0	1.37	62.7	1.39	69.7	1.49
40 to 49	51.8	1.27	53.5	1.29	70.9	1.50
50 to 59	56.4	1.32	54.3	1.30	73.7	1.54
60 to 69	54.9	1.30	53.0	1.28	67.0	1.45
70+	48.7	1.24	43.2	1.19	51.3	1.26

Table F. Estimates of total energy intake using alternative three-year weights for males

Age group (males)	Est. mean intake			CV (%) of estimate		
	Ann. wts.	Q-wts.	M-wts.	Ann. wts.	Q-wts.	M-wts.
20 to 29	2,825	2,821	2,852	2.25	2.22	2.21
30 to 39	2,667	2,664	2,664	1.69	2.04	1.91
40 to 49	2,444	2,435	2,422	1.53	1.54	1.59
50 to 59	2,270	2,269	2,265	1.47	1.44	1.58
60 to 69	2,073	2,071	2,050	1.44	1.50	1.53
70+	1,830	1,835	1,833	1.48	1.45	1.39

5. References

Goldman, J., Borrud, L., and Berlin, M. (1997), An Overview of USDA's 1994-96 Continuing Survey of Food Intakes by Individuals and the Diet and Health Knowledge Survey, *Proceedings of the Section on Survey Research Methods, American Statistical Association* (forthcoming).

Kish, L. (1992), Weighting for Unequal P_i, *Journal of Official Statistics*, Vol. 8, No. 2, 183-200.

Oh, H. and Scheuren, F. (1978), Multivariate Raking Ratio Estimation in the 1973 Exact Match Study, *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 716-722.

Verma, V. (1993), *Sampling Errors in Household Surveys*, United Nations Department for Economic and Social Information and Policy Analysis, Statistical Division, New York.

Westat (1997), Continuing Survey of Food Intakes by Individuals/Diet and Health Knowledge Survey 1994-96, Cumulative Survey Operations Report, prepared by Westat under contract to the USDA Agricultural Research Service.