WEIGHTING SCHEMES FOR HOUSEHOLD PANEL SURVEYS[1]

Graham Kalton, J. Michael Brick, Westat Inc.
Graham Kalton, Westat Inc., 1650 Research Blvd., Rockville, MD 20850

Key Words: Cross-sectional estimates, fair share weighting, multiplicity weighting

1. Introduction.

National panel surveys of household economics have been mounted in many countries in recent years. The U.S. Panel Study of Income Dynamics (PSID) began in 1968 and has been collecting data on an annual basis since that time (see Hill, 1992). Household panel surveys similar to PSID are in progress or are being planned in most European countries. The U.S. Bureau of the Census started to conduct the Survey of Income and Program Participation (SIPP) in 1983 (Jabine *et al.*, 1990) and Statistics Canada introduced the Survey of Labour and Income Dynamics (SLID) in 1994 (Lavallée *et al.*, 1993).

A common feature to most of these surveys is that they start with a national sample of households, and then follow all the members of those households for the life of the panel. Over the course of time, some members of original sampled households leave those households to set up on their own or to join other households. In order to be able to describe the economic circumstances of sample members at different points of time, household panel surveys usually collect data not only for the sample members but also for the individuals living with sample members at the particular point of time. These individuals are termed associated persons in this paper.

As the panel duration increases, the proportion of associated persons in the sample at a wave rises. For example, in the SIPP, after a year associated persons comprised about 8.6 percent, and after two years they comprised about 12.6 percent, of the sample. With a long-term panel, the proportion of associated persons becomes substantial after several years.

The purpose of this paper is to consider how weights can be developed for the data collected from both original and associated sample persons to be used in producing unbiased (or approximately unbiased) estimates of cross-sectional population parameters. The estimates may relate to households or to individuals at a particular point of time during the course of the panel.

[1]The paper reports research undertaken for the U.S. Census Bureau. The views expressed are the authors'. They do not necessarily reflect the views of the Census Bureau.

In order to prepare for the discussion of weighting schemes, the next section elaborates on the population and household changes that can occur over time, and the types of individual involved. The following sections then discuss weighting schemes that may be used for the different forms of analysis. This discussion relies heavily on previous work by Ernst (1989), Gailly and Lavallée (1993), Huang (1984), Judkins *et al.* (1984), Lavallée and Hunter (1992), and Little (1989).

2. Changes in Population and Household Composition over Time.

In analyzing a panel survey, it needs to be recognized that survey populations change over time. With household panel surveys it is important to distinguish between changes in population composition and changes in household composition.

The composition of a survey population changes over time because some individuals leave the population, some enter the population, and some may leave and join the population more than once. Individuals leave the population through death, emigration, or entering an institution (for surveys of the noninstitutional population). They enter the population through birth (or reaching the specified minimum age), immigration, or leaving an institution.

Households change composition over time for many different reasons, including deaths, births, marriages and divorces. For example, a household at time 1 may contain several individuals who end up in a number of different households at time 2. These individuals may set up new households on their own, they may join individuals who were in one or more households at time 1, or they may join individuals who were not in the population at time 1. One or more of the individuals may leave the population during the intervening period.

Consider a simple sample design in which households are selected independently at time 1 with equal probability. At time 2, the sample of households comprises all the households that contain one or more individuals from the households sampled at time 1, and the sample of individuals at time 2 comprises all the members of the sampled households at time 2. The samples of households and individuals at time 2 are selected with unequal probabilities. For instance, the selection probability of a household at time 2 that contains individuals from three households at time 1 is three times greater than that of a household at time 2 that contains individuals from only one household at time 1. Similarly, the individuals in that household have three times the probability of selection. Thus weighting schemes that compensate for these unequal selection probabilities are needed for the analysis of the resultant data.

Changes in population composition occur when individuals leave or enter the population. An individual sampled at time 1 who leaves the population before time 2 reduces the sample size for time 2 but does not otherwise affect cross-sectional estimates at time 2. However, the situation with regard to entrants is less straightforward. The household panel survey enumeration rule described above incorporates new entrants who join households that contain individuals who were eligible for the initial sample into the population for cross-sectional estimates for later time points. However, new entrants who set up their own households are not represented in later waves of the panel. Equally, households composed of only new entrants are not represented in household analyses at later waves.

The failure of household panel surveys to cover households composed of only new entrants presents a problem for cross-sectional analyses of later waves of the panel. If these households and their members constitute a negligible proportion of the population, the solution to the problem may be to simply ignore it. However, if the proportion is appreciable, as can occur in later waves of a long-term panel, alternative solutions may be called for. One possibility is to add a sample of new entrants (e.g., immigrants) to the panel. This solution is, however, often impracticable. Another solution is to limit the population of inference to persons who were members of the population at the start of the panel. New entrants found living with sample members are then also excluded from the sample. This solution provides a clearcut definition of the population of inference. Whether the solution is appropriate depends on whether that definition can adequately satisfy the survey objectives.

3. Cross-Sectional Estimates for Households.

This section considers weighting schemes that may be used to produce cross-sectional estimates for households for any wave of a household panel survey after the first. At the first wave a sample of households is selected and all the individuals in the sampled households become panel members to be followed throughout the life of the panel or until they leave the survey population. At a subsequent wave, wave t, the household sample comprises all the households in which panel members reside. Households that consist of new entrants only are not represented in the sample at later waves. Such households are ignored here. Complications of nonresponse are also ignored.

Consider the estimation of the total Y for all H households in the population at time t:

$$Y = \sum_{i=1}^{H} Y_i . \tag{3.1}$$

A general estimator for this total can be expressed as

$$\hat{Y} = \sum_{i=1}^{H} w_i Y_i$$

where w_i is a random variable that takes the value $w_i = 0$ if household H_i is not in the sample. The expectation of $\hat{Y}$ is

$$E(\hat{Y}) = \sum_{i=1}^{H} E(w_i) Y_i . \tag{3.2}$$

By comparing equations (3.1) and (3.2), it can be seen that $\hat{Y}$ is unbiased for Y for any weighting scheme for which $E(w_i) = 1$.

There are many ways to satisfy the condition $E(w_i) = 1$. Three will be treated here. First, consider a standard *inverse selection probability weighting scheme*. The probability of a household being in the sample at time t is the probability of one or more of the households at time 1 from which it has drawn members being selected for the original sample. The probability of household H_i being in the sample at time t is then

$$\begin{aligned} P(H_i) &= P(h_j \cup h_k \cup h_\ell \cup ...) \\ &= \Sigma p_j - \Sigma\Sigma p_{jk} + \Sigma\Sigma\Sigma p_{jk\ell} - ... \end{aligned} \tag{3.3}$$

where $P(h_j \cup h_k \cup h_\ell \cup ...)$ is the selection probability of the union of original households h_j, h_k, h_ℓ, etc., for the original sample, p_j is the selection probability of h_j for the original sample, p_{jk} is the joint selection probability of h_j and h_k for the original sample, etc., and where households h_j, h_k, h_ℓ, etc., each contain at least one member who is currently in household H_i. The weight for each sampled household is then $w_i = 1/P(H_i)$. With this weighting scheme,

$$E(w_i) = P(H_i)[1/P(H_i)] + [1-P(H_i)]0 = 1,$$

satisfying the condition for an unbiased estimator of a population total.

In practice, the computation of $P(H_i)$ will generally not be as complex as equation (3.3) might suggest because the number of original households represented in household H_i is usually small. With, say, two original households involved, $P(H_i)$ reduces to

$$P(H_i) = P(h_1 \cup h_2) = p_1 + p_2 - p_{12}. \qquad (3.4)$$

A problem with the application of the inverse selection probability approach is that p_j may be known only for households selected for the original sample, and not for other households. Also the joint probability may not be known. Even when the original sample was an equal probability one, so that all the p_j are the same, the joint probability may depend on the sample design (for instance, whether the two households were in the same segment or not). The difficulty of obtaining $P(H_i)$ is a major drawback with the inverse selection probability approach.

An alternative strategy for developing the weights for time t is to base them only on the selection probabilities of households selected for the original sample, thus avoiding the difficulty in obtaining $P(H_i)$ noted above. One approach is to identify the set of households h_j at time 1 that would result in household H_i being in the sample at time t, and compute the weight for household H_i as

$$w_i = \sum_j \alpha_{ij} w_{ij}^{\prime} \qquad (3.5)$$

where $w_{ij}^{\prime} = 1/p_j$ if household h_j, which has at least one member in household H_i, was selected for the original sample and $w_{ij}^{\prime} = 0$ if not, and where α_{ij} are any set of constants satisfying $\sum_j \alpha_{ij} = 1$.

With this approach,

$$E(w_{ij}^{\prime}) = p_j(1/p_j) + (1-p_j)0 = 1,$$

and hence $E(w_i) = \sum_j \alpha_{ij} = 1$. Thus, the use of weights w_i will yield unbiased estimators of totals for the household population for any choice of constants α_{ij}, provided that $\Sigma\alpha_{ij} = 1$. As indicated above, the principal advantage of this type of scheme is that it requires information only on the initial selection probabilities of the original households that were sampled at time 1, which are known. It does not require information on the initial selection probabilities of the other original households that have members in the current household, which are often not known.

A natural choice of α_{ij} is to make them equal for all the original households that lead to the selection of household i at time t. Huang (1984) terms this scheme a multiplicity approach. Here this scheme will be called an *equal household weighting scheme*. With the equal household weighting scheme

$$w_i = \Sigma w_{ij}^{\prime}/C_i, \qquad (3.6)$$

where C_i is the number of original households represented in household H_i at time t.

An alternative version of the above approach is one based on original sample persons rather than households. In this case, let I_{ijk} denote individual k from original household h_j in household H_i. Then

$$w_i = \sum_j \sum_k \alpha_{ijk} w_{ijk}^{\prime}$$

where $w_{ijk}^{\prime} = 1/p_j$ if individual k in household h_j was in the original sample and $w_{ijk}^{\prime} = 0$ if not, and where the α_{ijk} are any set of constants satisfying $\sum_j\sum_k \alpha_{ijk} = 1$. Since the probability of an individual being selected for the original sample is the same as that of that individual's household,

$$E(w_{ijk}^{\prime}) = p_j(1/p_j) + (1-p_j)0 = 1.$$

In this case, the natural choice of the constants α_{ijk} is to make them equal for all members of the current household who were eligible for selection for the original sample. This produces what has been termed the fair share weighting scheme (Huang, 1984; Ernst, 1989). This scheme is termed here an *equal person weighting scheme*. With this scheme

$$w_i = \frac{1}{M_i} \Sigma M_{ij} w_{ij}^{\prime}$$

where $w_{ij}^{\prime} = w_{ijk}^{\prime}$ is constant for all individuals in household H_i emanating from the same sampled household at time 1, M_{ij} is the number of individuals in household H_i coming from household h_j, and $M_i = \sum_j M_{ij}$ is the number of individuals in household H_i who were eligible for the sample at time 1. The equal person weighting scheme is applied in the SIPP and is proposed for use in the SLID.

Although developed here in terms of persons rather than households, it is readily apparent that the equal person weighting scheme could equally have been generated in terms of households. As shown above the household weight $w_i = \sum_j \alpha_{ij} w_{ij}^{\prime}$ satisfies the condition $E(w_i) = 1$ for any set of constants α_{ij} such that $\sum_j \alpha_{ij} = 1$. The choice $\alpha_{ij} = M_{ij}/M_i$, with $\sum_j \alpha_{ij} = 1$, leads to the equal person weighting scheme.

It is instructive to compare the inverse selection probability weighting scheme with the equal household and equal person weighting schemes in a simple case. Following Little (1989), consider household H_i selected at time t with household members coming from two original households. Let p_1 and p_2 denote the selection probabilities for the original households, and let p_{12} denote their joint selection probability. Under the inverse selection probability approach, the household weight is

$$w_i^* = \frac{1}{p_1 + p_2 - p_{12}}.$$

Under the equal person weighting scheme the weight for household H_i depends on which household or households were selected for the original sample. With P_1 and P_2 being the proportions of members of household H_i who came from households h_1 and h_2, respectively (excluding any new entrants to the population), $w_i = P_1/p_1$ if only h_1 was selected; $w_i = P_2/p_2$ if only h_2 was selected; and $w_i = (P_1/p_1)+(P_2/p_2)$ if both h_1 and h_2 were selected. The probability of only h_1 being selected is (p_1-p_{12}), of only h_2 being selected is (p_2-p_{12}), and of both households being selected is p_{12}. The expected value of the weight conditional on household H_i being in the sample is thus

$$\frac{(p_1-p_{12})(P_1/p_1)+(p_2-p_{12})(P_2/p_2)+p_{12}[(P_1/p_1)+(P_2/p_2)]}{p_1+p_2-p_{12}}$$

i.e.,

$$E(w_i|H_i \text{ in sample}) = \frac{1}{p_1+p_2-p_{12}} = w_i^*.$$

As this result demonstrates, the weight for household H_i varies depending on which original households were selected, but in expectation the weight is the same as that obtained from the inverse selection probability approach.

Results for the expectation of the weight of H_i under the equal household weighting scheme can be readily obtained as a special case of the above derivation in which $P_1 = P_2 = ½$. In expectation, the weight is the same as that for the inverse selection probability approach.

Given that the weight $w_i = \Sigma\alpha_{ij}w_{ij}^{/}$ satisfies the condition $E(w_i) = 1$ for any set of α_{ij} such that $\Sigma\alpha_{ij} = 1$, the question arises as to the optimal choice of the α_{ij}. One approach is to choose the α_{ij} to minimize the variance of $\hat{Y}$.

The variance of $\hat{Y}$ may be expressed as

$$V(\hat{Y}) = VE(\hat{Y}|s) + EV(\hat{Y}|s) \qquad (3.7)$$

where s denotes the set of households in the sample at time t. Now

$$E(\hat{Y}|s) = \sum^{s} E(w_i|H_i)Y_i = \sum w_i^* Y_i = \hat{Y}^*,$$

where $\hat{Y}^*$ is the standard inverse selection probability estimator. Hence

$$VE(\hat{Y}|s) = V(\hat{Y}^*).$$

The first term in equation (3.7) is thus the variance of the standard inverse selection probability estimator, and the second term is the additional variance resulting from the use of weighting schemes from the class (3.5), $w_i = \Sigma\alpha_{ij}w_{ij}^{/}$. The α_{ij} may then be chosen to minimize $EV(\hat{Y}|s)$.

Consider the simple case discussed above in which H_i is composed of members from two original households, and let $w_i = \alpha_i w_{i1}^{/} + (1-\alpha_i)w_{i2}^{/}$. Then the optimum value of α_i, i.e., the value that minimizes $EV(\hat{Y}|s)$, is

$$\alpha_{oi} = \left(1 + \frac{p_2-p_{12}}{p_1-p_{12}}\right)^{-1}. \qquad (3.8)$$

In the special case of an equal probability (epsem) sample of households initially, with $p_1 = p_2$, $\alpha_{oi} = \frac{1}{2}$. Thus, in the two-household case, the equal household weighting scheme minimizes the variance of the household weights around the inverse selection probability weight when the initial sample is an epsem one.

The optimal choice of α_{oi} given by (3.8) requires knowledge of p_1, p_2 and p_{12}. If these probabilities were known, then the standard inverse selection probability weight could be employed and would be preferable. In the case of an approximately epsem sample, the equal household weighting scheme seems a reasonable choice, at least for the two-household case.

4. Cross-Sectional Estimates for Individuals.

This section considers weighting schemes that may be used to produce cross-sectional estimates for individuals for any wave of a household panel survey after the first. At a subsequent wave, wave t, the survey population has changed: some members of the original population will leave and some new entrants will join the survey population in the period from wave 1 to wave t.

Let there be N individuals in the population at time t, with N_i individuals in household H_i $(i = 1,2,...H)$ and $\Sigma N_i = N$ (new entrants living in households containing no members of the original population are not included here, and are ignored throughout the rest of this section). The members of household H_i come from households h_j, h_k, h_ℓ, etc., at time 1. Let M_{ij} denote the number of members of household i at time t who were in household h_j at the start of the panel. The sum $M = \Sigma\Sigma M_{ij}$ is less than the population size at time 1 because of leavers from the population in the period from time 1 to time t, and $M < N$ because of new entrants to the population who are in households containing members from the original population.

Consider now the estimation of a total for the population of individuals at time t:

$$Y = \sum_{i=1}^{H}\sum_{k=1}^{N_i} Y_{ik}, \qquad (4.1)$$

where Y_{ik} is the value for individual k in household

H_i. As in the household case discussed in the previous section, a general estimator for this total can be expressed as

$$\hat{Y} = \sum_{i=1}^{H} \sum_{k=1}^{N_i} w_{ik} Y_{ik} \qquad (4.2)$$

where w_{ik} is a random variable that takes the value $w_{ik} = 0$ if individual k in household H_i is not in the sample. The estimator $\hat{Y}$ is unbiased for Y provided that $E(w_{ik}) = 1$.

There are many ways to satisfy the condition $E(w_{ik}) = 1$. It is instructive to consider three of them. First, let $w_{ik} = 0$ for all individuals not in the original sample. In this case, the estimator $\hat{Y}$ discards associated persons. Let p_{ik} denote the probability of a member of the original population, individual k residing in household H_i at time t, being selected for the initial sample, and let $w_{ik} = 1/p_{ik}$. Then, for such an individual

$$E(w_{ik}) = p_{ik}(1/p_{ik}) + (1-p_{ik})0 = 1.$$

With this scheme, all new entrants to the population have $w_{ik} = 0$ with certainty. Thus $\hat{Y}$ in (4.2) provides an unbiased estimator of the total for the original population that is still present at time t, but does not include a component for the new entrants.

Modifications to the above procedure can be made to cover certain types of new entrants. For instance, births to sampled mothers can be included by assigning them the weight of their mothers, or if, as in the SIPP, the survey population is taken to be adults aged 16 and over, those under 16 at the start of the panel can be treated as sampled persons with assigned probabilities, and they can be included in the analyses of later waves after they have attained the age of 16. Such modifications do not, however, handle all types of new entrants. Provided that the proportion of other types of new entrants is small, this deficiency may not be a serious concern.

The weighting scheme that restricts the analysis to original sample persons, plus certain specified new entrants, is employed with the PSID. Its limitation is that it fails to make direct use of data collected for associated persons. Such data may be used to provide information on the situation of sample persons, but the associated persons are excluded from the sample for the analysis.

In order to include associated persons in cross-sectional analyses for time t, they need to be assigned positive weights. Noting that the probability of an individual being selected for the sample is the same as that of his or her household, weighting schemes for cross-sectional analyses of individuals at wave t can be obtained directly from those for households given in Section 3. Here we will develop the general strategy of producing weights for cross-sectional analysis at time t based only on the selection probabilities of members of the original sample, thus avoiding the problems with the inverse selection probability approach noted in Section 3.

Let I_{ijk} denote individual k from original household h_j who is now in household H_i. Let w_i denote the weight for every member of household H_i for cross-sectional analyses at time t, and let

$$w_i = \sum_j \sum_k \alpha_{ijk} w'_{ijk}$$

where $w'_{ijk} = 1/p_j$ if household h_j was in the original sample and $w'_{ijk} = 0$ if not. Then, as before, $E(w'_{ijk}) = 1$ for members of the original population. New entrants, for whom $p_j = 0$, may be handled by setting $\alpha_{ijk} = 0$. Then

$$E(w_i) = \sum_j \sum_k^{N_i} \alpha_{ijk} E(w'_{ijk}) = \sum_j \sum_k^{M_i} \alpha_{ijk} = 1$$

provided that $\sum\sum \alpha_{ijk} = 1$. Under this condition $\hat{Y}$ is unbiased for Y.

A natural choice of α_{ijk} is to set $\alpha_{ijk} = 1/M_i$ for all members of the original population. This is the equal person weighting scheme in which every member of household H_i at time t (including new entrants) receives the weight

$$w_i = \sum_j \sum_k w'_{ijk}/M_i .$$

Another choice of the α_{ijk} is that used for the equal household weighting scheme. Let C_i denote the number of original households that have members in household H_i at time t. Then $\sum_j \sum_k \alpha_{ijk} = 1$ can be divided equally between households, with each member of original household h_j being assigned a value of $\alpha_{ijk} = 1/C_i M_{ij}$. Then for original household h_j

$$\sum_k^{M_{ij}} \alpha_{ijk} = 1/C_i .$$

The derivation of the α_{ijk} to minimize the variance of the estimated total $\hat{Y}$ for the population of individuals follows directly from the corresponding derivation for the population of households. The estimated total for the population of individuals is

$$\hat{Y} = \sum_i^{s} \sum_k^{N_i} w_{ik} Y_{ik} = \sum_i^{s} \sum_k^{N_i} w_i Y_{ik}$$

since the weights for every individual in sampled household H_i are the same. This estimated total can be expressed as

$$\hat{Y} = \sum_{i}^{s} w_i Y_i,$$

where $Y_i = \Sigma Y_{ik}$ is the household total for H_i. Thus $\hat{Y}$ can be expressed as a household total, and the results of Section 3 can be applied directly.

Consider the example from Section 3 in which H_i is composed of members from only two original households, perhaps together with one or more new entrants. Then the person-level weight $w_i = \Sigma\Sigma\alpha_{ijk}w'_{ijk}$ reduces to $w_i = \alpha_i w'_{i1} + (1-\alpha_i)w'_{i2}$, where $w'_{ijk} = w'_{ij}$ and where $\alpha_i = \Sigma\alpha_{i1k}$. The optimum value of α_i is given by equation (3.8). The individual values α_{ijk} are not needed for computing the w_i; only the original household totals $\sum_k \alpha_{ijk}$ are required. If individual values are needed for the α_{ijk}, they may be simply assigned as $\sum_k \alpha_{ijk}/M_{ij}$.

As in the household case, the optimum weighting α_{oi} requires knowledge of p_1, p_2 and p_{12}. If these probabilities are known, the standard inverse selection probability weight w_i^* can be computed, and would be preferred. In the case of an approximately epsem sample, the equal household weighting scheme should fare well.

5. Summary and Concluding Remarks.

This paper has described weighting schemes that enable all households for which, and all individuals for whom, data are collected in the later waves of a household panel survey to be included in cross-sectional analyses of those waves. These weighting schemes can accommodate new entrants to the population who move in to live with members of the original population, but not other new entrants.

The usual inverse selection probability weighting scheme requires information on the household selection probabilities of all members of the households sampled at a later wave, as well as the joint selection probabilities of the original households that contribute members to the later wave households. The inverse selection probability weighting scheme can often not be applied because these probabilities are unknown. To deal with this problem, an alternative approach that requires information on only the selection probabilities of sampled original households is described.

This alternative approach produces a class of weighting schemes including the equal person (fair share) scheme used in SIPP and the equal household weighting scheme. All the schemes in this class produce weights that are in expectation equal to those produced by the usual inverse selection probability scheme. The variance in the weights around the inverse selection probability weights gives rise to an increase in the variance of the survey estimates. When the original households are selected with approximately equal probability, the equal household weighting scheme seems a reasonable choice for both household and individual level analyses to control this increase in variance.

The class of weighting schemes described has a broader range of application than that indicated here. It can in fact be usefully applied in any situation where an inverse selection probability weighting scheme would be appropriate, but where not all the inclusion probabilities and joint inclusion probabilities are known. Such situations occur frequently when multiple frames are used for sample selection.

6. References.

Ernst, L.R. (1989). Weighting Issues for Longitudinal Household and Family Estimates. In *Panel Surveys*, eds. D. Kasprzyk, G. Duncan, G. Kalton and M.P. Singh, New York: John Wiley, pp. 139-159.

Gailly, B. and Lavallée, P. (1993). *Inserer des Nouveaux Membres dans un Panel Longitudinal de Menages et D'Individus: Simulations*. Walferdange, Luxembourg: CEPS/Instead.

Hill, M.S. (1992). *The Panel Study of Income Dynamics: A User's Guide*. Newbury Park, CA: Sage Publications.

Huang, H. (1984). Obtaining Cross-Sectional Estimates from a Longitudinal Survey: Experiences of the Income Survey Development Program. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 670-675.

Jabine, T.B., King, K.E. and Petroni, R.J. (1990). *Survey of Income and Program Participation: Quality Profile*. Bureau of the Census, Washington D.C.: U.S. Department of Commerce.

Judkins, D., Hubble, D., Dorsch, J., McMillen, D., and Ernst, L. (1984). Weighting of Persons for SIPP Longitudinal Tabulations. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 676-687.

Lavallée, P. and Hunter, L. (1992). Weighting for the Survey of Labour and Income Dynamics. *Proceedings of Statistics Canada Symposium 92: Design and Analysis of Longitudinal Surveys*, 65-75.

Lavallée, P., Michaud, S. and Webber, M. (1993). The Survey of Labour and Income Dynamics, Design Issues for a New Longitudinal Survey in Canada. *Bulletin of the International Statistical Institute*, 49th Session, Contributed Papers, Book 2, 99-100.

Little, R.J.A. (1989). Sampling Weights in the PSID: Issues and Comments. Panel Study of Income Dynamics Working Paper, Ann Arbor: University of Michigan.