Journal of Official Statistics
Vol. 2. No. 3, 1986, pp. 303–314

Handling Wave Nonresponse in Panel Surveys

Graham Kalton[1]

Abstract: Panel surveys are subject to wave nonresponse which occurs when responses are obtained for some but not all waves of the survey. While weighting adjustments are routinely used to compensate for total nonresponse and imputations are used for item nonresponse, the choice of compensation procedure for wave nonresponse is not obvious. The choice depends on a number of factors, including: the number of waves of missing data; the type of analysis to be conducted; the availability of auxiliary variables with high predictive power for the missing values; and the work involved in implementing the procedures. The paper reviews the issues involved in compensating for wave nonresponse.

Key words: Nonresponse; weighting adjustments; imputation; panel surveys; panel attrition.

1. Introduction

Textbook discussions of missing data in surveys generally make only the simple distinction between unit (or total) nonresponse and item nonresponse, the former arising when no data are collected for a sampled unit and the latter when responses are obtained to some but not all of the survey items. The choice of procedures for attempting to compensate for nonresponse is then reasonably straightforward. As a rule weighting adjustments are used for unit nonresponse and imputation for item nonresponse.

This paper is concerned with the more complex situation of missing data in panel surveys, and in particular in the Survey of Income and Program Participation (SIPP). There are two features of the SIPP that complicate the simple distinction between unit and item nonresponse, and in consequence raise questions about the appropriate choice of compensation procedure for certain types of nonresponse. The main feature is that the survey is a panel survey that collects data from the same units in eight different waves. The second feature is that the SIPP collects data for all persons aged 15 and over in sampled households; the units of analysis are persons for some analyses, while for others they are households, families or other groupings of persons.

Units failing to respond on any wave in a panel survey clearly constitute unit nonresponse, and weighting adjustments may be employed in an attempt to compensate for them. Equally, missing responses to certain items from units that respond on all waves are item nonresponses which may be handled by

[1] Survey Research Center, University of Michigan, Ann Arbor, Michigan 48106-1248, USA. This paper is a slightly revised version of a paper given at the U.S. Bureau of the Census First Annual Research Conference, March 20–23, 1985. The author would like to thank the referees for their helpful suggestions.

imputation. The complication in a panel survey is that there are units that respond to some but not all waves of data collection. From a longitudinal perspective, wave nonresponse may be viewed as a set of item nonresponses in the longitudinal record, suggesting that imputation may be the appropriate compensation procedure. From a cross-sectional perspective, it may be viewed as unit nonresponse, for which a weighting adjustment may be appropriate.

Some missing data issues arising in household sampling mirror those raised by the panel design. In a cross-sectional survey, sample households in which no one responds clearly count as unit nonresponse, and missing responses to certain items in households in which data are collected for all eligible persons are clearly item nonresponses. The complication is how to treat cases where no data are collected for one or more persons in an otherwise cooperating household. For household-level analyses, such individual nonresponse may be viewed as a set of item nonresponses in the household record, suggesting that imputation may be used in compensation. For individual-level analyses, it may be viewed as unit nonresponse, which may be handled by a weighting adjustment.

This paper focuses on the question of what form of compensation procedure should be used to attempt to compensate for wave nonresponse. The next section reviews the general issues involved in the choice between weighting adjustments and imputation for handling missing survey data. The following two sections then discuss some special features that arise in the application of these procedures to panel surveys. The final section presents some concluding remarks. Where possible, the discussion is illustrated with data from the Income Survey Development Program's (ISDP's) 1979 Research Panel, a prototype for the SIPP (Ycas and Lininger (1981)).

2. Weighting or Imputation

Although weighting and imputation are often thought of as entirely distinct methods of compensating for missing survey data, they are in fact closely related for univariate analysis (Kalton (1983), Little (1984), Oh and Scheuren (1983)). As a simple illustration, consider the imputation scheme in which the sample is divided into adjustment cells based on auxiliary information available for both the respondents and nonrespondents to a given item, and then a nonrespondent is assigned the same response for that item as a respondent in the same cell. For univariate analyses, this imputation scheme is equivalent to the weighting scheme that adds the weight of the nonrespondent to that of the respondent who in the imputation scheme donated the imputed value: the distribution of respondent and imputed values from the imputation scheme is the same as the weighted distribution of respondent values from the weighting scheme, and hence summary statistics such as the mean and variance are also the same.

While this relationship between weighting and imputation is instructive, it nevertheless hides some major differences between the two procedures. For one, weighting does not require the assignment of increased weights to a sample of respondents as in the above example. Instead, fractional weights can be spread evenly across the respondents in a cell. This even spread of weights avoids the increase in the variances of survey estimates associated with the sampling of respondents. With imputation this increase is less easily avoided; however, it can be reduced to minor magnitude by the use of appropriate methods of sampling respondents to serve as donors (Kalton and Kish (1984)) or by the use of multiple imputations (Rubin (1979)).

The major differences between weighting and imputation stem not from this issue of sampling respondents but rather from the multivariate nature of survey data. Surveys are not concerned with a single variable as in the above example, but rather with many variables. This feature has a number of consequences for both weighting and imputation, and explains why unit nonresponse is generally treated by weighting and item nonresponse by imputation.

Usually the values of only a few survey design variables (e.g., strata, PSUs) are known for unit nonrespondents. These variables can all–or nearly all–be incorporated into the construction of the adjustment cells. The cells, reflecting everything that is known about the nonrespondents, can thus be used to predict all the missing survey variables as effectively as possible. This is efficiently done by increasing the weights of the respondents in the cells so that they represent the nonrespondents also.

In the case of item nonresponse, however, a great deal more is known about the nonrespondents. It is therefore rarely possible to find a respondent who exactly matches a nonrespondent in terms of all the data available for the nonrespondent. In this circumstance, three alternative approaches are possible:

(1) Discard enough of the less important data about the nonrespondents to enable matches to be made and a cell weighting adjustment to be used;

(2) Attempt to incorporate the important data about the nonrespondents in a model of response propensities, which can then be used to develop weighting adjustments;

(3) Employ an imputation procedure to assign values for the missing responses.

The first approach may be appropriate for nonrespondents for whom only limited data are available. Discussion of the second approach is deferred to the next section. An important difference between the weighting and imputation approaches for item nonresponse is that with weighting some of the data for item nonrespondents has to be discarded, whereas with imputation the nonrespondents' responses to other items are retained intact. This is an obvious advantage of imputation; however, it also has some undesirable consequences.

Weighting has the notable advantage over imputation that it preserves the observed associations between the survey variables. Increasing the weight of a respondent record by an amount w can be regarded as the creation of a new record taking the complete set of variables from the one respondent record and giving the new record a weight of w. Thus the relationships between the survey variables in the respondent record are reproduced in the new record. Imputation, however, fails to have this desirable property. In general, imputation preserves the covariances of a variable subject to imputation with the auxiliary variables used in the imputation scheme, but attenuates the covariances with other variables (Santos (1981), Kalton and Kasprzyk (1982)). Unless safeguards are taken, imputed values may even turn out to be inconsistent with other responses on the record. Since most of survey analysis involves studying relationships between variables, such as by crosstabulation and regression analysis, this failure of imputation to preserve covariances is a serious disadvantage.

Another concern with imputation is that it fabricates data to some extent. There is the risk that analysts will treat the imputed values as real values and compute sampling errors accordingly. They will thus attribute greater precision to the survey estimates than is justified. The extent of fabrication depends on the situation. If there is some redundancy in the survey data so that a missing response can be deduced without error from other responses, the imputation involves no fabrication. If the variable subject to imputation is highly correlated with the auxiliary variables used in the imputation scheme, the amount of fabrication is small. If, however, the variable subject to imputation is only slightly correlated with the auxiliary variables, the amount of fabrication is sizable. Often, the situation corresponds most closely to the last of these alternatives. The amount of fabrication also affects the attenuation of covariances: the larger the amount of fabrication, the greater the degree of attenuation.

Another important difference between weighting and imputation is that weighting is a global strategy, treating all variables simultaneously, whereas imputation can be item-specific. To the extent that there is a choice of auxiliary variables to use for the global weighting adjustments, the choice is mainly made in terms of their ability to predict the response propensities. For instance, adjustment cells are generally determined to compensate for differences in response rates across different subgroups of the sample.

On the other hand, the choice of auxiliary variables to use in imputing for a specific variable is generally governed by their abilities to predict that variable since, as noted above, the greater their predictive power the smaller are the problems of covariance attenuation and fabrication. (Sometimes a slight modification is made to this choice to deal with the problem of several associated missing items on a given record. If each imputation was conducted independently, the covariances between these variables would be attenuated. This problem can be dealt with by imputing for the several missing items from the same donor: this can be readily done if the same set of auxiliary variables is used for the several items.)

A factor to be taken into account in choosing between weighting and imputation is the auxiliary information available for use in making the nonresponse adjustments. Weighting tends to be favored when the auxiliary variables are only weakly related to the variables with the missing values, because imputation gives rise to serious problems of fabrication of data and attenuation of covariances in this case. On the other hand, imputation tends to be favored when auxiliary variables with high predictive powers for the variables with missing values are available; in this case the problems of fabrication of data and attenuation of covariances are less significant, and imputation can make much more effective use of the auxiliary information than can weighting.

Having reviewed the general issues relating to the choice between weighting and imputation, we now address the specific issue of handling wave nonresponse in a panel survey. The next section discusses the use of weighting adjustments for this purpose and the following one discusses the use of imputation.

3. Weighting Adjustments for Wave Nonresponse

Panel surveys are subject to many forms of analysis. Some analyses yield cross-sectional estimates from a single wave while others relate variables across two or more waves (e.g., measuring changes between waves or adding four-monthly income components across three waves to produce annual totals). In conducting such analyses, it needs to be recognized that the population is dynamic, changing its composition between waves as "births" and "deaths" occur (Kasprzyk and Kalton (1983), Kalton and Lepkowski (1985)). This feature itself can lead to complications in the weights used, but for simplicity we will ignore these complications by treating the population as essentially static. We will further assume that the sample elements are selected with equal probability so that no sampling weights are required. We will thus be concerned only with the development of weights to compensate for total and wave nonresponse.

For illustrative purposes, consider a three-wave survey (as, for instance, will occur when the first three waves of a SIPP panel are merged to create an annual file). There are then eight different patterns of response/nonresponse for the sampled units. Denoting response as 1 and nonresponse as 0, these eight patterns are;

111	110	101	011
100	010	001	000

The last pattern represents the total nonrespondents; for any form of analysis a weighting adjustment can be made for them. For a particular form of analysis, the patterns that provide the requisite data can be identified, and weights can be developed to compensate for the sample units in the other patterns. Thus, for instance, sample units in patterns 111, 110, 101 and 100 provide data for a cross-sectional analysis of wave 1 and they can be weighted up to compensate for units in the other four patterns; similarly, sample units in patterns 111 and 011 provide data for measuring changes between the second and third waves, and they can be weighted up to compensate for the units in the other six patterns. There are potentially seven combinations of waves for different forms of analysis, thus implying the need for seven different sets of weights.

Little and David (1983) distinguish three types of wave nonresponse: attrition, reentry, and late entry. Attrition nonresponse occurs when a unit drops out of the survey at one wave and remains out thereafter, reentry occurs when a unit drops out for one or more waves but reenters at a later point, and late entry occurs when a unit is not interviewed at the first wave but enters later. With a three-wave panel, the patterns 110, 100 and 000 constitute attrition nonresponse, the pattern 101 constitutes reentry, and the patterns 011 and 001 constitute late entry. There is also the possibility of dropping out more than once: the pattern 010 represents a late entry which drops out later.

If all the missing wave data were in the form of attrition nonresponse, the resultant data would form a nested pattern, with fewer of the same set of respondents at each successive wave. With only four of the above patterns arising, namely 111, 110, 100 and 000, just three sets of weights are needed. There would be one set of weights for each wave; these could be used straightforwardly for cross-sectional analysis, and any analysis involving more than one wave would employ the weight of the latest wave used in that analysis. With more waves of data, the reduction in the number of sets of weights required based on all patterns of wave nonresponse to the number based on attrition nonresponse only is more substantial. For instance, making allowance for analyses of all possible combinations of wave data from the eight waves of a SIPP panel would require $2^8-1 = 255$ sets of weights with all possible patterns of wave nonresponse, but just 8 sets when only attrition nonresponse occurs.

Little and David propose a method for developing weights to compensate for attrition nonresponse that attempts to take account of all the auxiliary data available at each successive wave. At the first wave, the only auxiliary data available for both the nonrespondents and the respondents are the design variables z, such as strata and PSUs. These may be employed to form adjustment cells, using the inverses of the response rates within the cells as the weights, or else the response indicator ($r = 1$ for a respondent, $r = 0$ for a nonrespondent) can be regressed on the design variables, using a logistic or probit regression, with the weights for the respondents then being the inverses of the predicted means from the regression for their specified values of z.

The auxiliary variables available for units lost at the second wave are both the z variables and their responses at the first wave, x_1. Little and David propose regressing the response indicator for wave 2 on these auxiliary variables, z and x_1, for all the sampled units that responded at the first wave. The inverses of the predicted means from this regression then give the adjustments needed to compensate for the loss from the first to second waves. Thus the overall weight for the second wave respondents is $w_2 = w_1 w_{2.1}$, where w_1 is the weight for the first wave and $w_{2.1}$ is this further adjustment.

For the third wave, the auxiliary data comprises z, x_1 and the responses at the second wave, x_2. The regression of the response indicator for the third wave is then run for all those units that responded at the second wave, and the inverses of the predicted means are used for the further adjustment to compensate for units lost at the third wave; i.e., the weight at the third wave is $w_3 = w_2 w_{3.12}$. The same procedure is used for all subsequent waves.

Unfortunately, the simplicity of the above procedure is lost when non-attrition losses are included. In practice there are likely to be a fair number of non-attrition cases. Table 1 gives the relative frequency of the various response patterns (excluding the total nonrespondents, pattern 000) for the first three waves of the ISDP 1979 Research Panel. As can be seen from the table, 80.2% responded on all three waves, 13.9% were attritors and 6.0% were non-attritors. Little and David provide a corresponding table for persons who responded to at least one of the first five waves of the ISDP 1979 Research Panel. Of such persons, 74% responded to all five waves, 15% were attritors, and 11% nonattritors.

Little and David describe a weighting scheme for the non-nested situation, but the scheme has some unattractive features. As a simple illustration, consider a two-wave panel with a respondents to both waves, b respondents to the first but not the second wave, c respondents to the second but not the first wave, and d nonrespondents to both waves.

Table 1. Person Response/Nonresponse in the First Three Waves of the 1979 ISDP Research Panel (Excluding Total Nonrespondents)

Response (1)/Nonresponse (0)	%
Respondents	
111	80.2
Attritors	
110	7.2
100	6.7
Non-attritors	
101	2.3
011	2.2
010	0.6
001	0.9
Total	100.0
Number of persons	20676

For wave 1 cross-sectional analyses, the $(a+b)$ first-wave respondents are weighted up using the design variables z as the auxiliary variables. For wave 2 cross-sectional analysis, the set of a respondents at wave 2 are weighted up to represent the b nonrespondents, using the z and first-wave variables as auxiliary variables, and the set of c respondents at wave 2 are weighted up to represent the d nonrespondents, using just the z variables as auxiliary variables. Longitudinal analysis of waves 1 and 2 combined are conducted with the set of a respondents to both waves, weighted up by the product of the cross-sectional weights. Note that, in determining the wave 1 weights, this scheme does not utilize the responses available for the c respondents at wave 2 who fail to respond at wave 1. These responses could be incorporated by performing a reverse weighting scheme like the forward one for wave 2, using auxiliary data from wave 1 where available, but then the longitudinal weight would not be the simple product of the cross-sectional ones.

The scheme involves matching respondents and nonrespondents in terms of their response patterns on previous waves (e.g., the fourth-wave nonrespondents with the pattern 1010 are matched with respondents with the pattern 1011), and then weighting up the respondents to represent the nonrespondents. If the number of respondents of a matched pattern is small and the number of nonrespondents large (as might well occur with the patterns 1001 and 1000), that set of respondents will have a large weight. The resulting wide variation in weights would have an adverse effect on the precision of the survey estimates. To avoid this effect, it may be advisable to sacrifice some of the earlier wave data, for instance matching respondents 1101 and 1001 together with 1000, ignoring the second-wave responses in the first of these respondent patterns, or forcing non-attrition response patterns into nested patterns by ignoring responses to waves after a missing wave (e.g., treating 1101 as 1100).

The development of wave nonresponse weights that attempt to take into account all the auxiliary information available from other waves is clearly a substantial task, but probably much less extensive than the task required for imputation.

4. Imputing for Wave Nonresponse

Imputation assigns values for missing responses by making use of auxiliary variables. In general, the value imputed for the ith nonrespondent on variable y is

$y_i = f(x_{1i}, x_{2i}, \ldots, x_{pi}) + e_i$, where $f(x)$ is a function of the p auxiliary variables and e_i is an estimated residual. Often $f(x)$ is a linear function $\beta_0 + \Sigma\beta_j x_{ji}$, and the β's are estimated from the respondents' data. This formulation covers regression imputation in an obvious way and also cell imputation – such as the widely used hot-deck procedure – by defining the x's as dummy variables to represent the cells. If the e_i are set at zero, the imputation scheme may be termed a deterministic one; if the e_i are estimated residuals, the imputation scheme may be termed a stochastic one. See Kalton and Kasprzyk (1982) for further discussion.

The auxiliary variables for use in imputing for wave nonresponse are the survey design variables and the responses to items on other waves. In most panel surveys, many of the same items are repeated at each wave. When the responses to a repeated item are highly correlated over time, the response on one wave will be a powerful predictor of a missing response on another wave. Kalton and Lepkowski (1983) found, for example, that for respondents reporting hourly rates of pay on each of the first two waves of the ISDP 1979 Research Panel, the correlation between the two rates was 0.97. This suggests that if a person's hourly rate of pay is available for one wave but that person is a nonrespondent on an adjacent wave, the missing rate can be imputed almost without error. Note, however, that a high correlation for the respondents does not guarantee that the nonrespondents' values will be predicted well. It could be, for example, that the respondents' rates of pay remain the same on the two waves, giving a correlation of 1, but that the nonrespondents' rates change between waves. The use of the respondents' correlation to measure the predictive power for nonrespondents depends on the assumption that, conditional on the auxiliary variables, the missing values are missing at random.

Kalton and Lepkowski describe a variety of procedures that can be employed for crosswave imputation in a two-wave panel, using the value of the variable on one wave to impute the missing value of the same variable on the other. One such procedure is hot-deck imputation. For instance, in imputing for hourly rate of pay on wave 2, hourly rate of pay on wave 1 would be categorized into a number of cells, and an individual with a missing wave 2 rate would then be assigned the wave 2 rate of an individual who came from the same wave 1 cell. When the variable's crosswave correlation is extremely high, however, the categorization into cells throws away valuable information: a wave 2 nonrespondent at one end of a wave 1 cell may be matched with a wave 2 respondent from the other end of the cell. While this loss of information may be reduced by increasing the number of cells, the number of cells that can be used is limited by the need to ensure that matches can be made.

The categorization with the hot-deck procedure can be avoided by using some form of regression imputation. Thus, for instance, the imputed hourly rate of pay of individual i on wave 2 (y_i) may be obtained from the regression $y_i = a + bx_i + e_i$, where x_i is the individual's wave 1 hourly rate of pay and e_i is a residual term. Regression imputation can be viewed as constructing a new variable, the predicted value $a + bx_i$, for all individuals in the second wave. The values of the errors e_i can then be calculated for the respondents, and the imputation problem reduces to assigning e_i values for the nonrespondents. The e_i may be set to zero, as in deterministic imputation, or they may be assigned in a variety of ways, such as a hot-deck imputation procedure, using the variable in question or other variables as the auxiliary variables for creating the cells. The selection of the residuals for several variables from the same donor will help to maintain the relationships between the variables.

One way to choose the values of a and b is to use the least-squares estimates obtained from the regression based on those who responded on both waves. Sometimes it may be appropriate to force the regression through the origin, setting $a = 0$; this is then a model of proportionate change. An alternative model is to set $b = 1$, which is a model for additive change. The proportionate and additive change models are simple to implement. For variables that are extremely stable over time, the simple imputation of directly substituting the value on one wave for the missing value on the other may serve well for many purposes. This is the special case of regression imputation with $a = 0$, $b = 1$ and $e_i = 0$. However, this procedure suffers from the disadvantage that it understates the amount of change between waves, and measurement of change is often of interest in panel surveys. This understatement can be avoided by using the stochastic imputation model $y_i = x_i + e_i$, where e_i is assigned from some respondent. If the variable is very stable, the assigned e_i will mostly be 0, but nonzero values will occur when donors have values that change between waves.

The above regression imputation procedures are applicable for continuous variables. One possible wave nonresponse imputation procedure for categorical variables is to assign the modal response category among respondents who gave the same response to the variable on the other wave. As an illustration, consider a respondent who reported his work status in the first wave of the ISDP 1979 Research Panel, but who was a nonrespondent at the second wave. Among respondents to the first two waves of the Panel, 94.4% of those who were working in the first wave were also working in the second wave, and 90.1% of those who were not working in the first wave were also not working in the second wave. Thus, if the second-wave nonrespondent had been working in the first wave, the modal category imputation procedure would assign him a status of "working" in the second wave. If, however, he had not been working in the first wave, he would be assigned a status of "not working" in the second wave.

When, as in this example, a categorical variable is highly stable over time, the modal category imputation procedure reduces to assigning the value from the other wave. In this case, the use of this imputation procedure leads to an understatement of the change across waves. This understatement can be avoided by using a stochastic imputation procedure. In the above example, for instance, the second-wave nonrespondent who worked in the first wave could be assigned a second-wave status of "missing" not with certainty, but only with a probability of 0.94. He would have a probability of 0.06 of being assigned a second-wave status of "not working."

These imputation procedures for categorical variables can be readily extended to take account of additional auxiliary information by confining the procedures to specified subgroups of the sample. For instance, the missing second-wave work status for a man of a given age could be imputed from respondent data that related only to men in the same age group.

The preceding discussion has been cast in terms of two waves of data, one of which is missing. In a three-wave panel, the wave nonresponse patterns are 110, 101, 011, 100, 010 and 001. With pattern 110, the missing third-wave data could be forecast from the second wave by one of the procedures discussed; it would probably be satisfactory to ignore the first-wave data, since they are unlikely to add much explanatory power to that given by the second-wave data alone. In the same way, with 011, the first-wave data could be backcast from the second-wave data. The missing first and third waves of data in the pattern 010 could be backcast and forecast respectively. The second wave's data in 100 and 001 could similarly be forecast and backcast, but the other missing waves are two waves apart: these could equally be imputed by one of the preceding procedures, but probably less well. The final pattern, 101, has the missing wave surrounded by nonmissing waves. In this case, it should be possible to develop a stronger imputation method, using both adjacent waves' data in the imputation scheme.

The imputation schemes described above use the response for a variable on one wave in imputing for a missing response to that variable on another wave. These schemes are especially effective when the variable is highly stable, or at least the values are highly correlated between waves, for then the observed value on one wave is a powerful predictor of the missing value on the other. A limitation to these schemes is that the value of the same variable on another wave must be available. Kalton and Lepkowski found that in many cases these schemes could not be used because a person with a missing hourly rate of pay on one wave also had a missing rate on the other wave, or was a non-wage earner or not part of the panel on the other wave. An alternative back-up imputation procedure is needed to deal with such cases, adding to the complexity of the imputations and lowering their overall quality.

Another situation giving rise to the unavailability of responses to an item on another wave is when the item is included on the questionnaire for only one wave. The so-called "topical modules" on the SIPP questionnaires fall into this category. When crosswave imputation based on the same item on another wave cannot be applied, other forms of crosswave imputation, using other variables, may be employed. However, the quality of the resultant imputations will rarely compare with that of crosswave imputations based on the same item.

If imputation is used to handle wave nonresponse, the possibility of collecting data on additional auxiliary variables to improve the predictive power of the imputation models is worth considering. In particular, if a unit is a nonrespondent in one wave, additional data may be collected at the next wave. These data could include answers to topical items that are stable over time and answers to retrospective questions about nonstable issues.

5. Discussion

For simplicity of analysis, imputation is preferable to weighting as the method of handling wave nonresponse. It does not require choosing the appropriate set of weights to use for a particular form of analysis, and it avoids the inconsistencies that could occur when different weights are used for different analyses. With the weighting solution, it is for instance possible that the distribution of a variable on one wave will differ from its marginal distribution in a crosstabulation involving a variable from another wave.

An important factor in the choice between weighting and imputation is the amount of work required to implement the procedures. The work required to set up a wave nonresponse imputation procedure depends heavily on the number of variables in the survey. The task can be daunting with surveys like SIPP that collect data on very large numbers of variables. This factor thus favors weighting adjustments for such surveys. The development of efficient crosswave imputation procedures and associated edit checks is much more manageable in surveys that collect data on only a handful of variables, and imputation is consequently relatively more attractive in this case.

When imputation is based on a model with high predictive power, it is more efficient than weighting, even when the latter makes effective use of the auxiliary data. The development of good imputation models for all the many survey variables is, however, a substantial task. Moreover, the task is compounded by the need to have fall-back strategies for cases when the main auxiliary variables are unavailable. Yet imputation models will be required anyway for the item nonresponses within a wave. Models for item nonresponses also need to be developed carefully, and they should involve crosswave imputations for efficiency and to avoid distortion in measuring changes.

The potentially seriously harmful effects of imputation are the fabrication of data and the attenuation of the covariances between variables. The magnitude of these effects depends on the predictive power of the imputation models employed. When powerful models are used, as may often be the case when the imputation of a missing response is based on the response to the same item in another wave, these effects may not be appreciable. On the other hand, when weak models are used, as is likely to be the case for the topical items in the SIPP, these effects may be severe.

The severity of the effects of imputation depends not only on the predictive power of the imputation models but also on the form of analysis being conducted. The case for imputation rather than weighting is often stronger when the data are aggregated. Thus, for instance, a likely error of $1000 in an imputed four-month income of $8000 may be serious, but this error may be acceptable for an annual income of $24000, when only one of the incomes for the three four-month periods is imputed. Similarly, an error of $1000 may be serious for an individual's four-month income, but acceptable for the household annual income of $40000, when the incomes of other earners in the household and of that individual for the other four-month periods are known. With weighting adjustments, units with any missing components of an aggregate are excluded from the analysis.

Since both imputation and weighting have their disadvantages, it may be that a combination is the best solution. One combination would be to impute for variables for which powerful imputation models can be developed and to use weighting for other variables, such as those in the topical SIPP modules. While this approach has attractions, it creates the serious complication that for any wave or combination of waves two sets of weights would be required, one set to apply for those analyses that were restricted to variables for which missing waves were handled by imputation, and the second set to apply to analyses involving the other variables.

A second combination of weighting and imputation is to use weights to compensate for some patterns of wave nonresponse and to use imputation for others. In a three-wave panel, weighting could, for instance, be used to compensate for those that responded on only one wave and imputation could be used for the missing wave of those responding on two waves. On the one hand, this scheme avoids the deletion of units with two waves of data that occurs with the simple weighting approach that weights up the respondents to all three waves to represent the wave nonrespondents. On the other hand, it avoids the fabrication of two waves of data that occurs with the imputation approach. For the first three waves of the ISDP Research Panel, 11.7% of the persons responding on at least one wave had a single wave of missing data, which under this scheme would be handled by imputation. Another 8.2% had two waves of missing data which would be handled by weighting. This form of combination seems an attractive one.

A variant of this last procedure is to use imputation to complete the data in the non-nested patterns 011 and 101, and to discard the data in the non-nested patterns 001 and 010, thereby forcing the outcomes to nested patterns only. Then the nested weighting adjustments described earlier could be applied (Little and David (1983)).

6. References

Kalton, G. (1983): Compensating for Missing Survey Data. Survey Research Center, University of Michigan, Ann Arbor.

Kalton, G. and Kasprzyk, D. (1982): Imputing for Missing Survey Responses. Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 22–31.

Kalton, G. and Kish, L. (1984): Some Efficient Random Imputation Methods. Communications in Statistics, Theory and Methods, 13(16), pp. 1919–1939.

Kalton, G. and Lepkowski, J. M. (1983): Cross-wave Item Imputation. In Technical, Conceptual and Administrative Lessons of the Income Survey Development Program (ISDP), edited by M. H. David, pp. 171–198. Social Science Research Council, Washington, D.C.

Kalton, G. and Lepkowski, J. M. (1985): Following Rules in the Survey of Income and Program Participation. Journal of Economic and Social Measurement, 13, pp. 319–329.

Kasprzyk, D. and Kalton, G. (1983): Longitudinal Weighting in the Income Survey Development Program. In Technical, Conceptual and Administrative Lessons of the Income Survey Development Program (ISDP), edited by M. H. David, pp. 155–170. Social Science Research Council, Washington, D.C.

Little, R. J. A. (1984): Survey Nonresponse Adjustments. Proceedings of the Section of Survey Research Methods, American Statistical Association, pp. 1–10.

Little, R. J. A. and David, M. H. (1983): Weighting Adjustments for Non-response in Panel Surveys. Working paper, U.S. Bureau of the Census, Washington, D.C.

Oh, H. Lock and Scheuren, F. J. (1983): Weighting Adjustment for Unit Nonresponse. In Incomplete Data in Sample Surveys, Vol. 2, Theory and Bibliographies, edited by W. G. Madow, I. Olkin and D. B. Rubin, pp. 143–184. Academic Press, New York.

Rubin, D. B. (1979): Illustrating the Use of Multiple Imputations to Handle Nonresponse in Sample Surveys. Bulletin of the International Statistical Institute, 48(2), pp. 517–532.

Santos, R. (1981): Effects of Imputation on Regression Coefficients. Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 140–145.

Ycas, M. A. and Lininger, C. A. (1981): The Income Survey Development Program: Design Features and Initial Findings. Social Security Bulletin, 44(11), pp. 13–19.

Received January 1986
Revised July 1986