THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

THE TREATMENT OF PERSON-WAVE NONRESPONSE IN LONGITUDINAL SURVEYS

No. 26

G. Kalton, J. Lepkowski, S. Heeringa, Ting-Kwong Lin and M. E. Miller
Survey Research Center, University of Michigan

U. S. Department of Commerce BUREAU OF THE CENSUS

TABLE OF CONTENTS

INTRODUCTION

This report contains the findings of research conducted under a Joint Statistical Agreement between the Bureau of the Census and the Survey Research Center, University of Michigan, entitled "The Treatment of Person-Wave Nonresponse in Longitudinal Surveys". In longitudinal, or panel, surveys missing data can arise in three ways: unit nonresponse, when no data are collected for a sampled unit; item nonresponse, when a unit takes part in the survey but fails to provide acceptable answers to one or more of the items on the questionnaire; and wave nonresponse, when a unit provides data for some but not all waves of data collection. The choices of compensation procedure for missing data caused by unit and item nonresponse are generally straightforward; as a rule, weighting adjustments are used to compensate for unit nonresponse and imputation is used for item nonresponse. The choice of procedure to compensate for wave nonresponse is, however, less clear. It is this choice that is the subject of this report.

Viewed from a longitudinal perspective, wave nonresponse may be considered to be a set of item nonresponses in the longitudinal record, suggesting that imputation may be the appropriate compensation strategy. Viewed from a cross-sectional perspective, however, it may be considered to be unit nonresponse for which a weighting adjustment is appropriate. These two alternative strategies for handling wave nonresponse are examined in the following chapters.

The main focus of this research is on the choice of an appropriate compensation procedure for handling wave nonresponse at the person level in a longitudinal file created from the first three waves of a panel of the Survey of Income and Program Participation. At the outset of the research, data for the first three waves of the first SIPP Panel (the 1984 Panel) were not available. In consequence, the initial empirical investigations were performed using the 1979 Income Survey Development Program (ISDP) Research Panel, a large-scale panel survey that was conducted as part of the development of the SIPP. Subsequently, when cross-sectional data files for the first three waves of the 1984 SIPP Panel became available, these files were merged; and the empirical investigations were conducted with this merged file.

This report is a collection of five papers resulting from the research. The paper by Kalton, Lepkowski and Lin reproduced as Chapter 1 reports the results of investigations of wave nonresponse in the 1979 ISDP Research Panel, and provides an initial discussion of the alternative compensation strategies of weighting adjustments and imputation. Chapter 2, by Kalton, presents a general discussion of the issues involved in choosing between weighting adjustments and imputation for handling wave nonresponse. In Chapter 3, Kalton and Miller present the results of a simulation study conducted with the first three waves of the 1984 SIPP Panel to examine the effects of compensating for the wave nonrespondents by weighting adjustments or by a simple "carry-over" imputation procedure. In most panel surveys, many of the same items are repeated on each wave, and when the responses to such items are stable over time, a simple and fairly effective imputation procedure for wave nonresponse is just to impute the responses from the preceding wave. This is the "carry-over" or "direct substitution" imputation procedure. This procedure may also be employed for item nonresponses, and may produce much better imputations than are obtained from a standard item imputation procedure applied within a wave. In Chapter 4, Heeringa and Lepkowski compare carry-over and cross-sectional hot-deck imputations for item nonresponses to some wage and salary items in the first three waves of the 1984 SIPP Panel. The final paper, Chapter 5 by Lepkowski, provides a general review of issues involved in compensating for wave nonresponse in panel surveys.

As discussed in the report, a number of considerations are involved in making the choice between weighting adjustments and imputation for handling wave nonresponse. If weighting adjustments are used, it is attractive on the grounds of simplicity to employ a single set of weights to compensate for all patterns of wave nonresponse. However, this leads to discarding much of the data provided by the wave nonrespondents. On the other hand, the imputation solution has the attraction of retaining all the data collected, but may lead to distortions in the relationships between variables. The development of an effective imputation procedure for wave nonresponse is a major undertaking, whereas by contrast, the development of a weighting adjustment procedure is straightforward.

Given the pattern of wave nonresponse experienced in the first three waves of the 1984 SIPP Panel, the general conclusion here is that the weighting adjustment solution is preferable for the three wave file. In this file the loss of data associated with the weighting solution is not great, and it seems preferable to accept that loss rather than employ imputation with the consequent risks of distortions to covariances. However, this conclusion applies only to the three wave file. With files containing more waves of data, the loss of data associated with the simple single weighting adjustment solution for wave nonresponse will be greater, and it may therefore be preferable to employ imputation for at least some of the patterns of wave nonresponse.

CHAPTER 1

COMPENSATING FOR WAVE NONRESPONSE IN THE 1979 ISDP RESEARCH PANEL[1]

Graham Kalton, James Lepkowski, Ting-Kwong Lin

1. Introduction

The choice between weighting adjustments and imputation for handling missing survey data is generally straightforward: as a rule, weighting adjustments are used for total nonresponse and imputation is used for item nonresponses. There are, however, several situations where the choice is debatable. In general, these are situations of what might be termed partial nonresponse, where some data are collected for a sampled unit but a substantial amount of the data is missing. These situations include cases where the respondent terminates the interview prematurely, where data are not obtained for one or more members of an otherwise cooperating household (for household level analysis), and where an individual provides data for some but not all waves of a panel survey.

If weighting is used for partial nonresponse, the available responses for that unit may be employed in the determination of the weights, but the unit itself is discarded, resulting in a loss of data. On the other hand, if imputation is used, a sizeable number of responses for a partially nonresponding unit will need to be imputed, giving rise to concerns about the fabrication of much of the data and the effect of this fabrication on the relationships between variables. This paper examines the choice between weighting and imputation for handling the partial nonresponse that occurs when a respondent fails to provide data on one or more waves of a panel survey. Kalton (1985) provides further discussion of the issues involved in choosing between weighting and imputation to handle wave nonresponse, and Cox and Cohen (1985) report the results of an experimental investigation of these alternatives in the National Medical Care Expenditure Survey.

[1]From *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 1985, 372-377.

The objective of this study is to provide evidence on the choice between weighting and imputation for handling wave nonresponse in the Survey of Income and Program Participation (SIPP). The SIPP is a panel survey in which households are interviewed every four months over a period of about two-and-a-half years (Herriot and Kasprzyk, 1984). One major product of the SIPP will be an annual file combining three waves of data, and the focus of the present study is on this annual file. Since a longitudinal file for the first three waves of the first SIPP panel is not yet available, the empirical investigation reported here is based on the first three waves of the 1979 Income Survey Development Program (ISDP) Research Panel, a large-scale panel survey that was conducted as part of the development of the SIPP. All the results reported here relate only to original sample persons aged 16 and over in the area frame part of the 1979 Research Panel sample; persons sampled from the special list frames and persons joining the panel after the first wave are excluded from all the analyses.

In a three-wave panel there are eight different patterns of response/nonresponse for the sampled units. Denoting 1 as response and 0 as nonresponse, one of these patterns is 000, representing the nonrespondents to all three waves. The form of adjustment for these total nonrespondents is unproblematic, namely a weighting adjustment, and hence they will not be considered further here. The distribution for the other seven patterns for the 1979 Research Panel is given in Table 1.

The first pattern in Table 1 represents those who responded on all three waves of the panel, whereas the other six patterns represent those who failed to respond on one or two of the waves. The issue under study is whether weighting or imputation should be used to handle each of these six patterns. The next section of the paper discusses how weighting adjustments might be applied, and the following one discusses the use of imputation. The final section presents some concluding remarks.

2. Weighting Adjustments for Wave Nonresponse

The use of weighting adjustments for partial nonresponse presents two additional complications beyond those that apply with weighting adjustments for total

Table 1

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

Pattern	Response (1) Nonresponse (0)	%
1	111	80.2
2	110	7.2
3	101	2.3
4	011	2.2
5	100	6.7
6	010	0.6
7	001	0.9
	Total	100.0
Number of persons		20,676

nonresponse. One results from the fact that there is a great deal more information available about partial nonrespondents than about total nonrespondents. Often only a limited amount of auxiliary information is available for total nonrespondents (such as the PSUs and strata in which they are located), whereas for partial nonrespondents there is also the information provided by their responses to the questions they have answered. The complication raised by these extra data is how they should be taken into account in determining the weighting adjustments for partial nonrespondents.

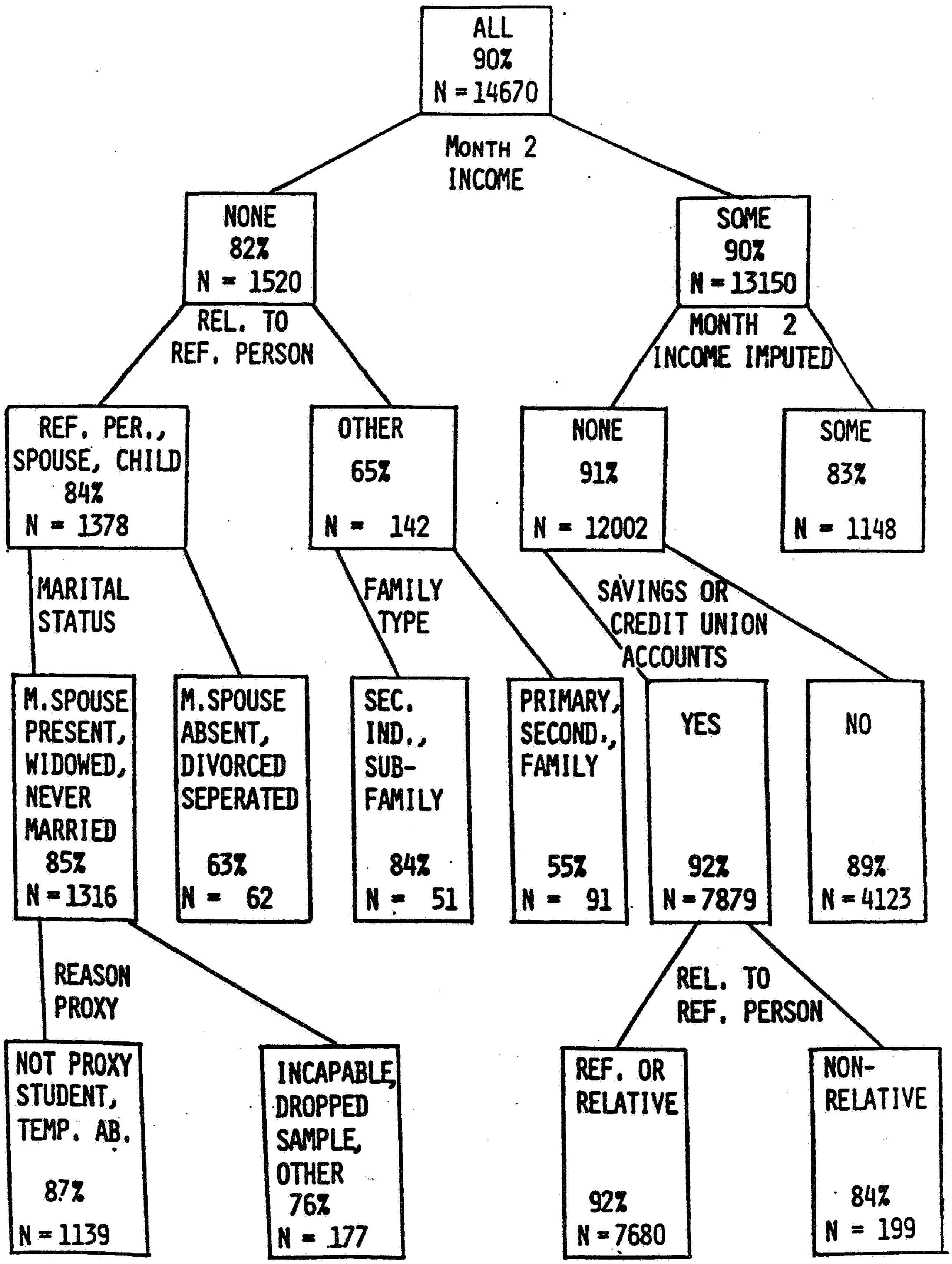
The second complication arises from the fact that surveys are subject to many different forms of analyses. Some partial nonrespondents will have provided all the data needed for certain analyses, and hence can be included in them, but they will not have provided all the data needed for some other analyses. If all those providing the requisite data for a particular analysis are included in that analysis, different analyses will be based on different subsets of the sample. This raises the complication that different sets of weights are needed according to what subset of the sample is included in a particular analysis. These two complications are discussed in turn subsequently in relation to handling wave nonresponse by weighting adjustments.

As an illustration of the first complication, consider the simple case of compensating for the second wave nonrespondents in the 1979 Research Panel. The auxiliary variables available for these partial nonrespondents are the design variables (PSUs and strata, etc.) and their wave 1 responses. The aim is to discover which, if any, of these variables are associated with response status at wave 2, and then to develop weights to compensate for differential wave 2 response rates in different parts of the sample. With the large number of wave 1 response variables, the first step in the analysis is to reduce those to be investigated in detail to a manageable number. This was done by examining the bivariate associations of each of the auxiliary variables in turn with the wave 2 response status variable. All but a few of the auxiliary variables were found to have virtually no association with wave 2 response status, and these variables were therefore excluded from the further analyses.

The next step was to employ the remaining auxiliary variables as joint predictors of wave 2 response status using SEARCH analyses (Sonquist, Baker, and Morgan, 1973) and logistic regressions. Figure 1 presents the results of a SEARCH analysis, one which explains 2.3 per cent of the variation in the wave 2 response status variable. Examination of this tree diagram shows that 88 per cent of the sample falls in cells with response rates between 87 and 92 per cent, and that 98 per cent falls in cells with response rates between 83 and 92 per cent. Only three small cells have distinctly lower response rates. In terms of weighting adjustments, giving the cell with the 92 per cent response rate a weight of 1, the weights for 88 per cent of the sample would be between 1 and 1.06 and for 98 per cent would be between 1 and 1.11. The use of these weights, with their slight variation, would be unlikely to have any appreciable effects on analyses of the data.

As an alternative to the SEARCH analysis, logistic regression analyses with wave 2 response status as the dependent variable were also conducted. For one of these regressions, the independent variables from wave 1 were the reason for proxy interview (1), the recipiency of interest income (2), the amount of personal earnings in

SEARCH ANALYSIS FOR WAVE 2 RESPONSE STATUS



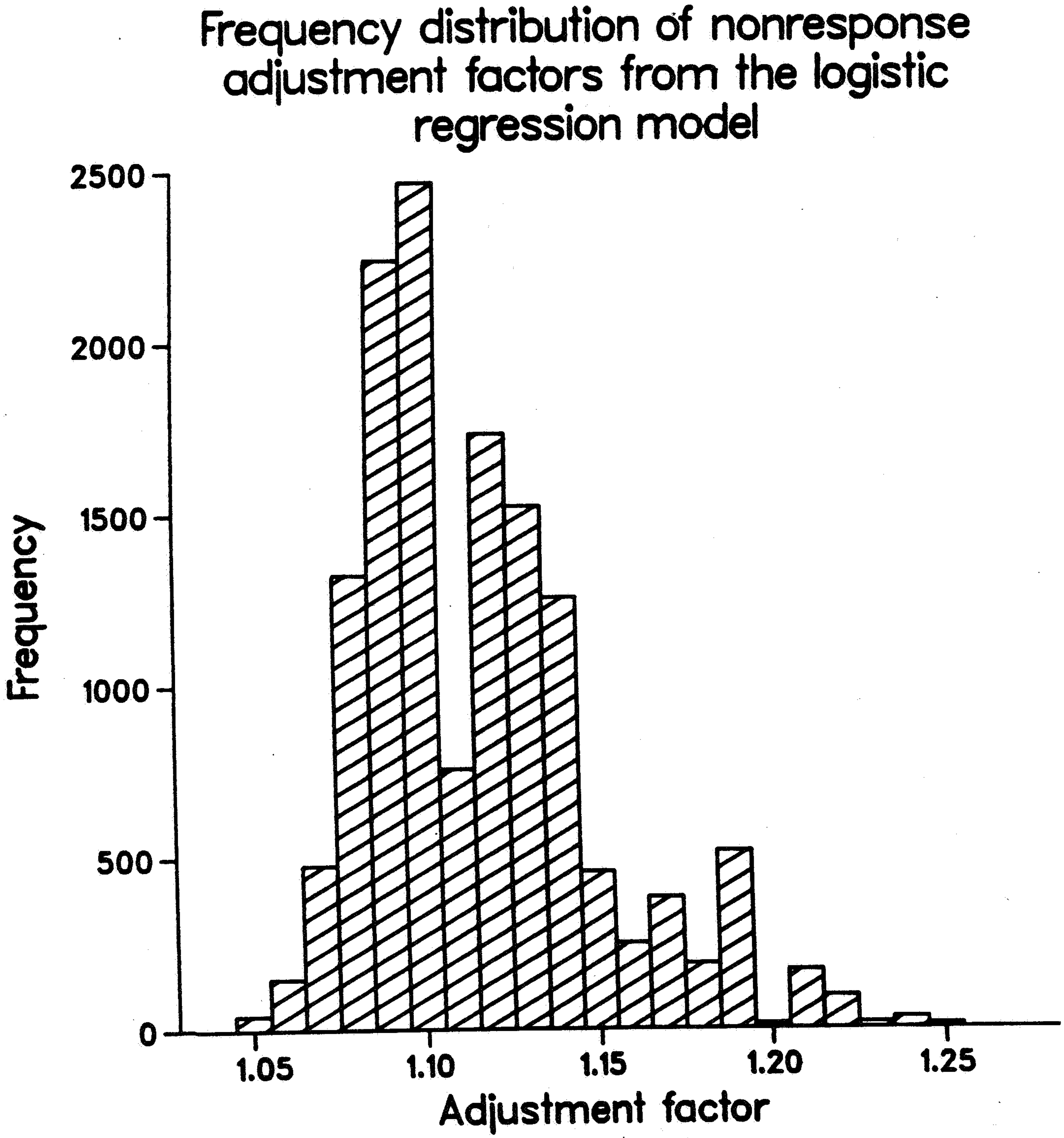
2.3% VARIATION EXPLAINED

month 2 (3), the relationship to the reference person (4), the type of family (5), marital status (6), and the two-factor interactions (1,2), (1,3), (1,4), (1,6), (4,5) and (5,6). Following Little and David (1983), the weights for wave 2 respondents were then set to be the inverses of their individual predicted means from this regression. Figure 2 shows the resulting distribution of weights. This distribution has a similar spread to that obtained from the SEARCH analysis, but in this case there are a few outliers. In practice, these outliers would probably be trimmed back to avoid the increase in sampling error associated with relatively large weights.

The results of the above analyses are fairly reassuring about the nature of wave 2 nonresponse. Comparisons of wave 2 respondents and nonrespondents show that the two groups are generally very similar in terms of their wave 1 responses. The differences that have been identified are not major ones, and weighting adjustments can be employed to compensate for them. Since the variation in these weights is not great, their use will not result in much loss of precision in the survey estimates. The weights from the SEARCH analysis, for example, would be likely to lead to an increase of less than 1/2 per cent in the variance of the survey estimates.

The second complication noted above concerns the need to employ different sets of weights for different types of analyses in the presence of partial nonresponse. For instance, considering the patterns of wave nonresponse in Table 1, it can be seen that patterns 1, 2, 3 and 5 provide data for cross-sectional analyses of wave 1, patterns 1, 2, 4 and 6 provide data for cross-sectional analyses of wave 2, patterns 1 and 2 provide data for analyses of changes between waves 1 and 2, and only pattern 1 provides data for forming aggregates across all three waves (*e.g.*, income over the period). For any particular analysis, the respondents in the patterns that provide the requisite data need to be weighted up to represent the other patterns. There are potentially seven combinations of waves that could be used for different forms of analysis, thus implying the need for seven different sets of weights. With more waves in the panel, the potential number of sets of weights increases rapidly. For instance,

FIGURE 2.


Frequency distribution of nonresponse adjustment factors from the logistic regression model
2500
2000
1500
1000
500
0
Frequency
1.05
1.10
1.15
1.20
1.25
Adjustment factor

with the eight waves from a full SIPP panel, there are 255 possible combinations of waves, and hence as many as 255 different sets of weights could be required.

The number of sets of weights needed would be reduced if not all the patterns of response/nonresponse occurred. In many panel surveys the major type of nonresponse is attrition nonresponse, which refers to the situation in which a unit drops out on one wave and remains out of the panel for all subsequent waves. If the only form of nonresponse was attrition nonresponse, there would be just four response/nonresponse patterns for a three wave panel, namely 111, 110, 100 and 000, and only three sets of weights would be needed. There would be one set of weights for each wave: these weights would apply straightforwardly for cross-sectional analyses of data from single waves, and an analysis incorporating data from two or more waves would use the weights applicable to the latest wave involved in that analysis.

Little and David (1983) propose a method for developing weights to compensate for attrition nonresponse that attempts to take account of all the auxiliary data available on the nonrespondents. The only information known about nonrespondents at the first wave (*i.e.,* the total nonrespondents) is their values on the design variables (*e.g.,* PSUs and strata), z; the information available for those who drop out at the second wave comprises their z-values and their responses at the first wave, x_1; the information available for those who drop out at the third wave comprises their z- and x_1-values and their responses on the second wave, x_2; and so on. Little and David propose running the following series of logistic or probit regressions with the response indicators r_i(r_i = 1 for a respondent, r_i = 0 for a nonrespondent at wave i) as the dependent variables:

(1) Regress r_1 on z_1 for the total sample
(2) Regress r_2 on z_1 and x_1 for respondents at wave 1
(3) Regress r_3 on z_1, x_1 and x_2 for respondents at wave 2; and so on.

The inverses of the predicted means from these regressions then give the weights needed to compensate from one wave to the next. Let these weights be denoted by w_1, $w_{2.1}$, and $w_{3.12}$. The overall weights for first wave respondents are then w_1; for second wave respondents they are $w_2 = w_1 w_{2.1}$; for third wave respondents they are $w_3 = w_2 w_{3.12}$; and so on.

Little and David (1983) also describe a weighting scheme for nonattrition nonresponse, but the simplicity of the above procedure is lost, and their scheme also has some unattractive features. As can be seen from Table 1, there were in fact a fair number of nonattrition nonrespondents in the 1979 Research Panel: the patterns 101, 011 and 001 account for 6.0 per cent of the total sample and comprise almost one-third of the partial nonrespondents. An approach that can be used to avoid the complications of the nonattrition nonresponse patterns is to convert them into attrition patterns. This can be done either by discarding some waves of data, by imputing some waves of data, or by a combination of these procedures. Thus, for instance, one might impute for the missing wave in the 011 pattern, discard the data in the 001 pattern, and either impute for the middle wave or discard the last wave in the 101 pattern. Note that if discarding is the chosen solution, the data need not have been collected in the first place (except for its potential use for methodological checks).

3. Imputing for Wave Nonresponse

When wave nonresponse is handled by imputation, all the missing items for a wave nonrespondent are assigned values, making use of responses on other waves in doing so. As Kalton and Kasprzyk (1982) discuss, the value imputed for the ith nonrespondent on variable y may in general be expressed as $y_i = f(x_{1i}, x_{2i},...,x_{pi}) + e_i$, where $f(\mathrm{x})$ is a function of the p auxiliary variables used in the imputation, and e_i is an estimated residual. If the e_i are set equal to zero, the imputation scheme assigns the predicted means, and the scheme may be termed a deterministic one. On the other hand, if the e_i are estimated residuals, the scheme may be termed a stochastic one. Deterministic imputations distort the shape of the distribution of y, and attenuate its variance. For this reason, stochastic imputation schemes are generally preferred.

In the SIPP and the 1979 ISDP Research Panel, in common with most panel surveys, many of the same items are repeated on each wave. Often the responses to a repeated item are highly consistent over time, and when this occurs the response on one wave can serve as a powerful auxiliary variable to use for imputing the missing response on another wave. To illustrate this point, we consider first some categorical variables and then some continuous variables from the 1979 Research Panel.

For the categorical variables we examine the consistency of responses across the first two waves of the 1979 Research Panel. The upper part of Table 2 presents unweighted cross-wave distributions of responses to whether the person worked in the quarter and to two recipiency items for original sample persons aged 16 and over who responded on both waves. The lower part of the table gives corresponding distributions of reasons for not working for those who were not at work on both waves. As the first row of the table shows, 58.2 per cent of persons reported that they worked on both waves and 34.5 per cent reported that they did not work on either wave. Thus, a total of 92.8 per cent of the respondents were consistent in their responses across the first two waves of the panel.

Table 2

Distribution of sample persons across Waves 1 and 2 for selected variables for original sample respondents for both waves ages 16 and older from the area frame, 1979 ISDP Research Panel

Item / 1st wave / 2nd wave	Yes / Yes	Yes / No	No / Yes	No / No	Consistency	Sample size
Worked in quarter	58.2	3.5	3.8	34.5	92.8	13,119
Receiving Soc. Sec.	18.4	0.4	0.9	80.3	98.7	13,151
Receiving Fed. SSI	3.2	0.3	0.3	96.2	99.5	13,151
Reasons for not working:						
Going to school	11.0	0.9	0.7	87.4	98.4	4,520
Didn't want to work	4.9	6.5	8.5	80.1	84.9	4,520
Retired	15.3	5.0	6.5	73.2	88.5	4,520

The degree of consistency of response for all the items in Table 2 is high, with the lowest level of consistency being 84.9 per cent for the responses to the item "Didn't want to work" as a reason for not working. That the "Didn't want to work" item exhibits the lowest level of consistency is perhaps not unexpected, given its greater degree of subjectivity than the other items. It is likely that all these consistency measures are underestimates, because of measurement errors, possible mismatches of respondents across waves, and other reasons. Even items like race and marital status show some degree of inconsistency. The former item has a consistency measure of 99.6 per cent, and the latter item has one of 97.8 per cent; several of the inconsistencies in marital status were in fact logical impossibilities, such as married, widowed or divorced at wave 1 and never married at wave 2.

The high levels of consistency found in Table 2 suggest that the response to one of these items on one wave is a good predictor for a missing response on the other wave. In order to illustrate how the quality of imputations based on responses to the same item on another wave may be assessed, consider the item in the first row of the table, whether the respondent worked in the quarter or not.

Among the respondents to both waves, 94.4 per cent of those who answered "Yes" to this item at wave 1 (*i.e.,* said they worked in the quarter) also said "Yes" at wave 2, and 90.1 per cent of those who answered "No" at wave 1 also answered "No" at wave 2. There were 1518 persons who answered this question on wave 1, but failed to answer it on wave 2; of these, 922 answered "Yes" at wave 1 and 596 answered "No". Using a deterministic imputation scheme, all those answering "Yes" at wave 1 would be assigned "Yes" answers at wave 2 (this being the modal wave 2 response amongst those answering "Yes" at wave 1); similarly, all those answering "No" at wave 1 would be assigned "No" answers at wave 2. Assuming that nonrespondents at wave 2 are missing at random conditional on their wave 1 responses, one can expect that 94.4 per cent of the 922 responding "Yes" at wave 1 will be correctly assigned "Yes" at wave 2 (*i.e.,* an expected 870 persons) and 90.1 percent of the 596 answering

"No" at wave 1 will be correctly assigned "No" answers at wave 2 (*i.e.*, an expected 537 persons). Thus this imputation scheme may be expected to correctly assign the responses of 92.7 per cent of the wave 2 nonrespondents. Without using the wave 1 responses in the imputation scheme, all the 1518 wave 2 nonrespondents would be assigned "Yes" responses with a deterministic imputation scheme, since "Yes" is the modal answer among wave 2 respondents. Again assuming wave 2 nonrespondents are missing at random conditional on their wave 1 responses, an expected 61.2 per cent of them would be correctly assigned "Yes" responses for wave 2.

The above deterministic scheme based on wave 1 responses suffers the disadvantage that it imputes only 60.7 per cent of "Yes" wave 2 responses, whereas 61.2 per cent of "Yes" responses should be imputed to generate the correct distribution of "Yes" and "No" answers under the missing data model adopted. (The difference here is small, but it could be greater in other cases.) In addition, the deterministic imputation scheme leads to a greater stability of responses over the two waves than is implied by the model: there are no changes in responses from wave 1 to wave 2 for those with imputed wave 2 responses.

A stochastic imputation scheme can avoid these disadvantages. A stochastic scheme for the above example would assign "Yes" responses to 94.4 per cent of wave 2 nonrespondents who answered "Yes" at wave 1 and "No" responses to the other 5.6 per cent, and it would assign "No" answers to 90.1 per cent of wave 2 nonrespondents who answered "No" at wave 1 and "Yes" answers to the other 9.9 per cent. A disadvantage of the stochastic scheme, however, is that it reduces the quality of the imputations: based on the missing at random conditional on wave 1 response model, the expected percentage of correct imputations with this scheme is only 86.6 per cent.

It should be emphasized that all the measures of the quality of the imputations are based on a model for the nonrespondents. The measures may be misleading if the model fails to hold. The model used here assumes that the wave 2 nonrespondents have the same distribution of wave 2 responses as the wave 2 respondents, conditional on their wave 1 responses. Thus, for instance, it is estimated that 94.4 per cent of the wave 2 nonrespondents who answered "Yes" at wave 1 would answer "Yes" at wave 2. This estimate may be seriously in error if the model is inappropriate, and if so, the measures of imputation quality will be invalid.

Consider now the imputation of continuous variables across waves of a panel survey. Kalton and Lepkowski (1983) describe a variety of procedures that can be employed for crosswave imputation in a two-wave panel, using the value of a variable on one wave to impute the missing value of the same variable on another wave. The widely used hot-deck imputation procedure does not work well when the auxiliary variable and the variable to be imputed are very highly correlated, as will often be the case with crosswave imputation. With the hot-deck procedure, the auxiliary variable is categorized into cells, and an individual with a missing value on the variable under consideration is assigned the value of a respondent from the same cell. Thus an individual from one end of a cell may be assigned the value from a respondent at the other end of that cell. Closer matches between nonrespondents and donors can be obtained by increasing the number of hot-deck cells, but the number of cells has to be limited to ensure that matches can be made.

The categorization with the hot-deck procedure can be avoided by using some form of regression imputation. Consider, for example, the imputation of the hourly rate of pay of individual i on wave 2 (y_i) given the individual's hourly rate of pay on wave 1 (x_i). A simple regression imputation model is $y_i = a + bx_i + e_i$, where e_i is a residual term. The e_i's do not need to have a zero mean, and no restriction need be placed on their distribution. Regression imputation can be viewed as constructing a new variable $\hat{y}_i = a + bx_i$ for all individuals, imputing the e_i's for the nonrespondents, and then calculating y_i as $\hat{y}_i + e_i$. The e_i's may be assigned by any appropriate imputation scheme. They may, for instance, be imputed by a hot-deck procedure, selecting respondents' e_i's within imputation cells formed by, say, age, sex, and categorized wave 1 hourly rate of pay to assign to the nonrespondents. The choice of regression imputation model is not critical, since the assignment of the e_i's can protect against a misspecified model. The better the choice of model, however, the smaller is the variance of the e_i's, and hence the better is the quality of the imputed y_i's.

Obvious choices for a and b are the least squares estimates obtained from a regression of respondents on both waves, but simpler alternatives may also work well. The simplest model is to take $a = 0$, $b = 1$, which specifies the wave 2 value as the wave 1 value plus the change between waves: the imputation is then made for changes. Other relatively simple models set either $a = 0$ or $b = 0$; the first is a proportionate change model and the second an additive change model. There is in fact no need to include the a term in the model, since it can be incorporated as part of the residual (*i.e.*, the residual is taken to be $a + e_i$).

The quality of crosswave imputations depends on (1) the correlation between the values of the item from one wave to the next and (2) the quality of the imputations for the residuals obtained by using other auxiliary variables. We present some findings from the 1979 Research Panel relating to the first of these factors.

First consider the hourly rate of pay variable. For original sample respondents aged 16 and older in the area frame reporting hourly rate of pay on each of the first two waves of the Panel, the correlation between the two waves is 0.976. Similarly, from waves 2 to 3 the correlation is 0.964 and from waves 1 to 3 it is 0.965. (All these correlations are computed after 28 cases of apparent keying errors had been removed.) These high correlations suggest that if a person's hourly rate of pay is available for one wave but not for a neighboring wave, the missing rate can be imputed with little error (even before considering the use of auxiliary variables in the imputation of the residual term).

Unlike hourly rate of pay, most of the amounts items in the 1979 Research Panel were reported on a monthly basis, so that there are three amounts reported for each wave. The cross-month correlations for one amount item, wage and salary income, for the first three waves of the 1979 Research Panel are given in Table 3. The data are again limited to original sample persons aged 16 and older from the area sample, and only persons reporting that they received wage and salary income are included in the correlation estimates. The correlations were computed using a pairwise missing data deletion algorithm so that the numbers of records used for different correlations may vary. Several records in the data file had apparent keying errors for the wage and salary amount (*e.g.,* the amount increased from one month to the next exactly by a factor of 10 or 100, suggesting a decimal place shift in the keying process). Since these potential errors substantially reduced cross-month correlations, the data values in error were excluded from the pairwise correlations.

Table 3

Cross-month correlations for wage and salary income amount for original sample persons ages 16 and older from the area frame, 1979 ISDP Research Panel

	1	2	3	4	5	6	7	8
2	0.903							
3	0.878	0.894						
4	0.840	0.858	0.834					
5	0.839	0.854	0.833	0.955				
6	0.828	0.853	0.816	0.945	0.944			
7	0.800	0.804	0.802	0.832	0.843	0.849		
8	0.809	0.797	0.784	0.826	0.843	0.822	0.952	
9	0.795	0.809	0.787	0.825	0.828	0.835	0.949	0.949

The correlations across months are generally high, ranging from 0.784 to 0.955. The highest correlations are between months within waves, while the lowest tend to occur for months that are more than 6 months apart. Looking down the main diagonal of the lower triangular matrix in Table 3, it can be seen that correlations between adjacent months in different waves are lower than those between adjacent months in the same wave. There are several possible explanations. One is that respondents tend to give falsely consistent responses within a wave, leading to unduly high within wave correlations. It seems more likely, however, that it is the between wave correlations that are too low. This could arise because of response variation between waves, including cases of proxy reports on one wave and self-reports on another. Also, a close examination of the records suggests that there may be some mismatched records in the file, giving rise to large differences in wage and salary income between waves.

Correlations for other amounts items in the 1979 Research Panel demonstrate similar high cross-month correlations. The correlations for wage and salary income and six other amounts items are summarized in Table 4. Average correlations were computed for the same difference between months, and separately for reports within the same wave and between different waves. For example, the average within wave correlation for a one month difference for the wage and salary amount is the average of months 1 and 2, months 2 and 3, months 4 and 5, months 5 and 6, months 7 and 8, and months 8 and 9 correlations from Table 3. The corresponding average between wave correlation is the average of the months 3 and 4 and months 6 and 7 correlations.

As observed for wage and salary income amounts, the average correlations between months in different waves for the other items are always smaller than those between months in the same wave. The correlations also decrease as the number of months between reports increases. But generally the correlations for these income items are high, indicating the kind of stability that may be used to provide accurate imputed values for missing data by using cross-month and cross-wave imputation strategies.

One of the items in the table has appreciably lower correlations than the rest, namely unemployment compensation amounts. The correlations for this item start by falling as the number of months between reports increases, but then rise for longer intervals: the correlations for months six or more months apart are in fact higher than the correlation for one month apart. This pattern of correlations may indicate that short-term unemployment receives unstable compensation while longer-term employment receives relatively stable amounts of compensation. In any case, the lower

Table 4

*Average cross-month correlations for seven amount items for original sample persons ages 16 and older from the area sample, 1979 ISDP Research Panel**

	One month difference			Two month difference								
	Within wave	Between wave	Within and Between	Within wave	Between wave	Within and Between	Three month	Four month	Five month	Six month	Seven month	Eight month
Wage and salary amount	0.933	0.842	0.910	0.890	0.839	0.861	0.837	0.830	0.810	0.794	0.809	0.795
Personal earnings	0.910	0.760	0.872	0.900	0.753	0.816	0.741	0.724	0.699	0.672	0.675	0.661
Social Security	0.983	0.921	0.968	0.978	0.924	0.946	0.919	0.913	0.902	0.890	0.892	0.900
Federal SSI	0.931	0.886	0.919	0.912	0.856	0.880	0.829	0.812	0.810	0.762	0.717	0.596
AFDC	0.961	0.897	0.945	0.931	0.887	0.906	0.859	0.831	0.799	0.572	0.693	0.715
Unemployment compensation	0.651	0.408	0.590	0.645	0.448	0.532	0.428	0.527	0.436	0.760	0.745	0.695
Food stamps	0.966	0.900	0.949	0.937	0.892	0.911	0.883	0.867	0.849	0.820	0.814	0.790

*Excluding apparent keying errors as missing data.

correlations for this item indicates the need for greater efforts to employ effective auxiliary variables in imputing for the residuals for unemployment compensation.

The preceding discussion has been 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 (1983) found that in many cases these schemes could not be used in imputing for hourly rate of pay in the 1979 Research Panel 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 responses to the item being unavailable on another wave is when the item was 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 on one wave, additional data may be collected at the next wave. Such a strategy is being adopted in the SIPP, with the addition of a "Missing Wave" section to the questionnaire for the fourth and subsequent waves of data collection (Bailey, Chapman and Kasprzyk, 1985). This section collects information on labor force participation, income sources and asset ownership/nonownership of respondents who, although eligible, did not respond to the preceding wave.

4. Concluding Remarks

The choice between weighting adjustments and imputation for handling wave nonresponse is not a simple one. Each method has its advantages and disadvantages. Imputation creates a completed data set that is easy for the analyst to use and, when based on a model with high predictive power, imputation is more efficient than weighting. The development of good imputations for all the variables in a missing wave is, however, a major undertaking. Unless the overall imputation scheme is constructed with great care, taking account of the cross-sectional and longitudinal interrelationships between all the variables, inconsistent or otherwise unacceptable imputed values may be assigned. In any event, imputation fabricates data to some extent and it will cause an attenuation in some of the covariances between variables. The amount of fabrication and attenuation is slight when powerful crosswave imputation models are used, but such models cannot be used in all cases. On the other hand, while weighting avoids the attenuation problem, the need to use different sets of weights for different types of analyses creates complexities for the analyst and can lead to inconsistent results. With both imputation and weighting having their advantages and disadvantages, it may be that some combination of the two methods, such as that outlined at the end of Section 2, is the best solution.

References

Bailey, L., Chapman, D.W. and Kasprzyk, D. (1985). Nonresponse adjustment procedures at the Census Bureau: a review. *Proceedings of the Bureau of the Census First Annual Research Conference*, 421–444.

Cox, B.G. and Cohen, S.B. (1985). *Methodological Issues for Health Care Surveys*. New York: Marcel Dekker.

Herriot, R.A. and Kasprzyk, D. (1984). The Survey of Income and Program Participation. *Proceedings of the Social Statistics Section, American Statistical Association*, 107–116.

Kalton, G. (1985). Handling wave nonresponse in longitudinal surveys. *Proceedings of the Bureau of the Census First Annual Research Conference*, 453–461.

Kalton, G. and Kasprzyk, D. (1982). Imputing for missing survey responses. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 22–31.

Kalton, G. and Lepkowski, J. (1985). Following rules in the Survey of Income and Program Participation. *Journal of Economic and Social Measurement*, 1, (forthcoming).

Kalton, G. and Lepkowski, J. (1983). Cross-wave item imputation. In *Technical, Conceptual, and Administrative Lessons of the Income Survey Development Program (ISDP)*, M. David, ed., pp. 171–198. New York: Social Science Research Council.

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)*, M. David, ed., pp. 155–170. New York: Social Science Research Council.

Little, R. and David, M. (1983). Weighting adjustments for non-response in panel surveys. Bureau of the Census working paper.

Sonquist, J.A., Baker, E.L., and Morgan, J.N. (1973). *Searching for Structure*. Ann Arbor, Michigan: Institute for Social Research.

Chapter 2

HANDLING WAVE NONRESPONSE IN PANEL SURVEYS[2]

Graham Kalton

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 used for item nonresponses, 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 types 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

---

[2]From *Journal of Official Statistics*, 1986, **2**, 303–314.

the survey is a panel survey that collects data from the same units on 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 imputation. The complication with 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 person nonresponse may be viewed as a set of item nonresponses in the household record, suggesting that imputation may be used in compensation. For person-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 attempting to compensate 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 the item in question, and then a nonrespondent is assigned the response for that item from 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 need to take a sample of respondents to whom to assign increased weights, 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 serves to explain 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 sizeable. 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 for, as noted above, the higher their predictive power the lesser 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 turn to 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 (1982); 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 apply 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 1 as response and 0 as nonresponse, these eight patterns are:

111	110	101	011
100	010	001	000

The last pattern represents the total nonrespondents; for any form of analyses 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 analyses, 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 analyses, 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 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% non-attritors.

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. For wave 1 cross-sectional analyses, the $(a+b)$ first wave respondents are

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	20,676

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 analyses 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 a matching of 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 for instance 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 account for 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 *i*th 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 βs 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 the 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 rates of pay of respondents 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 for imputing 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, 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 by 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 for 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 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 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 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 no responses to the item being available on another wave is when the item was 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 the 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 the choice of 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 cross-tabulation 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 cross-wave imputation procedures and associated edit checks is much more manageable for 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 $1,000 in an imputed four-month income of $8,000 may be serious, but this error may be acceptable for an annual income of $24,000, when only one of the incomes for the three four-month periods is imputed. Similarly, an error of $1,000 may be serious for an individual's four-month income, but acceptable for the household annual income of $40,000, 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.

With both imputation and weighting having 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 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 would apply for those analyses that were restricted to variables for which missing waves were handled by imputation, and the second set would 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 weighting approach and, 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 on 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, Volume 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.

Chapter 3

EFFECTS OF ADJUSTMENTS FOR WAVE NONRESPONSE ON PANEL SURVEY ESTIMATES[3]

Graham Kalton and Michael Miller

1. Introduction

Nonresponse in a panel survey can be classified into three components: total nonresponse, when a sampled unit does not take part in any wave of the survey; wave nonresponse when a unit takes parts in some but not all waves of data collection; and item nonresponse, when a unit takes part in a particular wave but fails to provide acceptable responses for some of the items. Total nonresponse and item nonresponse are routinely handled by weighting adjustments and imputation respectively. The choice of adjustment procedure for wave nonresponse is, however, less straightforward (Kalton, 1985). If weighting is used, data provided by the wave nonrespondents on waves for which they did respond are discarded, causing a loss of data. On the other hand, if imputation is used, complete waves of data have to be imputed, causing concerns about the fabrication of large amounts of data and the effect of the imputations on the relationships between variables. This paper examines the effects of these alternative strategies for handling wave nonresponse on survey estimates by means of a simulation study.

The simulation study is based on the 1984 Panel of the Survey of Income and Program Participation (SIPP). A description of the SIPP is provided by Nelson, McMillen and Kasprzyk (1985). The data set for this study was created by merging the public use files for the first three waves of the 1984 SIPP Panel. To create the simulation data set, the respondents on all three waves were taken from the merged file, and some waves of their data were deleted in a way that reflected the missing waves of data in the complete file. Details of the construction of the simulation data set are given in

---

[3] An abbreviated version of this chapter appears in the *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 1986, forthcoming.

Section 2. Imputation and weighting adjustments were then each applied to compensate for the missing waves of data.

The imputation of missing wave responses was carried out by a simple cross-wave imputation procedure: a wave nonrespondent's responses on a missing wave were assigned the values of that nonrespondent's responses to the same items on the most recent earlier wave for which data were available. The use of the responses to the same items on another wave as auxiliary information in an imputation procedure is effective when the responses to the items are stable over time, as is often the case. The stability of some items across the first three waves of SIPP is examined in Section 3. Section 4 then examines the quality of the imputations produced by the simple "carry-over" imputation procedure.

The weighting adjustments were applied to the three-wave respondents to compensate for those who missed either the second or the third wave, or both. (In the 1984 SIPP Panel no attempts were made to interview first wave nonrespondents on subsequent waves; hence all first wave nonrespondents are total nonrespondents, and as such are excluded from the present investigation.) The auxiliary variables used for determining the weighting classes were responses to certain items at the first wave.

Survey estimates have been computed from (i) the weighted sample of respondents to all three waves, (ii) from the data set with carry-over imputations assigned for missing wave responses, and (iii) from the data set with the actual responses (i.e., with the deleted values in the simulation data set replaced). Section 5 compares the estimates obtained from these three procedures. The final section of the paper presents some conclusions from this study.

2. The Simulation Data Set

A sample of households is selected for the first wave of a SIPP panel, and all persons aged 15 and over in the selected households become panel members who are followed even if they change addresses or move out of their sampled households. Children under 15 in sampled households become panel members at later waves after reaching the age of 15 provided that they are still living with a panel member at that time. Persons who were not in the initial sample but who subsequently reside with panel members – termed associated persons – are included in the survey while they continue to live with panel members. Panel members and associated persons are interviewed every four months about their income and program participation in the preceding four months.

For the purposes of this study a number of exclusions have been made from the total data set for the first three waves of the 1984 SIPP Panel. First, rotation group 4 was excluded because data were not collected from this group in the second wave. Second, all associated persons have been excluded. Third, all children aged under 15 at the first wave have been excluded. Fourth, all panel members leaving the survey population (e.g., through death, entering an institution, or emigration) have been excluded. Fifth, all nonrespondents at the first wave have been excluded; this category includes both nonresponding households (type A nonrespondents) and individual nonrespondents in cooperating households (type Z nonrespondents). The study is thus confined to panel members aged 15 and over at the first wave who were respondents at that wave and who remained in the survey population throughout the first three waves. There were 30,004 such persons in the data set. The patterns of response/nonresponse for these 30,004 persons are shown in Table 1.

The first step towards the creation of the simulation data set was to seek predictors for the four response patterns exhibited in Table 1. This step was conducted using SEARCH analyses, employing the option that maximizes the variation explained in terms of a χ^2 statistic (Sonquist, Baker, and Morgan, 1973). The predictor variables included in these analyses were any first wave variables that had some degree of association with the response patterns. (Unfortunately it was not possible to include a variable relating to the degree of urbanization of the panel member's area of residence in these analyses; this potentially important predictor variable had to be excluded because there was no suitable indicator relating to degree of urbanization

Table 1

*Person Response/Nonresponse Patterns Across the First Three Waves of the 1984 SIPP Panel for Respondents at the First Wave who Remained Eligible for the Panel for Three Waves**

Response (X)/Nonresponse (0)	%
XXX	90.0
XX0	4.9
X0X	1.0
X00	4.2
Total	100.0
Number of persons	30,004

*Rotation groups 1, 2 and 3 only.

available for all first wave respondents from the public use data files.) The objective for the SEARCH analyses was to develop a detailed and complex model for the response patterns. Since the purpose of the model was for constructing the simulation data set, not for substantive analysis, a complex but unstable model was preferred to a simpler, more stable, one.

The results of the SEARCH analysis adopted for the creation of the simulation data set are given in Table 2. As can be seen from the table, the analysis divided the sample into 41 groups. The largest group, group 14, contains 28% of the sample; four groups contain over 2000 panel members, and in combination they cover 55% of the sample. The percentage of respondents on all three waves (XXX) varies from 61.6% (group 1) to 98.6% (group 36), the percentage of the XX0 pattern varies from 0% (groups 13, 35 and 36) to 18.6% (group 39), the percentage of the X0X pattern varies from 0% in several groups to 12.3% (group 35), and the percentage of the X00 pattern varies from 0% in several groups to 22.2% (group 1).

The simulation data set was formed from respondents to the first three waves in the following manner. First, within each of the 41 SEARCH groups, a random sample of the XXX respondents was taken. The sample size in each group was set at 61.6% of

TABLE 2
SEARCH ANALYSIS OF RESPONSE PATTERNS

Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Level 8	Group Number	RESPONSE PATTERN 111	110	101	100	Group Size
Imputation of any other financial investments (72.30)	Total Household Earned Income For Month 1 ≥ $3000							1	61.62	16.16	0	22.22	99
	Total Household Earned Income For Month 1 < $3000 (75.54%)	Marital Status Codes 3, 4, 6						2	76.19	1.90	4.76	17.14	105
		Marital Status Codes 1, 2, 5 (75.23%)	Total Household Earned Income For Month 4 < $900					3	72.27	15.13	8.40	4.20	119
			Total Household Earned Income For Month 4 ≥ $900					4	78.64	8.74	0	12.62	103
No Imp. of Any Other Financial Investments (90.22)	Marital Status Codes 1, 3 (91.46)	Living Quarters owned or Being Bought (92.11)	Imputation of any Certificates of Deposits or other Savings Certificates					5	84.25	2.74	0	13.01	146
			No Imp. of Any CD's or other Savings (92.19)	>5 persons in household during Int. month (89.47)	6 persons in HH during Int. month (89.00)	Owned a Savings Account		6	91.69	8.00	0	0.31	325
						Did Not own a Savings Account		7	84.46	8.29	0	7.25	193
					>6 persons in HH during Int. month (80.14)	Person Belongs to a Related Sub-family		8	75.00	11.54	9.62	3.85	52
						Person Belongs to Primary Family or Unrelated Sub-family		9	92.65	3.51	1.92	1.92	313
				≤ 5 persons in household during Int. month (92.36)	Age 16-39 (93.84)	1-3 persons in HH during Int month (92.53)	Afro-American Ethnic Background	10	81.54	12.31	0	6.15	65
							All other Ethnic Backgrounds	11	92.95	2.82	0.88	3.35	1703
						4 or 5 persons in HH during Int. Month		12	94.75	3.08	0.43	1.74	2533
					Age 40-85 (91.69)	Age 40-75 (91.31)	Mexican, Mex-Am, Chicano Ethnic Background	13	98.31	0	0	1.69	118
							All other Ethnic Backgrounds	14	91.21	3.88	.84	4.07	825
						Age 75-85 (94.66)	Attended ≤ 3 years College	15	94.92	1.69	.42	2.97	944
							Attended > 3 years College	16	92.31	5.77	1.92	0	104
		Living Quarters Rented or Occupied without cash Payment (89.38)	Owned a Savings Account (91.61)	Attended ≤ 2 years College				17	91.05	5.57	.29	3.10	1743
				Attended > 2 years College				18	93.56	1.81	.60	4.02	497
			Did not own a savings Account (87.28)	Attended ≤ 8 yrs. school (92.79)	Age 40-85			19	95.44	1.88	.27	2.41	373
					Age 15-39 (85.71)	≤ 4 persons in HH during Int. month		20	78.75	7.50	1.25	12.50	80
						> 4 persons in HH during Int. month		21	95.00	5.00	0	0	60
				Attended > 8 years school (85.78)	Ethnic Group "not listed"			22	87.58	7.49	.56	4.36	894
					All other Ethnic Groups (84.12)	Total HH Earned Income month 1 < $2000 (84.77)	Ethnic Groups: Russians, Af-Am, 'Don't Know'	23	87.44	2.51	3.02	7.04	199
							All other Ethnic Groups	24	83.77	9.25	1.13	5.85	530
						Total HH Earned Income month 1 ≥ $2000 (82.19)	2-3 persons in HH during Int. month	25	78.91	4.76	6.12	10.20	147
							> 3 persons in HH during Int. month	26	87.00	1.00	0	12.00	100
	Marital Status Codes 2, 4, 5, 6 (87.88)	Relationship to Reference person Codes 5, 6, 7 (82.23)	Total HH earned income month 3 < $1600 (84.03)	Age 15-59				27	82.03	10.98	.33	6.66	601
				Age 60-85				28	94.12	.84	0	5.04	119
			Total HH earned income month 3 ≥ $1600 (80.22)	Sex: Male				29	75.95	12.90	2.35	8.80	341
				Sex: Female				30	85.05	6.98	4.32	3.65	301
		Relationship to Reference Person Codes 1, 2, 4 (88.75)	Living Quarters owned or Being Bought (89.98)	≤ 12 years school attended				31	90.97	4.64	.62	3.77	3235
				> 12 years school attended (88.42)	Age 15-44 years			32	87.51	4.86	1.60	6.02	1810
					Age 45-85 years			33	94.88	1.97	0	3.15	254
			Living Quarters Rented or Occupied w/o cash Payment (86.91)	> 5 persons in HH during Int. month (92.25)	7 or 8 persons in HH during Int Month			34	93.80	5.43	.78	0	129
					6 persons in HH during Int month (90.70)	Afro-American Ethnic Group		35	80.70	0	12.28	7.02	57
						All other Ethnic Groups		36	98.61	0	1.39	0	72
				≤ 5 persons in HH during Int. month (86.50)	Total HH Earned Income Month 4 ≥ $4000			37	70.30	15.84	4.95	8.91	101
					Total HH Earned Inc. Month 4 < $4000 (87.01)	Ethnic Groups: PR, Cuban Spanish, other Spanish (75.17)	Total HH Earned Inc. Month 3 < $600	38	80.23	16.28	3.49	0	86
							Total HH Earned Inc. Month 3 ≥ $600	39	67.80	18.64	0	13.56	59
						All other Ethnic Groups (87.57)	Age 15-44	40	86.55	7.22	1.22	5.01	2453
							Age 45-85	41	91.85	5.43	.85	1.87	589
Totals									90.00	4.90	1.00	4.20	30009

the total number of panel members in that group. The 61.6% figure was chosen because it is the lowest percentage of XXX respondents across the 41 groups. Thus in group 1, where the 61.6% figure applies, all the XXX respondents were kept in the sample; in group 2, the sample size of the XXX respondents was set at 61.6% of 105, i.e., 65; etc. The purpose of this procedure was to generate a sample of XXX respondents that has the same distribution across the 41 groups as the total sample. The sample of XXX respondents thus created comprises 18,481 persons.

The last stage in producing the simulation data set was to assign a response pattern to each of the 18,481 members of the sample of XXX respondents. The response patterns were assigned at random within the SEARCH groups, according to the percentage distributions for the group response pattern distributions given in Table 2. A variable was added to each data record to indicate the record's assigned response status. When the variable indicated that a sample member was a nonrespondent on one or more waves, the data for those waves are then treated as missing in the data set. In the analysis, weighting adjustments or imputation methods are used in an attempt to compensate for these missing data. Estimates made from the data set using the alternative methods of compensation for the missing data are then compared with the estimates based on the complete data set.

Although the simulation data set was constructed from respondents to all three waves, it needs to be recognized that not all the data are actual responses. Some respondents failed to answer some of the items, and in these cases the values in the data set are the values imputed by the Bureau's cross-sectional imputation procedures. Since these imputed values may distort the survey estimates – particularly estimates of change across waves – some of the results presented below relate only to records with no imputed values on the variables employed in the particular analysis.

There are weights on the original SIPP records that include an allowance for the nonrespondents at the first wave, that is the total nonrespondents. These weights are not employed for any of the analyses in this paper. The only weights used here are the weights developed to handle wave nonrespondents, as described in Section 4.

3. Stability of Responses

The effectiveness of a cross-wave imputation scheme that uses the response to an item for an available wave in imputing for a value on a missing wave depends on how well the missing value can be predicted from the available response. This section examines for several SIPP items how well responses for one point of time can be predicted from responses to the same item for another point of time. The data set employed for these analyses is the simulation data without the deletion of any of the responses. Two types of data items need to be distinguished: those that are measured once each wave and those that are measured monthly, that is four measurements are collected each wave.

The upper part of Table 3 provides the distributions of responses for a selection of items measured once each wave with simple Yes (Y) or No (N) responses. The lower part of the table gives: the average percentage responding "Yes" on a wave, P(Y); the percentage responding "Yes" on the second wave given a "Yes" on the first wave, $P(Y_2|Y_1)$; the percentage responding "Yes" on the third wave given a "Yes" on the first wave, $P(Y_3|Y_1)$; and corresponding percentages for the "No" responses. The conditional percentages provide an indication of the predictive ability of first wave responses for later wave responses. Thus, for instance, without prior information, there is a 37.6% chance that a person does not have a job, whereas if it is known that the person did not have a job on the first wave, there is a 91.6% chance that the person did not have a job on the second wave, and a 86.5% chance that the person did not have a job in the third wave. As can be seen from the table, there is a fair degree of stability in the responses to all the items across waves, but the degree of stability varies between items. The quantities $P(X_1 \neq X_2)$ and $P(X_1 \neq X_3)$ denote the overall percentages of responses on the first wave that are different from those on the second and third waves, respectively.

Table 3

Distributions of Responses across Waves for (a) Having a Job (b) Looking for Work (c) Receiving Social Security Payments (d) Receiving Food Stamps (e) Having Savings Accounts (f) Having Certificates of Deposit, all in the Past Four Months

Y=Yes N=No	Having a Job %	Looking for Work* %	Receiving Social Security %	Receiving Food Stamps %	Having Savings Account %	Having Certif. of Deposit %
YYY	55.8	1.7	17.6	3.3	49.7	13.0
YYN	2.3	1.1	0.1	0.5	2.6	0.9
YNY	2.2	0.2	0.1	0.2	1.2	0.2
YNN	2.8	1.9	0.2	0.8	3.3	1.1
NYY	2.5	0.9	0.6	0.5	2.7	1.1
NYN	0.7	1.4	0.1	0.2	0.4	0.1
NNY	2.5	1.3	0.6	0.5	2.2	0.9
NNN	31.2	91.5	80.7	94.1	38.0	82.8
	100.0	100.0	100.0	100.0	100.0	100.0
$P(Y)$	62.4	4.8	18.5	4.5	56.0	15.1
$P(Y_2 \mid Y_1)$	91.9	56.8	98.5	80.2	92.2	91.2
$P(Y_3 \mid Y_1)$	91.9	39.1	98.1	73.3	89.6	86.8
$P(N)$	37.6	95.2	81.5	95.5	44.0	84.9
$P(N_2 \mid N_1)$	91.6	97.6	99.1	99.3	92.9	98.6
$P(N_3 \mid N_1)$	86.5	97.7	98.5	99.0	88.7	97.7
$P(X_1 \neq X_2)$	8.2	4.5	1.0	1.6	7.5	2.5
$P(X_1 \neq X_3)$	10.1	5.2	1.6	2.2	10.8	3.9
No. of persons	18481	11271	18481	18481	18481	18481

*Only for those in the labor force at all waves.

For all items, a very high proportion of those who did not have the item on the first wave (i.e., the "No's") also did not have the item on the second or third waves. For most items, the corresponding result also holds for those who had the item on the first wave: a high proportion of these individuals also had the item on the second and third waves. The two items showing least stability in having the item across waves are looking for work and receiving food stamps. Only 56.8% of those looking for work in the first wave were still looking in the second wave, and only 80.2% of those who received food stamps in the first wave also received for stamps in the second wave.

The results in Table 3 suggest that, in imputing for a missing response, the use of a previous wave's value in the imputation procedure will often lead to a good imputed value. Moreover, the results in Table 3 underestimate the true stability of the items over time for a variety of reasons. First, some of the responses analyzed in Table 3 are themselves imputed values because of item nonresponses: these imputations were carried out by the Bureau of the Census on a cross-sectional basis, and hence are likely to introduce instability across waves. Second, a variety of other aspects of the survey operation are likely to give rise to variability in measurement errors, and hence an overstatement of instability across waves. These include simple response variability, changing informants across waves (e.g. self-report on one wave, proxy report on another), matching errors, and keying errors. Kalton, McMillen and Kasprzyk (1986) demonstrate the existence of instability induced by measurement errors with examples of inconsistencies in race, sex and age (more than a one year change) between adjacent waves in the 1984 SIPP Panel.

Other evidence on the existence of variability in measurement error between waves of the SIPP Panel comes from the items that are measured on a monthly basis. Burkhead and Coder (1985) have noted that for a number of items on recipiency status for various sources of income more changes in status between adjacent months occur when the data are collected in different waves than when they are collected in the same wave. Moore and Kasprzyk (1984) report the same finding with the Income Survey Development Program 1979 Panel. In the ISDP 1979 Panel, Kalton, Lepkowski and Lin (1985) found that monthly income amounts from various sources were more highly correlated from one monthly to the next when the month amounts were obtained in the same wave than when they were obtained in adjacent waves.

Table 4 provides an example of the stability in response for an amounts item in the SIPP. The table presents the correlation matrix for the monthly amounts of Social Security income received. Each correlation is based on the sample of persons reporting Social Security income in both of the two months. Imputed values are treated as missing data and influential outliers of $1500 or more have been removed (ten records had amounts of $1500 or more in one or more months). The correlations are based on samples of about 3000 respondents.

Table 4

*Cross-month correlations for Social Security income amounts from the simulation data set**

	1	2	3	4	5	6	7	8	9	10	11
2	0.99										
3	1.00	0.99									
4	0.99	0.98	0.99								
5	0.92	0.92	0.92	0.92							
6	0.92	0.91	0.92	0.92	0.99						
7	0.92	0.91	0.92	0.92	0.99	0.99					
8	0.93	0.92	0.93	0.92	0.99	0.99	1.00				
9	0.92	0.91	0.92	0.92	0.94	0.93	0.94	0.94			
10	0.92	0.91	0.92	0.92	0.94	0.94	0.94	0.94	1.00		
11	0.92	0.91	0.92	0.92	0.94	0.93	0.94	0.94	0.99	1.00	
12	0.92	0.91	0.92	0.92	0.93	0.93	0.94	0.94	0.99	1.00	1.00

*Excluding imputed amounts and monthly amounts of $1500 or more.

The results in Table 4 exhibit the same pattern of correlations that Kalton, Lepkowski and Lin (1985) found with the ISDP 1979 Panel: for a given difference in the months, the correlations when both amounts are obtained in one wave are appreciably higher than when they are obtained from different waves. The matrix in fact seems to partition into two parts: the correlations in amounts for months within a wave are on average about 0.99 whereas those between amounts for months in different waves are on average about 0.92. There is no evidence of a decline in the correlations as the difference between the months increases. A 3.5% increase in the level of Social Security payments was introduced in January, 1984. This increase thus applied in month 8 for rotation group 1, month 7 for rotation group 2, and month 6 for rotation group 3. As a consequence, the data for months 6 and 7 are mixtures of amounts from before and after the increase. This should lead to lower correlations between these months and other months, but this effect is not discernible.

The within-wave correlations are extremely high, and suggest that cross-wave imputations can be extremely effective. The between-wave correlations are lower, but are still high. It is not clear what the true correlation is. On the one hand, the between-wave correlations are probably attenuated by variation in measurement errors while, on the other hand, the within-wave correlations may be too high because respondents tend to overstate the consistency within a wave.

Table 5 provides another way of showing the consistency of Social Security income across months. This table gives the distribution of the percentage change in the amount received from one month to the next. Outliers and imputed values are excluded. The results are given by calendar months in order to remove the effect that the panel months relate to different calendar months for different rotation groups. The table brings out clearly the marked differences between the situation when the amount for the previous month is obtained in the same wave or from the previous wave. Excluding January, when the 3.5% increase came into effect, on average 98.5% of amounts showed no change from the previous month when data for both months were obtained in the same wave. In contrast, on average only 34.8% of amounts showed no change when the data for the previous month were collected in the previous wave.

The results in Table 5 exclude imputed values and outliers. When these are included, the instability of monthly amounts between adjacent months in different waves increases appreciably. On average only 31.0% of amounts showed no change in this case, and 19.1% of amounts changed by more than 10%, as compared with 12.1% when the imputed values were excluded.

The changes in Social Security amounts between December and January should reflect the 3.5% increase that occurred at that time. Table 5 shows that the majority of the respondents did indeed report increases in this period, but still 35.7% reported the same amount as in December. The percentage reporting no change from December varies appreciably with rotation group: for the rotation group for which December was the first and January the second month in the second wave, 43.5% of respondents

Table 5

*Percentage Change in Amount of Social Security Income in Current Month Compared to Previous Month**

Month	Within (W) or Between (B) wave	Percent change from previous month: Reduction: More than 10%	Reduction: 10% or less	No Change	Increase: 10% or less	Increase: More than 10%	Total	Sample Size
September	W	0.2	0.2	99.2	0.3	0.1	100.0	2970
October	W	0.2	0.3	99.1	0.4	0.1	100.0	1990
	B	5.9	22.5	36.2	26.8	8.6	100.0	948
November	W	0.3	1.0	97.7	0.5	0.6	100.0	1978
	B	5.4	23.2	28.3	36.2	7.0	100.0	962
December	W	0.3	0.5	97.7	1.0	0.5	100.0	2030
	B	5.7	25.9	23.3	37.8	7.2	100.0	938
January	W	0.3	1.0	35.7	61.3	1.8	100.0	3021
February	W	0.2	0.1	97.0	2.3	0.3	100.0	2008
	B	5.5	22.0	35.9	30.8	5.8	100.0	961
March	W	0.0	1.1	98.1	0.6	0.1	100.0	2002
	B	4.6	21.5	42.5	26.1	5.3	100.0	970
April	W	0.1	0.2	99.1	0.2	0.4	100.0	2040
	B	5.4	18.1	42.6	27.1	6.8	100.0	951
May	W	0.4	0.1	99.1	0.2	0.3	100.0	3042

*Excluding imputed amounts and monthly amounts of $1500 or more.

reported no change; for the rotation for which December was the second and January the third month, 34.8% reported no change; and for the rotation group for which December was the third and January the fourth month, 29.0% reported no change.

These results are consistent with the theory that respondents tend to forget changes and especially those that occurred longer ago.

In summary, the results in this section show that some SIPP items have a good deal of stability over waves. The exact extent of stability is however hard to assess because of measurement error problems. These measurement errors confound the assessment of cross-wave imputation procedures. In our simulation study we evaluate the carry-over imputation procedure by measuring how well it reproduces the values that we deleted. Variability in measurement error will cause this evaluation to understate the effectiveness of cross-wave imputation. Indeed, if much of the change between waves is attributable to variability in measurement error, it may be the case that for some purposes the carry-over imputations are in fact superior to the actual responses.

4. Quality of Carry-Over Imputations

A standard procedure for evaluating the quality of an imputation scheme in a simulation study is to examine how well the scheme reproduces the actual, but deleted, values. As noted in the previous paragraph, this procedure is problematic in the present case because of the probable variation in measurement errors between waves, but it is nevertheless applied in this section.

We use two indices to measure the quality of the imputations, the mean (MD) and either the mean square deviation (MSD) or its square root, the root mean square deviation (RMSD). The mean deviation is given by

$$MD = \frac{1}{n}\Sigma(\hat{y}_i - y_i),$$

where $\hat{y}_i$ is the imputed value, y_i is the actual value for the ith missing response and n is the number of imputed responses. The mean deviation is the difference in the means of the imputed and actual values, and is a measure of the bias of the imputation procedure. The mean square deviation is given by

$$MSD = \frac{1}{n}\Sigma(\hat{y}_i - y_i)^2$$

It measures the closeness of the imputed to the actual values.

For items with simple Yes/No responses, "Yes" answers can be scored 1 and "No's" scored 0. Then the MD is the difference in the proportions of "Yes" answers between the imputed and actual responses, and the MSD is the proportion of incorrect imputations. Table 6 gives the MD's and MSD's for the items considered in Table 3 for imputed values at the second wave (responses patterns XOX and XOO) and at the third wave (response patterns XXO and XOO).

The mean deviations in Table 6 represent the differences between the percentages of "Yes" answers in the imputed values and in the actual, but deleted, values for those assigned for the simulation to represent wave nonrespondents. Thus, for instance, the figure of 1.7% in the top left-hand corner of the table relates to the 173 respondents who had their second wave responses deleted in the simulation data set. With the carry-over imputation procedure, they were then assigned their first wave responses for the missing second wave responses. Based on these imputed values, 73.4% of them were classified as having a job in the second wave. Based on their actual second wave responses, the corresponding percentage is 71.7%. The difference between these percentages is the mean deviation of 1.7% in the table.

With the carry-over imputation procedure, a mean deviation of 0 occurs with a given response pattern when the percentage of the nonrespondents endorsing the item is the same at the missing wave as at the wave from which the carry-over imputed values are taken. A review of the distributions of the items in Table 3 shows that for the total sample the percentages endorsing the items under consideration here are mostly stable from one wave to the next. It is therefore not surprising that most of the mean deviations for the wave nonrespondents in Table 6 are close to 0. Only four of the mean deviations are significantly different from zero, and they can be readily explained. Consider, for instance, the 'having a job' item. From Table 3 it can be calculated that 63.1% of respondents had a job in the first-wave, 61.3% had a job in the second wave, and 63.0% had a job in the third wave. If the same percentages of second wave nonrespondents had jobs in the first wave as the total sample, 63.1% of

Table 6

Mean Deviations and Mean Square Deviations for Several Items for Second and Third Wave Imputations by Response Pattern

Item	Second Wave Imputations		Third Wave Imputations	
	XOX %	XOO %	XXO %	XOO %
		Mean Deviations		
Having a Job	1.7	3.0*	-2.6**	1.6
Looking for Work+	0.0	-1.0	2.1**	-1.0
Receiving Social Security	-0.6	0.1	-0.4	0.0
Receiving Food Stamps	0.6	0.1	0.4	0.0
Having Savings Accounts	1.7	3.3**	0.0	1.4
Having Certificates of Deposit	0.6	0.7	-0.3	0.0
		Mean Square Deviations		
Having a Job	7.5	10.0	8.4	10.7
Looking for Work+	3.3	6.0	4.8	6.0
Receiving Social Security	0.6	0.7	1.1	1.8
Receiving Food Stamps	0.6	1.7	1.3	1.3
Having Savings Accounts	7.5	10.0	5.3	12.9
Having Certificates of Deposit	1.7	4.0	2.8	4.7
Number of imputations	173	767	906	767
(Number of imputations for looking for work item)	(123)	(484)	(578)	(484)

+Only for those in the labor force at all waves
*Significant at the 5% level using McNemar's test
**Significant at the 1% level using McNemar's test

them would have jobs imputed to them in the second wave; if they also had the same percentage of jobs at the second wave as the total sample, 61.3% would in fact have jobs. Thus the mean deviation, or bias, of the imputation values would be 1.8%. The XXO wave nonrespondents have their third wave missing responses imputed from their second wave responses. Assuming that they behave as the total sample, the percentage with jobs increase from 61.3% to 63.0% between the second and third waves, but the imputations will show only 61.3% of them with jobs at the third wave. The bias is then – 1.7%. Similar explanations apply with the other items where the imputed values have significant biases. The carry-over imputation procedure risks serious bias when the level of endorsement of an item varies appreciably over waves. Some other form of cross-wave imputation may be needed in this case.

The mean square deviations in Table 6 represent the percentages of incorrect imputations (e.g., imputing having a job when the respondent has no job or vice versa). As might be expected, for the second wave imputations these percentages are broadly similar to the percentages of responses that change between the first two waves in Table 3, that is, $P(X_1 \neq X_2)$. In the same way, the MSD's for the third wave imputations for the XOO pattern are similar to the $P(X_1 \neq X_3)$ percentages in Table 3. The percentage of correct imputations is generally high, but there is nevertheless a not insignificant number of errors made.

We now turn to consider the quality of the carry-over imputation procedure for a numerical variable, Social Security income, that is obtained monthly. In this case, the first carry-over imputation we use assigns the amount for the latest available month for each missing month. The analysis reported here is restricted to those who receive Social Security income in the latest available month and in the months for which the responses are deleted. The analysis does not therefore reflect the effect of changes in recipiency status for Social Security income. Records with Bureau of the Census cross-sectional imputations for item nonresponses on Social Security income are deleted because they would distort the analysis. Monthly amounts of $1500 or more and changes of more than $200 between months are also deleted (six records had changes of more than $200 between months).

Table 7 presents the mean deviations (as percentages of the actual monthly means) and root mean square deviations for Social Security amounts that qualify after the above exclusions are made. A notable feature of the mean deviations is the significant negative biases in the imputed amounts from month 7 onwards for the XOX and XOO patterns. These biases may be explained by the fact that with these patterns the imputed values are carried over from months prior to January, 1984, and therefore

Table 7

Mean Deviations and Root Mean Square Deviations for Social Security Imputed Monthly Incomes in the Second and Third Waves by Response Patterns

Month	XOX		XOO		XXO	
	MD+ %	RMSD \$	MD+ %	RMSD \$	MD+ %	RMSD \$
5	0.1	10.8	-0.1	23.7	-	-
6	-0.6	13.2	-1.0	24.3	-	-
7	-2.1**	17.9	-1.4*	24.7	-	-
8	-3.2**	18.8	-2.2**	29.6	-	-
9	-	-	-3.8**	32.9	0.5	16.2
10	-	-	-3.8**	32.8	0.9	25.1
11	-	-	-4.0**	33.5	0.5	16.2
12	-	-	-4.3**	38.0	0.6	15.8
Approximate No. of Imputations	20	20	97	97	110	110

+As a percentage of the mean of the actual responses.
*Significant at the 5% level using a matched sample 't' test
**Significant at the 1% level using a matched sample 't' test

do not take account of the 3.5% increase that occurred in that month. With the XXO, the imputed values are taken from months after January and hence include the increase.

The root mean square deviation bears some similarity to a residual standard deviation around the predicted values. The standard deviations of the Social Security monthly amounts in this restricted data set are around \$180. The small magnitudes of the RMSD's compared with this standard deviation indicate the effectiveness of the carry-over imputation procedure for Social Security amounts (once the outliers have been removed).

An obvious modification to make to the carry-over imputation procedure for Social Security amounts is to increase all amounts carried over from months before January to January or later by 3.5%. This modification affects only the XOX and XOO response patterns. Table 8 gives the mean deviations and root mean square deviations for this modified carry-over imputation procedure for these two patterns for the same set of records as Table 7. As can be seen from the table, there are now no significant biases and the RMSD's are slightly lower than the corresponding ones in Table 7. The modification thus produces a useful improvement in the imputed values.

Table 8

Mean Deviations and Root-Mean Square Deviations for Social Security Imputed Monthly Incomes, Adjusted for January Increase, in the Second and Third Waves by Response Pattern

	XOX		XOO	
Month	MD+	RMSD	MD+	RMSD
5	0.1	10.8	−0.1	23.7
6	0.3	9.4	0.3	23.4
7	0.1	9.2	0.8	23.4
8	0.2	10.7	0.6	20.2
9	–	–	−0.5	29.9
10	–	–	−0.5	29.9
11	–	–	−0.6	30.5
12	–	–	−1.0	35.3
Approximate No. of imputations	20	20	97	97

+As a percentage of the mean of the actual responses

5. Comparison of Imputed and Weighted Estimates

One way to handle wave nonresponse is by some form of imputation, such as the carry-over imputation procedure discussed in the previous section. An alternative way is by a weighting adjustment. This section compares a selection of survey estimates computed under these alternative adjustment procedures with the estimates computed from the actual values.

It is possible to develop a number of different sets of weights to compensate for wave nonresponse, with the choice of the weights to be used in a particular analysis depending on the waves from which data are needed for that analysis (Kalton, 1985). The use of different sets of weights enables use to be made of all the responses on the waves for which data are available, but it adds to the complexity of the data set. For this investigation, we have developed a single set of weights to compensate for all wave nonrespondents; this is the approach being adopted by the

Bureau in creating an annual file for the SIPP. The use of a single set of weights has the attraction of simplicity, but it is wasteful of the data collected on wave nonrespondents.

The weighting scheme used for this study assigned weights to the 16,635 respondents to all three waves (pattern XXX) to compensate for the 1846 wave nonrespondents (patterns XXO, XOX and XOO). Data collected at the first wave were used to form weighting classes within which the three-wave respondents were weighted up to represent the wave nonrespondents. The weighting classes were formed by a classification according to sex, four age groups, three household income levels, race, three educational levels, whether receiving certain types of welfare or not, whether in the labor force or not, and whether unemployed or not. The classification was collapsed until all weighting classes contained a minimum of 20 three-wave respondents. The weights for the resultant classes vary between 1.0 and 1.5.

Table 9 presents the distributions in the total simulation data set for the patterns of having and not having a job in the three waves (a) for the actual data before the simulated wave nonrespondents' values were deleted, (b) for the data with wave nonrespondents' missing values imputed by the carry-over imputation procedure, and (c) for the data with the three-wave respondents weighted up to represent the wave nonrespondents. Comparisons of these three distributions show that they are very similar: the actual and weighted distributions are virtually identical, with the imputed distribution exhibiting some small differences. The noteable feature is that the imputed distribution overstates the percentages in the consistent patterns YYY and NNN compared with the actual distribution.

The close similarity of the distributions in Table 9 is not surprising given the relatively small amount of wave nonresponse. The imputation and the weighting adjustments have little effect on total sample estimates. A more insightful analysis is to examine how well these two forms of nonresponse adjustments represent the wave nonrespondents. In the case of imputation, this analysis can be readily conducted by

Table 9

Distributions of Responses across Waves for Having a Job in the Wave for the Total Sample in the Simulation Data Set (a) with the Actual Responses, (b) with Imputed Responses for Wave Nonrespondents, and (c) with Weighting Adjustments for Wave Nonrespondents.

Y=Yes N=No	(a) Actual %	(b) Imputed %	(c) Weighted %
YYY	55.8	56.3	55.7
YYN	2.3	2.1	2.3
YNY	2.2	2.0	2.2
YNN	2.8	2.8	2.8
NYY	2.5	2.3	2.5
NYN	0.7	0.6	0.7
NNY	2.5	2.3	2.5
NNN	31.2	31.6	31.3
Total	100.0	100.0	100.0
No. of persons	18,481	18,481	16,635

comparing the estimates obtained for the wave nonrespondents (i) from the actual values and (ii) from the combination of actual and imputed values, where imputed values are assigned when missing waves occur. In the case of weighting adjustments, the wave nonrespondents are represented by increases in the weights to the three-wave respondents. Weighted estimates for the wave nonrespondents can therefore be obtained from weighted analyses of the three-wave respondents' data set, where the weights are now taken to be just the increases in the weights assigned to represent the wave nonrespondents. Since, for the purposes of this study, all respondents in the data set were given an initial weight of 1, the increase in weight allocated to the ith three-wave respondent is simply $(w_i - 1)$, where w_i is the weight assigned to compensate for the wave nonresponse.

Table 10 compares the response distributions across the three waves for three items for wave nonrespondents for (a) the actual responses, (b) the data with wave nonrespondents' missing values imputed by the carry-over imputation procedure and (c) the data with the three-wave respondents' values weighted by $(w_i - 1)$. Several features of the imputed results may be noted. First, the distributions for the imputed data have zero entries for the patterns YNY and NYN; in fact, these patterns cannot occur among wave nonrespondents with the carry-over imputation procedure. Secondly, the patterns YYN and NNY occur rarely in the imputed data set; they can arise only from the XOX response pattern, and this pattern occurs infrequently. Thirdly, the imputed data set consistently overestimates the frequencies of the consistent patterns YYY and NNN: these patterns are indeed the only patterns that can occur with the response pattern XOO. As a result of these effects, the imputed distributions deviate systematically from the actual distributions.

On the other hand, the weighted distributions show no systematic deviations from the actual distributions. There is, for instance, no tendency to overrepresent the consistent patterns at the expense of the inconsistent ones. The weighted distributions do, however, differ from the distributions of actual values in a few places.

As a summary of Table 10, Table 11 presents the percentages of "Yes" responses for each of the three items by wave. As can be seen from the table, the percentages of "Yes" responses from the actual and imputed data sets are the same at the first wave, despite the differences in the distributions across waves noted in Table 10. In fact, these two percentages are necessarily equal, because first wave responses are available for all, both three-wave respondents and wave nonrespondents. Hence no imputations are needed at the first wave. On the other hand, with weighting adjustments, the first wave responses are not retained. In consequence, the percentages of first wave "Yes" responses do differ between the actual and weighted data sets.

As noted in Section 3, the carry-over imputation procedure leads to biased estimates when the level of endorsement of an item changes across waves. Evidence of this bias can be seen in the imputed second wave percentages having a job and having savings accounts. In both cases, the actual percentages having the attribute

Table 10

Distributions of Responses across Waves for Three Items for the Wave Nonrespondents (a) with the Actual Responses (b) with Imputed Responses for Missing Waves and (c) with Weighting Adjustments for Wave Nonrespondents

Y=Yes N=No	(a) Actual %	(b) Imputed %	(c) Weighted %
Having a Job			
YYY	58.1	63.3	57.4
YYN	2.4	0.4	2.4
YNY	2.5	–	2.5
YNN	3.2	2.6	3.1
NYY	2.5	1.5	2.6
NYN	0.7	–	0.7
NNY	2.7	0.4	2.7
NNN	27.8	31.8	28.6
	100.0	100.0	100.0
Receiving Social Security Income			
YYY	14.4	14.8	14.7
YYN	0.3	–	0.1
YNY	0.1	–	0.1
YNN	0.3	0.2	0.2
NYY	0.3	0.3	0.6
NYN	0.2	–	0.1
NNY	0.6	0.1	0.6
NNN	83.9	84.6	83.7
	100.0	100.0	100.0
Having Savings Accounts			
YYY	45.1	49.9	48.9
YYN	2.4	0.7	2.7
YNY	1.2	–	1.2
YNN	4.4	2.4	3.3
NYY	2.8	1.3	2.7
NYN	0.2	–	0.4
NNY	2.4	0.8	2.3
NNN	41.5	44.9	38.5
	100.0	100.0	100.0
No. of persons (sum of weights)	1846	1846	(1846)

Table 11

Percentages of "Yes" Responses at Each Wave for Three Items for the Wave Nonrespondents (a) with the Actual Responses, (b) With Imputed Response for Missing Waves, and (c) with Weighting Adjustments for Wave Nonrespondents

	(a) Actual %	(b) Imputed %	(c) Weighted %
<u>Having a Job</u>			
Wave 1	66.2	66.2	65.4
Wave 2	63.7	65.2	63.1
Wave 3	65.8	65.2	65.2
<u>Receiving Social Security</u>			
Wave 1	15.1	15.1	15.1
Wave 2	15.2	15.1	15.5
Wave 3	15.4	15.2	16.0
<u>Having Savings Accounts</u>			
Wave 1	53.1	53.1	56.1
Wave 2	50.5	51.9	54.7
Wave 3	51.5	52.0	55.1

declined from the first to second waves. The carry-over imputation procedure dampens down the amount of decline, so that the second wave imputed estimates are too high. As a consequence, the imputed data set gives underestimates of the amount of net change: for instance the actual change between the first and second waves in the percentages having a job is − 2.5%, whereas the imputed data set shows a change of only − 1.0%. The weighted estimates of change do not suffer this distortion; although they appear less stable, they give better measures of net change.

An even more serious problem with the carry-over imputation procedure is its effect on gross change. All carry-over imputations involve no change, so gross change is underestimated. As an illustration, the actual percentage of wave nonrespondents changing between having and not having jobs from the second to third waves is 8.3%.

The estimate from the weighted analysis is 8.3%, but that from the imputed data set is only 0.8% (arising from the XOX response pattern).

Table 12

*Monthly Mean Social Security Incomes for Wave Nonrespondents Receiving Such Income (a) with the Actual Responses, (b) with Carry-Over Imputed Values for Missing Waves, (c) with Carry-Over Imputed Values for Missing Waves Adjusted for the January Increase, and (d) with Weighting Adjustments for Wave Nonrespondents (Differences from actual monthly means in parentheses)**

(a) Actual $	(b) Imputed $	(c) Adjusted $	(d) Weighted $
388	388 (0)	388 (0)	386 (-2)
395	395 (0)	395 (0)	386 (-9)
389	389 (0)	389 (0)	385 (-4)
387	387 (0)	387 (0)	386 (-1)
381	382 (+1)	382 (+1)	388 (+7)
383	382 (-1)	384 (+1)	390 (+7)
387	386 (-1)	390 (+3)	394 (+7)
390	387 (-3)	393 (+3)	398 (+8)
400	391 (-9)	396 (-4)	399 (-1)
395	391 (-4)	396 (+1)	400 (+5)
398	391 (-7)	396 (-2)	401 (+3)
399	391 (-8)	396 (-3)	401 (+2)

*Excluding monthly amounts of $1500 or more.

Finally, Table 12 presents the means of the monthly Social Security amounts for the wave nonrespondents receiving such amounts or imputed to be receiving such amounts. The figures in this table represent the survey results that would be obtained by the different adjustment procedures for this class of individual. Unlike Tables 7 and 8, the columns do not relate to the same set of individuals. In particular, individuals starting to receive Social Security payments after the point at which they were simulated to be wave nonrespondents are included in the calculations of the means of the actual amounts in column (a), and individuals who ceased to receive amounts but were assigned amounts by the carry-over imputation procedures are included in the calculations of the imputed means in columns (b) and (c). Since those starting and ceasing to receive Social Security amounts tend to receive below average amounts, the means of the actual and the imputed amounts for the third wave in Table 12 are lower than those that applied for Tables 7 and 8. The general conclusions are, however, the same: the simple carry-over imputation procedure underestimates the means for the last six months, but the allowance for the January increase in the modified procedure (column (c)) provides a reasonable correction for this bias.

The weighted means deviate more from the actual means than do the means for the adjusted imputed amounts. In the first four months, the imputed means are necessarily equal to the actual means because there is no wave nonresponse at the first wave. In the second four months, the imputed means still include actual values for almost half of the wave nonrespondents (i.e., those in the pattern XXO). This fact helps to explain why the imputed means track the actual means more closely.

6. Discussion

The preceding results are extremely limited in scope, but they nevertheless do identify some factors involved in making the choice between cross-wave imputation and weighting for handling wave nonresponse. A prime consideration for imputation is the availability of auxiliary information with high predictive power for the missing waves. The few examples investigated in this study agree with other results (e.g., Kalton, Lepkowski and Lin, 1985) that many of the types of variables included in the SIPP are very stable over time. Thus, the values of the variables on a missing wave can be well predicted by the values of the same variables on another wave.

The carry-over, or direct substitution, imputation procedure is one way for utilizing the available wave data for cross-wave imputations. The procedure has a notable advantage of great simplicity, but as our analyses have illustrated it fails to track net changes in means or proportions when these vary over time. The extent of bias in the survey estimates caused by this failure depends on the degree of net change that occurs and the amount of wave nonresponse. It will be small when there is not much net change and a low level of wave nonresponse, as will often be the case. More seriously, the carry-over imputation procedure causes an underestimation of gross change, since all imputed values are assigned the same response as the last available wave. This simple procedure causes the amount of gross change to be underestimated by a proportion equal to the proportion of carry-over imputations.

Kalton and Lepkowski (1983) describe some alternatives to the carry-over imputation procedure that avoid the distortions caused by this simple procedure. These procedures take account of changes over time by imputing changes for some wave nonrespondents. Thus, for instance, if 8% of the respondents change from having to not having a job between the first and second waves, 8% of second wave nonrespondents with jobs at the first wave would be assigned changes (and this can be extended to be applied separately, with different rates of change, in a set of imputation classes). While these procedures are attractive for reflecting change, they suffer other disadvantages. Unless great care is taken, they may lead to the imputation of sets of responses that are inconsistent, and in any case they will cause distortions in the relationships between some of the responses (see Kalton and Kasprzyk, 1982, Section 3.3). The simple carry-over procedure retains the relationships between responses that occur on the wave used for imputation; provided that these relationships do not change over time, this is an attractive feature.

As our study of the imputation of Social Security amounts brought out, even the carry-over imputation procedure should not be applied uncritically with numerical variables. Social Security amounts in general fall within definite limits, but nevertheless some outliers do occur. In the simulation data set, there was, for instance, one person who received $4359 in one month, nothing in the previous month, and only $337 in each of the two subsequent months. Another person purportedly received $2242 in one month, $242 in the preceeding month, and $251 in each of the two subsequent months (an amount 3.5% larger than the $242 amount). While some of the outliers may be erroneous values (as seems probable in this second case), they cannot always automatically be treated as such because large payments in a single month are possible.

The assignment of these large amounts to subsequent months by the carry-over imputation procedure would however create unrealistic longitudinal records.

Weighting has the attraction over imputation that it avoids the above problems. The weighting scheme employed in the simulation study, however, suffers the disadvantage that it discards a good deal of information: first wave responses are available for all wave nonrespondents, but apart from those used in forming weighting classes, these responses are discarded; similarly, second and third wave responses are available for one-half and one-tenth of the wave nonrespondents, respectively, but they are also discarded. This discarding of data can be avoided by the use of several different sets of weights, but this solution adds to the complexity of the data set, and it can lead to inconsistencies in the results of different analyses. In addition to this discarding of actual responses, weighting does not take advantage of the high predictability of many of the wave nonrespondents' missing values that cross-wave imputation employs.

No measure of the effective sample size is available for the situation where imputation is used to handle missing responses. Table 1 shows that there was 10% of wave nonresponse in the first three waves of the 1984 SIPP Panel. However, only 4.7% of these responses were missing, and moreover many of the missing responses could be imputed with little error from other waves. Thus it seems that the effective sample size is only a few percentage points below the first wave sample size. The sample size when the simple single set of weights is used is 10% lower than that of the first wave, and in addition the use of weights decreases the effective sample size still further. This further decrease may be approximately measured by the multiplying factor $(\Sigma w_i)^2/(n\Sigma w_i^2)$, where w_i is the weight of the ith sampled element. In the simulation data set, this factor is very close to 1 because of the small variation in the weights. Thus, the effective sample size with the weighting solution is about 90% of the sample size at the first wave.

The choice between imputation and weighting for handling wave nonresponse is complicated by the fact that the survey data will be subjected to many types of analyses, involving different forms of estimates and being based on varying-sized subclasses of the total sample. Since imputation can distort some forms of estimates, weighting may be the preferred solution for large subclasses when the reduction in effective sample size is tolerable. However, imputation may be better for estimates based on small subclasses, when the loss in effective sample size matters and when any bias caused by imputation is less important relative to the sampling error. The choice of one or other of these adjustment procedures for multipurpose use must balance out these considerations. In the case of the three-wave SIPP file, the difference in the effective sample sizes between the imputation and weighting solutions is not great, and therefore weighting may be the safer general purpose solution.

References

Burkhead, D. and Coder, J. (1985). Gross changes in income recipiency from the Survey of Income and Program Participation. *Proceedings of the Social Statistics Section, American Statistical Association*, 351-356.

Kalton, G. (1985). Handling wave nonresponse in longitudinal surveys. *Proceedings of the Bureau of the Census First Annual Research Conference*, 453-461.

Kalton, G. and Kasprzyk, D. (1982). Imputing for missing survey responses. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 22-31.

Kalton, G. and Lepkowski, J. (1983). Cross-wave item imputation. In *Technical, Conceptual, and Administrative Lessons of the Income Survey Development Program (ISDP)*, M. David, ed., pp. 171-198. New York: Social Science Research Council.

Kalton, G., McMillen, D.B. and Kasprzyk, D. (1986). Nonsampling error issues in the Survey of Income and Program Participation. *Proceedings of the U.S. Bureau of the Census Second Annual Research Conference*, forthcoming.

Kalton, G., Lepkowski, J. and Lin, T. (1985). Compensating for wave nonresponse in the 1979 ISDP Research Panel. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 372-377.

Moore, J.C. and Kasprzyk, D. (1984). Month-to-month recipiency turnover in the ISDP. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 726-731.

Nelson, D., McMillen, D.B. and Kasprzyk, D. (1985). An overview of the Survey of Income and Program Participation: Update 1. SIPP Working Paper Series No. 8401, Washington, D.C.: U.S. Bureau of the Cênsus.

Sonquist, J.A., Baker, E.L. and Morgan, J.N. (1973). *Searching for Structure*. Ann Arbor, Michigan: Institute for Social Research.

Chapter 4

LONGITUDINAL IMPUTATION FOR THE SIPP[4]

Steven G. Heeringa and James M. Lepkowski

1. Introduction

The problem of item nonresponse in a survey arises when an otherwise cooperative respondent does not or cannot provide a response to one or more survey questions. Imputation, the estimation of a value for a missing response, is commonly used to compensate for such item missing data. Item nonresponse and its compensation methods become more complex in the case of a panel survey where a sample of respondents provides data at a series of points in time. In a panel survey, the item nonresponse problem can be extended to include wave nonresponse, that is, failure to obtain any data from a respondent at one or more waves of the data collection sequence. Whether the data are missing for an entire wave or only for specific items within a wave, longitudinal survey data can provide additional information which may be used to improve the quality of imputation for missing values (Kalton and Lepkowski, 1982).

Since panel data are usually collected and processed one wave at a time, imputation of missing values is often conducted for each wave separately using only the information available within a wave to derive an imputed value. Such "cross-sectional" imputations do not take advantage of the information collected at other waves of the panel. In contrast, longitudinal imputation methods have the capability to use data collected at other waves, data which may be highly correlated with the item to be imputed.

The purpose here is to examine longitudinal and cross-sectional imputation methods for item missing data in the Survey of Income and Program Participation (SIPP). The investigation reported in this paper uses selected survey variables from the first three

---

[4]From *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 1986, forthcoming.

waves of the 1984 SIPP panel to compare the effectiveness of a simple longitudinal direct substitution technique and that of the Census cross-sectional hot-deck imputation method. After describing the SIPP design and cross-sectional hot-deck imputation method in Section II, we review some longitudinal imputation methods that could be applied to the SIPP in Section III. In Secton IV, the simple longitudinal imputation method that is applied to the SIPP file is described, and Section V compares the longitudinal and cross-sectional imputations. The paper concludes with remarks about further investigations that might be conducted.

2. SIPP Design

The SIPP is a national survey of U.S. households conducted by the Bureau of the Census. It is designed to provide comprehensive information on both households' and individuals' economic status and participation in government programs. It is a panel survey in which households that participate in a baseline interview are followed and interviewed at 4 month intervals for a total of eight interviews. Interviewing for the 1984 SIPP panel began in October 1983 with an equal probability sample of about 20,000 households. (See Nelson, McMillen and Kasprzyk (1985) for a full description.)

The SIPP is designed to meet a range of analytic objectives. Some analyses involve the data for a single wave while others require data from several waves (e.g., analyses of annual incomes). Cross-sectional data collected at each wave of the SIPP are used to provide important estimates for quarterly reports on income and program participation. For this purpose, each wave of the SIPP panel is processed as a separate cross-sectional survey, and item missing data at each wave are handled by cross-sectional imputations.

The Bureau of the Census currently uses a cross-sectional hot-deck (CSHD) imputation for selected item nonresponse on individual waves of the SIPP (Nelson, McMillen and Kasprzyk, 1985). The first step in the CSHD procedure is to define a "hot-deck" matrix based on a cross-classification of characteristics that are correlated with the item being imputed. Based on the cross-classifying variables, each individual record is uniquely linked to a cell of the hot-deck matrix. To initialize the procedure, a "cold-deck" or starting value is assigned to each cell of the hot-deck matrix. The complete SIPP data file is then sorted by geographic characteristics and is passed through the hot-deck imputation program two times. In the first pass, no imputations are made, but if an observation has a non-missing value for an item to be imputed, that value "updates" the current value for the item stored in the hot-deck matrix.

In the second pass of the data, the actual imputation of missing values takes place. In the sequential order of the file, each record is examined and if the item is missing, the current value stored in the hot-deck cell for that item replaces the missing value on the record. If the value of the record is not missing, the non-missing value for that case replaces the current donor value for the hot-deck matrix cell. Thus, missing values for a record are, for the most part, replaced by values from another record that has the same characteristics used to define the hot-deck cell. For each item receiving imputations, an indicator variable is added to the SIPP file identifying which values have been imputed (Bureau of the Census, 1985).

3. Longitudinal Imputation Methods and Models

Longitudinal methods are designed to utilize cross-wave data in imputing the value of a missing item (Kalton and Lepkowski, 1982). However, the exact form in which the cross-wave information is used differs from one techinque to another. Five general classes of longitudinal imputation methods might be considered as an alternative to the CSHD method:

1) **Longitudinal direct substitution**. For items that are stable over time, the value of a nonmissing item is substituted from one time period to another where the same item is missing. Direct substitution can be a highly accurate form of imputation in some situations.

2) **Deterministic imputation of change**. Additive or proportionate change from one time period to another can be computed from the survey data or obtained

from an exogenous source. Imputed values are created by applying this change to a non-missing value from an another wave.

3) **Longitudinal regression imputation**. Missing values are predicted from a regression equation obtained by fitting a model to data with nonmissing values. In the prediction, the residual term in the model can be set to zero for a deterministic form of regression imputation, or it can be assigned a value through a hot-deck or other stochastic procedure.

4) **Longitudinal hot-deck**. Auxiliary cross-wave information available from the longitudinally linked records is used to form the cells of the hot-deck matrix, extending the characteristics used in the CSHD procedure. Continuous items must be categorized to form the cells of the hot-deck matrix, reducing the strength of the cross-wave correlations. Nonetheless, the strength of correlations over time for stable items improves the accuracy of the CSHD procedure.

5) **Longitudinal hot-deck imputation of change**. Longitudinal hot-deck procedures are used to impute change from a donor record to the case with the missing value. The imputed change can be added directly to a nonmissing value from a prior or succeeding wave or another wave's nonmissing value can be proportionately altered.

Under these five general longitudinal imputation strategies, the value imputed for the ith respondent with missing data is derived as $y_i = f(x_{1i}, x_{2i}, \ldots, x_{pi}) + e_i$ where $f(\cdot)$ is a function of p auxiliary variables and e_i is an estimated residual. For the five general strategies the function $f(\cdot)$ can be expressed as a linear function where $y_i = b_0 + b_1 x_{1i} + \ldots + b_p x_{pi} + e_i$, and the b_j's are estimated from data for respondents with no missing values for y_i or the auxiliary variables.

Figure 1 presents simple linear models corresponding to the five general strategies to illustrate the relative features of the longitudinal imputation strategies. The simplest model is associated with the longitudinal direct substitution (LDS) method in which a

Figure 1

Models for Longitudinal Imputation Methods

Method	Model	Component of Change		
		Proportionate	Additive	Stochastic
Direct Substitution	$y_i = x_i$	None	None	None
Deterministic Imputation of Change	$y_i = cx_i$ or $y_i = cx_i + a$	$(1-c)x_i$ $(1-c)x_i$	None a	None None
Longitudinal Regression	$y_i = b_o + b_1 x_i + e_i$	$(1-b_1)x_i$	b_o	e_i
Longitudinal Hot Deck	$y_i = b_o + b_1 x_i + e^*_{j \ddagger i}$	$(1-b_1)x_i$	b_o	$e_{j \ddagger i}$

nonmissing value is essentially "carried over" from another wave. Each of the other methods can be viewed as a modification of the LDS strategy incorporating proportionate change, additive change, and stochastic variation. For example, the deterministic imputation of change method can improve the LDS method by including an additive component of change (a), a proportionate change (cx_i), or both additive and proportionate change $(a + cx_i)$ to the "carry-over" LDS imputation.

From this perspective, the LDS method may be viewed as a base longitudinal imputation procedure to which modifications can be made to address deficiencies in the quality of the LDS imputations. As an initial investigation of the general longitudinal approach, a comparison of the LDS to the CSHD imputations will indicate whether longitudinal methods improve the quality of imputed values. Thus, the subsequent discussion examines the LDS as a base longitudinal imputation method relative to the CSHD imputations available in the SIPP data files.

Although the LDS method is conceptually simple, implementation can be complicated, because cross-wave information may not be available for each record with missing data on one wave. The general LDS strategy employed in this study was essentially a two step process. When an item could be carried over longitudinally, the imputation was made. Otherwise, the Census CSHD imputed value was used as the imputed value.

The LDS method has also been implemented somewhat differently for categorical and continuous types of variables. For categorical variables, the records with imputed responses (i.e., with missing data that has been replaced by the CSHD method) were scanned to determine if an actual value was available at a prior (or a subsequent) wave. If so, the actual value from the alternate wave was imputed for the missing item. If no value was available, the original CSHD imputed value was left unchanged. When two "donor" values were available, but different in value, the value from the "nearest" data collection wave was imputed for the missing item. For continuous variables, the LDS imputation algorithm also scanned the longitudinal data record to identify the full set of potential donor items for a missing value. But instead of selecting one member of the set as a "donor", the *average* of all nonmissing values was imputed for the missing item.

Finally, some SIPP variables such as earnings and wages undergo both systematic changes and random fluctuation across time. Therefore, short of performing the evaluation on a complete data set where both the amounts and patterns of missing values are simulated, it is difficult to choose an appropriate benchmark to measure the accuracy of imputations. Simulation can be a useful tool (Kalton and Lepkowski, 1982), but for the current study it has several drawbacks. First, since the simulation must operate on a data set with no missing values, the extension of the results to a full data set requires strong assumptions (or knowledge) about the distributions of the missing and non-missing values. Secondly, simulation of "missingness" would have to be carried out separately for each variable under study. This would require a large investment in set-up time and computing funds. By necessity then, the comparison of the CSHD and LDS imputation methods is presented here simply as a demonstration of what happens to actual distributions of these variables under the two imputation alternatives.

4. Implementation of the Longitudinal Direct Substitution Method

Using data from 1984 SIPP Panel, an empirical investigation was conducted to test the feasibility and effectiveness of simple longitudinal imputation as an alternative to imputations based solely on cross-sectional hot deck methods.

The empirical study used a longitudinal file created from the first three waves of the 1984 SIPP panel. The Bureau of the Census cross-sectional public use files of data collected in the first three waves were merged to create longitudinal records of various types. The fourth rotation group of the original 1984 SIPP sample was excluded from the longitudinal file because data were not collected for the group in the second wave. From the sample households included in the first three rotation groups, a total of 31,161 individuals aged 15 and older by the end of Wave 3 had data on at least one of the three waves. A total of 26,992 of these persons had data at all three waves.

Each person could have had up to four wage-earning jobs on each wave. Each job is represented by a Wage and Salary record which can be linked to a person in the file. CSHD imputations were made to a limited number of items on these Wage and Salary records. The empirical work reported here focuses on three categorical and two continuous variables from the Wage and Salary record for which CSHD imputations were made where needed. The categorical variables were 1) occupation code, 2) employer category, and 3) frequency of pay. The continuous variables were the wage rate for hourly paid jobs and total monthly earnings for each of four reporting months in a single wave. Each of the three categorical and five continuous items (wage rate plus four monthly earnings) can be reported for each of three waves in the 1984 SIPP Panel. The merged data set contains longitudinal Wage and Salary records for a total of 23,005 job reports: 19,223 reports for individuals' first jobs; 2978 for the second jobs; 684 for the third jobs; and 120 for the fourth jobs. To simplify the presentation, results from only the first job are given.

Table 1 presents counts of item responses, both total and missing, by wave for the Job 1 Wage and Salary variables of interest. Among these variables, the item missing data rates for the categorical items are very small, ranging from a low of 0.16% item missing data for the Wave 1 employer category variable to a high of 2.45% for the Wave 3 frequency of pay question. The percentages of item missing data among earnings items are higher, particularly at the first wave of data collection: 9.37% item missing data for reports of Job 1 monthly earnings in Wave 1 of the 1984 SIPP Panel. However, item missing data rates for Job 1 monthly earnings drop substantially at Waves 2 (2.97%) and 3 (3.16%).[5] At 9.7%, the Wave 1 item missing data rate for

[5]A comparison of the total response counts across waves shows a decline in the number of cooperating respondents who hold Job 1, the sharpest drop occurring between Waves 1 and 2 of the panel. From one wave to the next, the change in the number of Job 1 reporters is a function of both panel attrition due to Type A (household) and Type Z (individual) nonresponse and responding individuals who no longer hold Job 1 at a later wave.

Table 1

Item Nonresponse in the 1984 SIPP Wage and Salary Data

Job 1 Variable	Wave	Item Responses	Imputed Values		Percent of missing values for which longitudinal imputation is possible
			Number	%	
Occupation	1	17,110	90	.53	56.6
	2	15,766	101	.64	85.1
	3	15,196	85	.56	78.8
Employer Category	1	17,110	62	.36	70.7
	2	15,766	44	.28	86.4
	3	15,196	25	.16	76.0
Pay Frequency	1	17,110	316	1.85	56.9
	2	15,766	362	2.30	75.1
	3	15,196	373	2.45	82.5
Monthly Earnings*	1	68,440	6,410	9.37	68.9
	2	63,064	1,875	2.97	41.1
	3	60,784	1,918	3.16	61.1
Hourly Wage	1	10,258	993	9.70	39.4
	2	9,476	1,103	11.60	54.5
	3	9,141	1,038	11.40	57.2

*Item response totals are 4 monthly responses for each wave.

hourly wage reports is also relatively high but, unlike Job 1 monthly earnings, the missing data rate for this variable rises slightly at Waves 2 and 3.

It is important to know not only the rate at which responses are missing but also what proportion of these missing values can be imputed longitudinally. LDS imputation is possible only when the missing item has actually been observed at a preceding or succeeding wave. Table 1 also indicates the extent to which missing items can be imputed longitudinally. Among the categorical variables, the percentage of missing responses which can be imputed by direct substitution from another wave ranges from 56.5% to 86.4%. Similarly, longitudinal imputation of missing data on the earnings items appears promising. For example, almost 69% of the missing values for Job 1 monthly earning at Wave 1 could be imputed using the LDS procedure.

5. Comparison of the CSHD and LDS Imputation Methods

Once the LDS imputations were made, the effect of the CSHD and the LDS imputations on distributions of the categorical and continuous variables of interest could be examined. In this section simple frequency distributions and distributions of change in individual reports from one wave to the next are compared between CHSD and LDS imputed values for each of the variables of interest. *The LDS method examined here uses the original CSHD imputed value whenever a substitute value was not available on an alternate wave. Thus, results for the LDS method will incorporate a proportion of missing value cases which were imputed by the secondary CSHD technique.*

Due to the very low rates of item missing data, CSHD and LDS imputations should not be expected to have widely differing effects on the overall frequency distributions of the categorical variables. For Job 1 Wave 1 employer category and frequency of pay, there is no difference in the distributions whether the CSHD or the LDS method is used to impute for missing data. Similar results were observed for Waves 2 and 3 and for other categorical type variables.

The categorical variables are essentially job descriptors, and given that the job was not changed, their values should not be expected to change significantly from one wave to the next. However, Table 2 indicates that, even in instances where no imputation is involved, a wave to wave change in response value for these variables can occur in as many as 20% of cases. It is difficult to say what proportion of this observed change is real, as opposed to a reflection of response error or coding inconsistency.

Table 2

Wave to Wave Consistency in Categorical Variable Values Under the CSHD and LDS Imputation Methods

Job 1 Variable	Wave Comparison	No Imputation		One or Both Waves Imputed		
		n	% Agreement	n	CSHD % Agreement	LDS % Agreement
Frequency of Pay	1 to 2	14,079	81.3	475	45.8	89.3
	2 to 3	13,111	79.8	478	46.8	94.7
Employer Category	1 to 2	14,477	95.6	77	77.9	97.4
	2 to 3	13,546	95.4	43	69.7	100.0
Occupation Code	1 to 2	14,425	78.4	129	26.4	72.1
	2 to 3	13,470	78.8	119	19.3	78.8

If cases where one or both values have been imputed are compared to cases without imputations the CSHD imputations lead to a significant reduction in wave to wave response consistency. On the other hand, the LDS imputation method produces a high level of cross-wave consistency for these job descriptors. In fact, one might view the LDS method as overriding the observed natural variation in responses and thereby forcing an artifically high level of wave to wave consistency.

The drop in cross-wave consistency for the job records with CSHD imputed values is so large that it suggests that the level of agreement across waves might be explained by "chance" alone. In the case of the hot-deck method, a discrete response category is modeled as an ANOVA-type function of a series of categorical factors (e.g., hot deck variables such as age, sex, race, education). If the model is weak, the probability of a correct imputation degenerates to the multinomial probability of agreement between the true value and a "random" imputation. The greater the number of response categories and the more uniform the odds across categories, the more difficult it is to impute the correct (or matching) value. The Wage and Salary categorical variables for which the CSHD imputation results in high wave to wave consistency do in fact have either few categories or highly unequal odds across categories.

For example, the employer category variable with six response categories has as the largest category "private company" with 82% of the cases. By simply imputing the code value for this largest category to each missing item, we might expect to be correct about 82% of the time. For this variable, even a random imputation of respondents' values will, in expectation, impute a matching value 69% of the time. In Table 2, a two-wave comparison involving CSHD imputations for this variable shows 78% agreement from Wave 1 to 2 and 70% agreement from Wave 2 to 3.

Although the small sample sizes and limited set of variables prevent us from drawing any firm conclusions, the data suggest that the CSHD imputation of these job descriptors provides only small increases in accuracy relative to what we might expect by chance alone.

Considering the continuous variables, Tables 3 and 4 compare characteristics of the earnings variables after CSHD and LDS imputations have been made for item missing data. Table 3 compares the sum of up to four monthly Job 1 earnings values for each wave of data collection; Job 1 hourly wage rates are compared in Table 4. The upper panel of each table presents findings for all cases, both those with nonmissing data and those where a missing amount has been imputed. The lower panel of each table presents only those cases where one or more component earnings amounts have been imputed.[6]

The basic and not unexpected result found in Tables 3 and 4 is that even with item missing data rates of almost 10% at Wave 1, the choice of CSHD or LDS imputation appears to have only a small effect on the statistics examined.

---

[6]Since earnings reports are taken for each month of the reference period it is possible to have actual and imputed values in the same wave. In such cases, the wave earnings totals will be the aggregate of actual and imputed monthly amounts.

Table 3

Imputation of Job 1 Earnings. Comparison of Sample Earnings Distributions After CSHD and LDS Imputation for Item Missing Data

Wave	Statistic	Imputation Method	
		CSHD	LDS
All Job 1 Reports			
Wave 1 (n=16,895)	Mean	$4,796	$4,750
	Std.Dev.	4,199	4,140
	Skewness	1.94	1.94
	Kurtosis	6.95	7.00
Wave 2 (n=15,569)	Mean	$5,041	$5,051
	Std.Dev.	4,222	4,239
	Skewness	1.94	1.96
	Kurtosis	6.85	7.03
Wave 3 (n=14,994)	Mean	$5,128	$5,142
	Std.Dev.	4,260	4,294
	Skewness	1.88	1.93
	Kurtosis	6.54	6.88
All Job 1 Reports With One or More Imputed Amounts			
Wave 1 (n=2,688)	Mean	$5,510	$5,225
	Std.Dev.	4,491	4,174
	Skewness	2.32	1.70
	Kurtosis	8.23	9.38
Wave 2 (n=485)	Mean	$7,576	$7,896
	Std.Dev.	6,115	6,343
	Skewness	1.70	1.78
	Skewness	3.60	3.86
Wave 3 (n=494)	Mean	$7,447	$7,859
	Std.Dev.	5,943	6,485
	Skewness	1.73	1.82
	Kurtosis	3.62	3.94

Table 4

Imputation of Job 1 Hourly Wage Rates. Comparison of Sample Income Distributions After CSHD and LDS Imputation for Item Missing Data

Wave	Statistic	Imputation Method	
		CSHD	LDS
All Job 1 Reports			
Wave 1 (n=10,456)	Mean	$6.58	$6.60
	Std.Dev.	3.67	3.67
	Skewness	2.23	2.21
	Kurtosis	16.27	16.27
Wave 2 (n=9,416)	Mean	$6.73	$6.76
	Std.Dev.	3.84	3.84
	Skewness	3.41	3.38
	Kurtosis	37.74	37.96
Wave 3 (n=9,078)	Mean	$6.75	$6.76
	Std.Dev.	3.92	3.79
	Skewness	4.23	3.33
	Kurtosis	61.33	42.96
All Job 1 Hourly Wage Reports With One or More Imputed Amounts			
Wave 1 (n=993)	Mean	$7.23	$7.34
	Std.Dev.	3.81	3.73
	Skewness	1.40	1.25
	Kurtosis	2.29	1.73
Wave 2 (n=1,103)	Mean	$7.18	$7.39
	Std.Dev.	3.99	3.97
	Skewness	2.24	2.09
	Skewness	13.23	12.46
Wave 3 (n=1,038)	Mean	$7.47	$7.57
	Std.Dev.	4.96	3.98
	Skewness	6.32	2.33
	Kurtosis	83.97	15.50

The findings in Tables 3 and 4 indicate that univariate analyses of the SIPP Wage and Salary earnings data will not be greatly affected by the imputation methodology that is used. However, the data presented here give no indication of the effect these imputation methods have on univariate distributions for population subclasses or domains. Furthermore, the result for descriptive univariate statistics has no implicit generalization to bivariate and multivariate analyses.

One form of longitudinal analyses of SIPP data is to examine how and why individual income and earnings change over time. For this kind of analysis, information is needed on the effects of CSHD and LDS imputations on distributions of micro-level change in earnings. Table 5 presents the distribution of Wave 1 to 2 and Wave 2 to 3 changes in individual respondents' Job 1 earnings. Columns (3) and (4) compare the change distributions for all cases (actual and imputed) having a nonzero earnings amount at each wave. Column (5) restricts the change distribution to cases where two actual reports were obtained. Sample distributions of change involving actual-imputed and imputed-actual combinations of values are described in columns (6) - (9).

Over time, it is expected that the average wages and earnings of panel respondents should follow an increasing trend. Looking at Table 5, the overall distribution of change (columns 3 and 4) does show, as expected, a positive increment in Job 1 earnings between successive waves. For the Wave 1 to 2 change, the average amount of this increase is appreciably lower when the CSHD method is used to impute for item missing data. Examination of standard deviations and percentiles shows that CSHD imputation both increases the variability and elongates the tails of the sample distribution of wave to wave change in earnings. In fact, if the change computation is restricted to pairs with one CSHD imputed value and one actual response, the result is a distribution which is highly variable and has many extreme values. Because the number of extreme changes imputed by the CSHD method is so large, the distributional statistics -- particularly the means -- reported in Columns (6) - (9) should be viewed as highly unstable. These statistics are reported here primarily as evidence of the variability

Table 5

Wave to Wave Change in Job 1 Earnings. Comparison of Sample Distributions Under CSHD and LDS Imputation Methods

Change Estimate	Sample Distribution Statistic	All Data		Actual–Actual	Actual–Imputed*		Imputed–Actual*	
		All Pairs						
		CHD Imputation	LDS Imputation	No Imputation	CHD Imputation	LDS Imputation	CHD Imputation	LDS Imputation
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Wave 2 – Wave 1	Mean	$104.66	$165.43	$183.45	–$302.28	$24.65	–$948.75	$27.42
	Std.Dev.	2,561	2,243	2,205	2,902	1,679	6,913	1,091
	5th–%tile	–3,199	–2,766	–2,790	–4,490	–2,489	–12,640	–1,232
	95th–%tile	3,484	3,223	3,297	1,974	2,190	8,655	1,750
	(n)	(14,344)	(14,344)	(11,892)	(2,021)	(2,021)	(182)	(182)
Wave 3 – Wave 2	Mean	$139.97	$145.13	$148.69	–$214.76	$0.35	–$94.64	$124.09
	Std.Dev.	2,449	2,243	2,117	6,295	1,406	5,359	1,480
	5th–%tile	–2,974	–2,710	–2,659	–10,082	–1,917	–5,536	–1,709
	95th–%tile	3,400	3,199	3,149	8,939	1,766	–8,792	3,175
	(n)	(13,403)	(13,403)	(12,818)	(142)	(142)	(215)	(215)

*The CHSD imputations –– including default imputations under the LDS method –– produce a high degree of variability in the wave to wave change values. Therefore, the statistics reported in these columns are also highly variable and should be interpreted with caution.

which CSHD imputation can introduce to longitudinal measures such as change in earnings.

Given that a zero change model is implicit in the direct substitution imputations used in this exercise, the LDS method should be expected to compress the wave to wave change distribution about the zero value. In comparing differences between actual and imputed values, columns (7) and (9) indicate that the LDS method of imputing averages of actual values for a missing earnings report results in changes which average just slightly greater than zero. (An exception occurs in the estimates of change between Wave 2 actual and Wave 3 imputed values.) The "compression" effect which the LDS method has on estimates of change is evident in a comparison of percentile statistics for the change distributions. For example, in the sample distribution of change between Wave 2 CSHD-imputed and Wave 3 actual values, the 5th and 95th percentiles are -\$10,082 and \$8,939. For cases where earnings are not imputed at either wave, the 5th and 95th percentiles of the corresponding change distribution are -\$2,659 and \$3,149. The comparable percentiles for Wave 2 LDS-imputed to Wave 3 actual change are -\$1,977 and \$1,766.

6. Concluding Remarks

Cross-sectional hot-deck (CSHD) imputation is a practical and timely method for imputing missing item values on the SIPP Wage and Salary record for an individual wave. However, the evidence presented here suggests that the CSHD method may perform only slightly better than chance at imputing the correct response to a missing categorical item from the wage and salary variable set. CSHD imputations for continuous wage and salary earnings variables do not appear to appreciably alter the distributions of these items. However, the impact on both cross-sectional and longitudinal multivariate distributions is larger.

Given the cross-wave patterns of item missing data observed in the 1984 SIPP Wage and Salary record, the use of longitudinal imputation methods appears to be warranted for SIPP longitudinal files. For categorical variables, the direct substitution method is a practical approach to cross-wave imputations of missing items. For the continuous variables such as Job 1 earnings, the empirical tests clearly demonstrate the desirability of longitudinal imputations for missing data on these items. The LDS method of longitudinal imputation understates change, but this may be preferred to the gross overstatement of change resulting from the use of the CSHD method.

References

McMillen, David B. and Daniel Kasprzyk (1985). "Item Nonresponse in the Survey of Income and Program Participation," *Proceedings of the Survey Research Methods Section, American Statistical Association,* 360–365.

Bureau of the Census (1985). *Survey of Income and Program Participation User's Guide,* 2nd edition. Washington: U. S. Department of Commerce.

Kalton, Graham (1985). "Handling Wave Nonresponse in Longitudinal Surveys," *Proceedings of the Bureau of the Census Research First Annual Conference,* pp. 453–461. Washington: U. S. Bureau of the Census.

Kalton, Graham and James M. Lepkowski (1982). "Cross-wave Item Imputation," in *Technical, Conceptual, and Administrative Lessons of the Income Survey Development Program (ISDP),* Martin H. David, ed., pp. 155–170. Washington: Social Science Research Council.

Nelson, Dawn, David B. McMillen, and Daniel Kasprzyk (1985). "An Overview of the Survey of Income and Program Participation, Update 1," *SIPP Working Paper Series,* No. 8401. Washington: U. S. Bureau of the Census.

CHAPTER 5

THE TREATMENT OF WAVE NONRESPONSE IN PANEL SURVEYS[7]

James M. Lepkowski

Abstract: Unit nonresponse occurs in a survey when data are not collected for a sampled unit, while item nonresponse occurs when data are not collected for a given item for an otherwise responding unit. In panel surveys, partial nonresponse occurs when data are not collected for one or more waves from a unit which responds to at least one wave of the panel. Since loss of an entire wave of data is more extensive than item nonresponse but less extensive than unit nonresponse, wave nonresponse poses different problems for analysis, and different compensation strategies may be appropriate. The nature and extent of wave nonresponse in panel surveys is reviewed and discussed in the context of several large panel surveys. Weighting and imputation are reviewed as compensation strategies for wave nonresponse, and alternative combined weighting and imputation strategies are also described. The variety and complexities of panel survey designs prohibits the recommendation of a single strategy for wave nonresponse in general, but criteria for developing a suitable strategy are reviewed.

Keywords: Weighting, imputation, longitudinal analysis, nested nonresponse, attrition nonresponse

1. Panel Surveys and Wave Nonresponse

Missing data in surveys are generally considered to be of two types. In unit nonresponse survey data are not obtained for a sampled unit, while in item nonresponse an item or limited set of items is not completed for a unit which otherwise provides a complete set of responses. These broad categories of missing data are generally considered to include many other types of missing data patterns. For example, in household surveys in which all eligible persons in a sample household are to be

---

[7]This manuscript was presented to the International Symposium on Panel Surveys, 19 – 22 November 1986, Washington, D.C.

interviewed, missing data for one or more persons in the household is considered to be unit nonresponse for the individuals rather than blocks of item nonresponse for the household. On the other hand, a section of a questionnaire concerning income that is not completed because of respondent refusal or interviewer error is typically considered to be a series of item nonresponses.

Panel surveys have an additional type of missing data that has some of the characteristics of missing data due to both unit and item nonresponse. Partial or wave nonresponse occurs when one or more waves of panel data is missing for a unit that has provided data for at least one other wave.

The amount of missing data and the amount of information available about the nonresponding unit influences the type of imputation strategy employed to compensate for the missing data. All survey data are missing for unit nonresponses; limited sample design information may be the only available data about the nonresponding unit. Missing data from unit nonresponse is typically compensated by weighting. In contrast, item missing data has more extensive data available about the nonresponding unit (*e.g.*, sample design data and responses to other survey items), information that can be used to improve the quality of the compensation. Usually only a few items on a given record require compensation. As a result, imputation, which can be a much more intensive activity than weighting, is typically employed to compensate for item missing data.

The amount of missing data for a record with wave nonresponse is typically more than that encountered for item nonresponses, but data available from completed waves provide more detailed information about the partially nonresponding unit than is available for unit nonrespondents. Thus, weighting, imputation, or a combination of weighting and imputation may be considered as suitable methods for compensating for missing data due to wave nonresponse.

The purpose of this paper is to review the nature of and to examine missing data compensation strategies for wave nonresponse in panel surveys. The review is not intended to cover every conceivable wave nonresponse compensation strategy, but rather the review describes the characteristics, strengths, and weaknesses of several general strategies. Where possible, methods and issues are illustrated with examples from several large panel surveys.

After a review of the general panel survey and wave nonresponse issues in this section, Sections 2 and 3 describe characteristics of weighting and imputation as compensation methods for wave missing data. Section 4 reviews empirical comparisons of weighting and imputation compensations in data sets with simulated missing data. The paper concludes with an examination of combined weighting and imputation strategies in Section 5, and a discussion in Section 6 of criteria that might be applied in the selection of a suitable strategy for wave nonresponse missing data compensation.

Before reviewing specific wave nonresponse compensation strategies, it is useful to consider several issues in panel survey design that have an effect on the quality of various wave nonresponse strategies: the purposes of panel surveys, the types of data collected in surveys, and wave nonresponse patterns are briefly examined.

Panel surveys collect the same information from the same sample elements over several different data collection periods (Duncan and Kalton, 1985). The periods of data collection or waves of the panel may be temporally contiguous or they may be separated by periods with no data collection activity. The data may be collected each wave about the same reference period (*e.g.*, the period January 1 to December 31 , 1985) for all panel members, or the reference periods may be of varying lengths covering somewhat different time periods (*e.g.*, since the last interview). Data collection reference periods are generally contiguous reflecting the longitudinal purposes of data collection.

Although panel survey data are typically considered most useful for collecting longitudinal information, particularly the measurement of change over time, panel surveys may be used to collect other information as well. For example, the panel data may be analyzed at a fixed point in time to obtain cross-sectional information about a population, even though panel surveys are clearly not the most appropriate design for obtaining cross-sectional information. Panel surveys are also used to cumulate information over time. Information known to be subject to sizeable response errors due to recall loss is collected in a panel survey design to decrease reference periods and hence reduce recall errors. The panel data are cumulated over reference periods to provide more accurate information about the cumulated reference period. Cumulation may occur for intervally scaled data which can be summed across reference periods, or it may involve the cumulation of rare events or the construction of event histories over extended time periods. For the purposes of discussion of wave nonresponse compensation strategies, it will be important to distinguish these three purposes for panel surveys: longitudinal comparisons, cross-sectional estimation, and longitudinal cumulation.

The purpose of a panel survey must be taken into account when considering missing data compensation strategies for wave nonresponse. A compensation method which provides reasonably accurate results for one purpose may provide poor prediction for another. For example, an imputation suitable for cross-sectional analysis of panel data may introduce changes from imputed data on one wave to nonimputed data on another that are clearly inappropriate (Heeringa and Lepkowski, 1986).

Survey data may consist of many different types of measures. The elements themselves may consist of intervally scaled or continuous measures, such as income from several different sources. The data elements may also consist of categorical measures that represent characteristics of an event or condition of interest, such as the diagnosis of a medical event during a fixed time period. Limited measures characterized by a limiting lower value and a skewed distribution among nonlimiting values are also frequently collected in panel surveys. For example, wage and salary income would only be available from individuals with a wage earning or salary job. In this case the limited measures might also include categorical measures such as the type of employer or weekly indicators of when income was received. The type of data elements occurring in a survey, continuous, categorical, or limited, must also be considered in examining alternative wave nonresponse imputation strategies.

Missing data compensation for wave nonresponse typically will be concerned with several different types of measures at the same time. A suitable compensation strategy for continuous measures on a missing wave may not be suitable for categorical or limited measures. Compensation for a limited measure is probably the most difficult among the three types. For example, it might involve a two stage procedure. First, whether the individual had or did not have a condition or state represented by the lower limiting value may be imputed. If the individual is imputed to have the condition, the other categorical and intervally scaled measures associated with the nonlimiting values of several measures may have to be imputed. Extensive decision trees may need to be developed for even the simplest types of panel survey data structures.

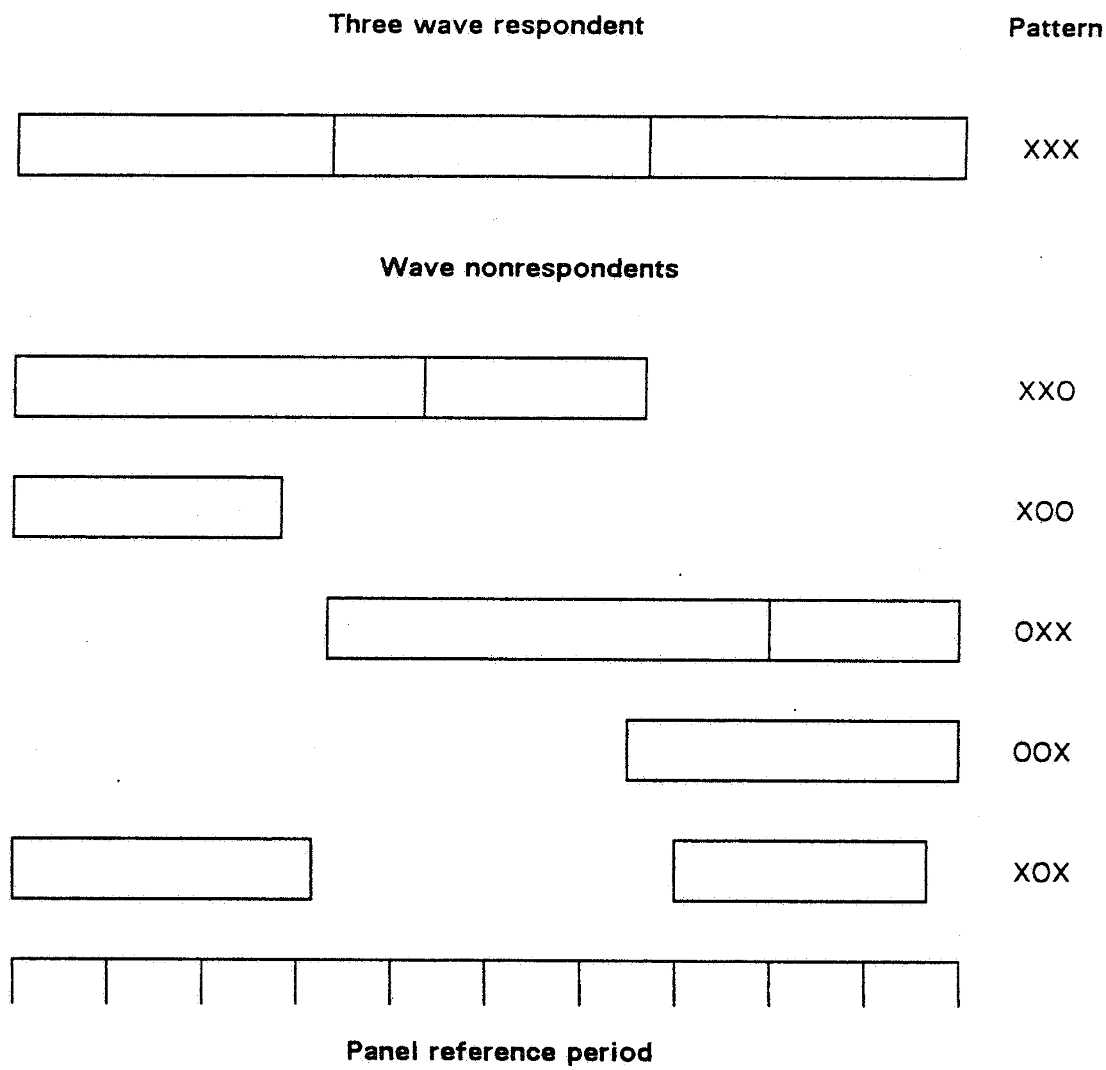
The complexity of missing data compensation for wave nonresponse is increased by the patterns of wave nonresponse that may occur. For example, the schematic in Figure 1 indicates several patterns of wave nonresponse that may occur in a three wave panel survey. Each wave of data collection may cover the same reference period, or as indicated by the boxes of varying length in Figure 1, they may cover different periods. In this case, the schematic represents the first three waves of a panel survey with more than three waves, and the reference periods for the various types of wave response patterns do not end at the same time point.

The patterns of wave response can be represented for this three wave panel as a pattern of X's (representing a wave response) and O's (representing wave nonresponse). There are $2^3 - 1 = 7$ possible wave nonresponse patterns for the three wave panel; all but the OOO pattern are represented in Figure 1.

The frequency of these wave nonresponse patterns will vary across surveys depending on the survey topic, survey organization, and a variety of other factors. Table 1 presents the frequency distribution of respondents to two panel surveys with similar topics and designs. The Income Survey Development Program 1979 Research

Figure 1

Wave Nonresponse Patterns for a Three Wave Panel Survey



Panel (ISDP) consisted of residents of approximately 9,000 households who were interviewed every three months for a total of 8 interviews concerning income and participation in government programs (Ycas and Lininger, 1981). The Survey of Income and Program Participation 1984 panel (SIPP) consisted of residents of approximately 20,000 households interviewed every 4 months for a total of 9 interviews about similar topics (Herriot snd Kasprzyk, 1984). The data in Table 1 are for the first three waves of each panel and for persons ages 16 and over in the ISDP panel and 15 and over in the SIPP panel.

Table 1

Percent Distribution of Person Response Patterns for the First Three Waves of the ISDP 1979 and SIPP 1984 Panels for Those Responding on at Least One Wave

Pattern (Response X/ Nonresponse O)	ISDP		SIPP
	All Persons	Wave 1 Respondents	
		Respondents	
XXX	80.2	83.3	90.0
		Attritors	
XXO	7.2	7.5	4.9
XOO	6.7	7.0	4.2
		Non-attritors	
XOX	2.3	2.4	1.0
OXX	2.2	–	–*
OXO	0.6	–	–
OOX	0.9	–	–

* Persons not responding at wave 1 in the SIPP were not followed for subsequent interviews.

In both surveys, the largest percentage of persons are three wave respondents. The attrition patterns, in which the respondent appears in an early wave and then fails to respond at later waves, are the next most frequent patterns. The nonattrition patterns are the least frequent, and three of these patterns do not appear in the SIPP 1984 panel since wave 1 nonrespondents were not followed at later waves. The percent distribution for ISDP wave 1 respondents is provided for comparison with the SIPP distribution, removing the nonattrition patterns OXX, OXO, and OOX.

As more waves are considered, the relative frequency of the "complete respondent" pattern decreases (usually slowly), while the number of patterns increases rapidly. For example, for a 6 wave panel survey there will be $2^6 - 1 = 63$ patterns. The complete respondent pattern (*i.e.*, XXXXXX) is likely to be the most frequent pattern, followed by the patterns with one and two missing waves. The other patterns are each likely to have small relative frequencies, but cumulatively the other patterns will not necessarily be a negligible percent of respondents. The more patterns that occur, the more difficult the task of developing a wave nonresponse adjustment.

In the subsequent discussions of compensation strategies, the number of patterns and their relative frequencies will be important in comparing different methods. Strategies which delete some or all records with wave nonresponse may delete very little data for panel surveys with only a few waves, but the precision of estimates may be greatly reduced if a large percent of records are deleted because of wave nonresponse.

Having examined several general issues in panel surveys and wave nonresponse, let us now examine two missing data compensation strategies and their application to missing data arising from wave nonresponse.

2. Weighting to Compensate for Wave Nonresponse

In a commonly used method of weighting for nonresponse, the sample is divided into a number of adjustment cells based on auxiliary information available for both respondents and nonrespondents. Within the adjustment cells, the weights, which may be the inverse of the probability of selection for the individual unit, are summed for all units and for responding units. The nonresponse adjustment is computed as the ratio of the sum of weights for all units to the sum for responding units. This ratio is applied to the weights of each of the responding units in the adjustment cell, while nonrespondents receive a weight of zero.

The use of adjustment cells serves several purposes. For one, to the extent that the auxiliary variables are correlated with the other survey variables, it is expected under a missing at random assumption that responding and nonresponding units in each cell will tend to have similar values for the missing survey items. If the response propensity tends to vary across the adjustment cells as well, the nonresponse adjustment weights will reduce the bias due to nonresponse. In addition, the effects of nonresponse adjustment are spread across many respondents in the same adjustment cell. There will be some increase in the variance of estimates due to the increased dispersion of weights, but it will generally be less than weighting methods which match nonrespondents to a single respondent and add the nonrespondent's weight to the matched respondent's.

Weighting is a global strategy that assigns a general type of adjustment for nonresponse to all the data elements in the survey. Separate weights might be considered for each data element to improve the predictive accuracy of the weights for each item. However, the complexity of analyzing a data set with multiple weights, especially for multivariable analyses, precludes the serious consideration of this approach further in this review. Nonetheless, while the global weighting strategy is a practical one for implementation, the ability of a set of auxiliary variables to serve as adequate predictors for all survey items is limited.

For wave nonresponse, the weighting strategy is more appealing. Not only are there many more auxiliary variables available, but also the information collected on other waves may be highly correlated with the information that should have been collected on the missing wave. However, more than one weight may be needed in order to satisfy multiple survey purposes.

Consider the three wave panel survey illustrated in Figure 1, and suppose that three cross-sectional estimates are to be calculated from the panel data: one each during waves 1, 2, and 3. For the wave 1 cross-sectional estimate, the responding patterns XXX, XXO, XOO, and XOX would have to be weighted to account for the nonresponding patterns OOO, OOX, OXX, and OXO. For the wave 2 cross-sectional estimate, the responding patterns are XXX, XXO, OXX, and OXO; the nonresponse adjustment weight will be different for the wave 2, and the wave 3, cross-sectional estimate than for wave 1. That is, three sets of weights are needed for cross-sectional estimation.

If longitudinal comparisons are of interest, other weights must be used. For example, comparing wave 1 and 2 requires weighting the XXX and XXO patterns to account for the other five wave nonresponse patterns and the complete nonrespondents (*i.e.*, OOO). Including the examination of three wave trends, the longitudinal comparisons require 4 sets of nonresponse weights.

For a panel survey designed to meet both cross-sectional and longitudinal comparison types of estimation, the three wave panel would require 7 sets of wave nonresponse weights. The number of weights rapidly increases with the number of waves: an 8 wave panel would require $2^8 - 1 = 255$ weights for both purposes.

If only longitudinal cumulation is of interest, yet a different weighting strategy may be employed. Each record may be viewed as data available for the portion of the combined reference period during which the respondent was eligible for the survey. When information about the respondent's period of eligibility is missing for some portion of the reference period, the nonmissing data for the respondent from periods of known eligibility may be inflated to account for the missing portion. That is, the respondent's available data is assumed to be the best predictor of the respondent's missing data. Thus, the nonmissing data for each wave nonrespondent is weighted to account for their own period of missing data. For example, the wave nonresponse patterns XXO, XOX, and OXX would receive a weight of 3/2 = 1.5 to account for the one missing wave in three. Patterns XOO, OXO, and OOX are weighted by a factor of 3/1. This longitudinal weighting strategy can be generalized from waves to other time units for which data may be recorded such as months, weeks, or days. For example, if a wave nonrespondent pattern XXO represented 3 months missing data during a one year reference period because of wave nonresponse, the nonresponse adjustment weight would be 12/9 = 1.33 (see Cox and Cohen (1985) for an illustration from the National Medical Care Utilization and Expenditure Survey.)

The number of weights needed for longitudinal analysis objectives can be reduced if fewer patterns occur. At one extreme, only the three wave respondents (*i.e.*, pattern XXX) are needed, weighting this group for the complete nonrespondents (*i.e.*, pattern OOO) and the six wave nonresponse patterns. The single weight for the XXX pattern respondents will meet all three panel survey objectives, cross-sectional, longitudinal comparison, and cumulation. To the extent that the frequency of the complete nonresponse and wave nonresponse patterns is small relative to the three wave respondents, this approach is attractive. But even for a three wave panel there may be sizeable amounts of data that are discarded with this approach. The precision of cross-sectional and longitudinal comparison estimates may be seriously reduced.

As illustrated for the ISDP and SIPP panel surveys, the majority of the wave nonresponse typically is attrition (*i.e.*, patterns XXO, XOO, and OOO for the three wave panel). With only the attrition patterns and the three wave respondents, one set of weights would be needed for each wave: one weight for each wave. Cross-sectional analysis would use the respective wave weights. Analysis using data from two or more waves would use weights from the latest wave involved in the analysis.

Little and David (1983) have referred to the attrition patterns of wave nonresponse as nested patterns, and they have proposed a method for incorporating all the auxiliary variables into the development of weights for nested patterns of wave nonresponse. The only auxiliary information available for both first wave respondents (patterns XXX, XXO, and XOO) and nonrespondents (OOO) are the sample design variables, denoted by the vector $\mathbf{z}$, such as strata, sampling unit, and characteristics of those units. As described previously, adjustment cells are created using the design variables $\mathbf{z}$, and the weight for the first wave is the inverse of the response rate within each cell. Alternatively, a wave 1 response indicator r_1 equal to 1 for wave 1 respondents and 0 otherwise can be regressed on the design variables $\mathbf{z}$ using probit or logistic regression. The wave 1 weights w_1 are the inverses of the predicted means for the wave 1 respondents given their specified values of $\mathbf{z}$.

For wave 2 both the design variables $\mathbf{z}$ and the responses obtained for the wave 1 respondents, say $\mathbf{x}_1$, are available for the wave 2 respondents (patterns XXX and XXO). A wave 2 response indicator r_2 is regressed on $\mathbf{z}$ and $\mathbf{x}_1$ to obtain weights $w_{2 \cdot 1}$ for the wave 2 respondents that compensate for the lost responses from wave 1 to 2. The weight for the wave 2 respondents is then computed as an adjustment to the wave 1 weight to compensate for the additional losses incurred at wave 2:

$w_2 = w_1 \cdot w_{2 \cdot 1}$

For a later wave, say the t *th*, the auxiliary data includes the sample design variables $\mathbf{z}$ and the responses at each previous wave for the respondents at the later wave: $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{t-1}$. The attrition compensation weight $w_{t \cdot 12 \ldots t-1}$ is computed as the inverse of the predicted mean for respondents at wave t from the regression of the response indicator r_t on $\mathbf{z}, \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{t-1}$. The t *th* wave weight is computed as $w_t = w_{t-1} \cdot w_{t \cdot 12 \ldots t-1}$.

Nested patterns of wave nonresponse can be created by eliminating nonnested patterns from the collected data, or, as implied in Table 1 for the SIPP, altering the data collection strategy to follow only those units that responded on a previous wave. The elimination of nonnested patterns may reduce sample sizes considerably for panels with a large number of waves. Thus, weighting strategies for nonnested patterns of wave nonresponse are also of interest.

Unfortunately, the simplicity of the nested situation is largely lost when nonnested patterns are used. For a three wave panel, the wave 1 weight is based on the regression of the response indicator r_1, that is 1 for respondents in patterns XXX, XXO, XOO, and XOX (*i.e.*, patterns with a wave 1 response) and 0 otherwise, on the sample design variables $\mathbf{z}$. For the wave 2 weight, two separate regressions are needed. The first regresses the wave 2 response indicator r_2 on $\mathbf{z}$ and the wave 1 auxiliary variables $\mathbf{x}_1$ for those respondents with wave 1 responses (*i.e.*, patterns XXX, XXO, XOO, and XOX). The second regresses r_2 on just the sample design variables $\mathbf{z}$ for the remaining patterns which do not have a wave 1 response (*i.e.*, OXX, OXO, OOX, and OOO). The wave 2 weight is the appropriate inverse of the predicted response probability for the wave 2 respondents.

Wave 3 weighting for nonnested patterns involves four regressions of the wave 3 response indicator r_3 on various combinations of sample design and previous wave auxiliary data x_1 and x_2:

1) r_3 regressed on z, x_1, and x_2 for patterns XXX and XXO.
2) r_3 regressed on z and x_1 for patterns XOX and XOO.
3) r_3 regressed on z and x_2 for patterns OXX and OXO.
4) r_3 regressed on z for patterns OOX and OOO.

Analyses involving more than one wave would use the product of these cross-sectional weights for waves used in the analysis. Thus, longitudinal comparisons of waves 1 and 2 would use the product $w_1 \cdot w_2$, which limits analysis to respondents in patterns XXX and XXO who have nonzero weights for both waves. Longitudinal analysis of three waves, such as a cumulation, would use the XXX pattern respondents and the product $w_1 \cdot w_2 \cdot w_3$.

The nonnested weighting scheme requires substantially more computation to develop cross-sectional weights than other schemes. The number of computations increases geometically as the number of waves increases. In addition, the nonnested weights suffer from another disadvantage. The number of respondents for a set of matched patterns may be small, while the number of nonrespondents is large. For example, for the matched patterns OOX and OOO for the wave 3 weights there are likely to be few OOX respondents but many OOO nonrespondents. The resulting weights could be quite large and variable, adversely affecting the precision of survey estimates. Some collapsing of the pattern matching will reduce this problem, but at the price of using all the auxiliary information available for some patterns. Thus, the OOX and OOO patterns might be matched with the XOX and XOO patterns for the wave 3 weight regression. But the predictor variables for this group would be limited to the sample design variables z since x_1 is not available for the OOX and OOO patterns.

Finally, the nested approach does not necessarily mean that all data for nonnested patterns must be discarded in analysis. In a three wave panel the nonnested pattern XOX could be used in the nested weighting scheme by discarding the wave 3 data to create a nested or attrition pattern for weighting. Although this is the only pattern which can be used in a three wave panel, panels with more waves will have a number of other nonnested patterns which can be converted to nested patterns by discarding later waves of data obtained from re-entrants to the panel.

The simplicity of weighting for unit nonresponse is reduced considerably for wave nonresponse by the need to handle multiple objectives. Some wave nonresponse weighting schemes may involve losses in precision of estimates due to the discarding of collected data necessary to simplify the weighting computations. However, the increased complexity of wave nonresponse weighting should be considered relative to the potentially more extensive tasks involved with imputation for wave nonresponse.

3. Imputation to Compensate for Wave Nonresponse

Imputation for missing data is a process in which values are assigned to replace the missing information, using auxiliary variables to determine the specific imputed value. In a panel survey, the imputations may be made within a wave by transferring data from a donor to a recipient. Imputation may also be made across waves within the same unit by replacing missing information on one wave with data from another wave, or information may be transferred from a donor using auxiliary information on another wave to determine the donor-recipient match.

Imputation completes a data set making analysis appear to be easier to conduct and results easier to present. For example, the form of the analysis has to be considered when deciding which weight to use from the wave nonresponse weighting. A data set completed by imputation can be used for any analytic purpose without such considerations. Imputation also assures that results from analyses that employ cross-sectional methods are consistent with those using longitudinal data, a feature not present with some of the wave nonresponse weighting schemes presented in the last section.

While imputation does have a number of attractive features, imputation as a method of compensating for missing data has major drawbacks as well. Imputation fabricates data. When there is a substantial amount of missing data, imputation gives the data set the appearance that it is complete. This appearance is misleading and can lead to a false level of confidence in the accuracy of the findings. In addition, imputed values tend to attenuate the covariance among survey items. That is, the observed covariation among the responses is biased toward zero by the presence of imputed values. For analysts concerned with examining the relationships among a number of variables, some of which may have imputed values, the amount of imputation may be an important problem to consider in determining appropriate analytic methods.

Following Kalton and Lepkowski (1983), imputation can be represented in a general way as the model $y_i = f(x_{1i}, x_{2i}, \ldots, x_{pi}) + e_i$ where y_i is the value imputed for the i th respondent, $f(\mathrm{x})$ is a function of p auxiliary variables in x, and e_i is an estimated residual. The function $f(\mathrm{x})$ is often expressed as a linear function $\beta_0 + \Sigma_j \beta_j x_{ij}$ where the β_j's are estimated from data for respondents. If the $e_i = 0$, the imputation method is a deterministic one, and the distribution of the imputed variable may be distorted and the variance attenuated. On the other hand, stochastic imputations add an estimated residual to the function $f(\mathrm{x})$, and are generally preferred to deterministic methods.

The functional form for imputation includes regression imputation through this linear expression, but it also includes other types of imputation procedures. For example, imputation class methods are represented by a set of x_j's that are indicator variables jointly defining the imputation classes. Similarly, the familiar hot-deck imputation can be viewed as a function of auxiliary variables defining imputation classes for which a mean value is to be assigned. The imputation consists of the imputation class mean plus a residual estimated from a randomly selected donor.

Kalton and Lepkowski describe a variety of methods for imputing across two waves of a panel survey. The hot-deck method can impute missing values on a wave by using the value of the missing variable on another nonmissing wave to determine imputation classes. For example, consider an income item from a survey such as SIPP that for a wave nonrespondent is known for wave 1 but missing for wave 2. The wave 1 income values can be categorized and classes formed. The missing wave 2 value for the wave nonrespondent would then be assigned the wave 2 value from a two wave respondent who comes from the same wave 1 income class.

Similarly, regression and other item imputation methods can be adapted to use auxiliary data from another wave in the imputation process. For instance, imputation of wave 2 income y_i given the individual's wave 1 income x_i can be represented by the regression model $y_i = a + bx_i + e_i$. Cross-wave regression imputation constructs a new variable $\hat{y} = a + bx_i$ for all individuals, imputing the e_i's for wave 2 nonrespondents, and calculating y_i as $\hat{y}_i + e_i$. The e_i's may be imputed by a hot-deck method, as indicated previously, or some stochastic mechanism can be used to generate them from a known distribution.

The constant and slope terms, a and b, can be obtained from least squares estimates for the regression of y_i on x_i for two wave respondents. Alternatively, assigning $a = 0$ and $b = 1$ without a residual term is a "carry-over" imputation, perhaps the simplest cross-wave imputation procedure to implement. Setting $a = 0$ and estimating b from the respondent data is a model of proportionate change; setting $b = 0$ and estimating a from the respondent data is a model of additive change. The quality of these cross-wave imputations depends on the strength of the correlation for the same item over time.

Many panel survey items are repeated on each panel wave, and the responses are highly correlated over time. Responses for an item on one wave will then be powerful auxiliary variables for imputing the missing response for the same item on another wave. Kalton, Lepkowski, and Lin (1985) have found strong cross-wave consistency and correlation for labor force and income items across the first three waves of the ISDP 1979 Research Panel. Table 2 illustrates the consistency of several quarterly categorical labor force and income recipiency items across the first two waves of the panel as presented by Kalton, Lepkowski, and Lin. (With the rotating panel design of the ISDP panel, the first three waves cover a nine month period for which three quarters of labor force data and nine months of income data were collected.) A consistent two wave response is one in which the response is identical for both waves. For receipt of Social Security income during the three month reference period, 18.3 percent of the 13,151 original sample respondents ages 16 and older with data on both waves were reported to receive income both waves, and 80.3 percent on neither wave. A total of 98.7 percent of the responses are consistent for the first two waves of the panel.

Table 2

Percent Distribution of Responses Across Waves 1 and 2 of the ISDP 1979 Research Panel for Original Area Frame Sample Respondents Ages 16 and Older

Item	Wave 1/Wave 2 Response				Consis- tency	Sample Size
	Yes/ Yes	Yes/ No	No/ Yes	No/ No		
Received Social Security Income	18.4	0.4	0.9	80.3	98.7	13,151
Received Federal SSI	3.2	0.3	0.3	96.2	99.5	13,151
Worked in Quarter	58.2	3.5	3.8	34.5	92.8	13,119
Reasons for not Working						
Going to School	11.0	0.9	0.7	87.4	98.4	4,520
Didn't Want to work	4.9	6.5	8.5	80.1	84.9	4,520
Retired	15.3	5.0	6.5	73.2	88.5	4,520

With consistencies indicated in Table 2, a simple carry-over imputation would assign the correct value a large proportion of the time. (The direction of the carry-over imputation could be forward or backward depending on which waves are missing.) If wave nonrespondents have the same cross-wave consistency as the wave respondents demonstrated in Table 2, the carry-over imputation is likely to assign a high proportion of missing wave income recipiency and labor force items correctly. However, the deterministic carry-over imputation has two basic problems which limit its usefulness. For one, the distribution of the carry-over imputations may differ from the nonmissing responses on the wave. The percent of "Yes" responses on wave 1 for Social Security income is 18.4 + 0.4 = 18.8, which would be the carry-over imputed percentage as well. But the actual wave 2 percent "Yes" is 18.4 + 0.9 = 19.3. In addition, the carry-over imputation forces stability of responses for wave nonresponses, attenuating the cross-wave changes in the data.

A stochastic cross-wave imputation method can avoid these deficiencies. For example, instead of carrying over the response, a stochastic mechanism can be employed to assign 18.4 / (18.4 + 0.4) = 97.9 percent of the wave 1 Social Security income recipients a "Yes" response at wave 2, and 2.1 percent a wave 2 "No" response. Similarly, 99.5 percent of the wave nonrespondents with data on wave 1 but not wave 2 with a "No" response on wave 1 would be assigned a "No" response on wave 2. The wave 2 stochastically imputed responses will have in expectation the same distribution on wave 2 as the wave 2 nonmissing responses, and they will have changes in Social Security income recipiency across the waves.

Kalton, Lepkowski, and Lin also present cross-wave correlations from the ISDP 1979 panel for several continuous income amount items, three of which are presented in Table 3. The 8 × 8 correlation matrix for these items has been summarized by computing average correlations when the income items were one, two, or three months apart. In addition, one and two month differences could either be between two monthly responses from the same wave or between different waves.

The average cross-month correlations are quite high for the earnings and Social Security items, and somewhat smaller for unemployment compensation. The

Table 3

Average Cross-month Correlations for Three Amount Items for ISDP 1979 Research Panel Original Area Frame Sample Respondents Ages 16 and Older

Item	One Month Difference			Two Month Difference			Three or More Month Difference
	Within Wave	Between Wave	Within and Between	Within Wave	Between Wave	Within and Between	
Wage and Salary Earnings	0.933	0.842	0.910	0.890	0.839	0.861	0.813
Social Security Income	0.983	0.921	0.968	0.978	0.924	0.946	0.903
Unemployment Compensation	0.651	0.408	0.590	0.645	0.448	0.532	0.599

correlations within a wave are higher than those between waves, presumably because of response errors caused by such survey procedures as the use of a proxy report on one wave and a self report on another. As the number of months between values increases for the earnings and income items, the correlations decrease. However, for unemployment compensation, there is a decrease followed by an increase in correlation. Unemployment compensation has lower cross-month correlations than the income items to begin with. Further, it appears that there is substantial short term change in the recipiency of unemployment compensation, but once the compensation has been received for a longer period, the amount of compensation begins to stabilize.

The data in Tables 2 and 3 primarily address two waves with wave nonresponse on one of them. Imputation could be made forward or backward, depending on which wave is missing. If both waves are wave nonresponses, the cross-wave imputation procedures are not applicable, and a back-up within wave imputation strategy may be needed to complete the imputations for wave nonresponse. This situation may arise in a panel with more than two waves of data when the processing is conducted wave by wave. Opportunities for imputation across several waves may not be available in such a processing environment.

For example, in a three wave panel, patterns such as XOO and OOX may appear to be inapplicable for cross-wave imputation for a given pair of waves, when in fact the imputations can be made across several waves. Thus, missing waves 2 and 3 for the XOO pattern can be forecast from wave 1, while missing waves 1 and 2 for the OOX pattern can be backcast from wave 3. Of course, the longer the period across which the imputations must be made, the lower the quality of the imputations that may be expected.

The other three wave patterns present other alternatives for cross-wave imputations as well. The XXO and OXX patterns could use data from the nearest wave to forecast or backcast, respectively, for the missing wave. Given the decreasing correlations with increasing time, data from the other nonmissing wave are unlikely to improve the explanatory power of the imputation regression model further. Data can be both forecast and backcast for the OXO pattern, while the XOX pattern can provide either a forecast or a backcast for the missing wave.

Of course, the complete nonrespondent pattern OOO cannot be imputed by cross-wave methods. These nonrespondents can be handled as unit nonresponses and weighted in the usual way, or some within wave imputation procedure such as a hot deck can be employed. Weighting is easier and probably nearly as accurate a method as imputation for complete nonrespondents.

Not all panel survey items will be highly positively correlated over time, and hence would not be good candidates for cross-wave imputation. Health care utilization and expenditure items collected over a one year period are not likely to be highly correlated over time for most persons. The National Medical Care Utilization and Expenditure Survey (NMCUES) was designed to collect such data from a panel of approximately 17,000 persons interviewed four or five times during a one year period (Bonham, 1983). The primary purpose for the panel design employed in NMCUES was to improve the quality of data collected by decreasing the length of the reference periods through repeated interviews with the panel over a one year period. Cross-wave imputation was not considered appropriate to compensate for wave nonresponse because of the low correlations over time for the principle survey items (Cox and Cohen, 1985).

A few items may actually be negatively correlated over time. For example, expenditures for an automobile may be negatively correlated over short time periods, since once such a major purchase of a durable good has been made, the expenditure will not be made again for some time. Similarly, events such as births will be negatively correlated over limited time periods. These negative correlations may be useful for developing cross-wave imputation methods by indicating when an expenditure or event will not occur in a preceding or succeeding wave.

This discussion, and the functional imputation model described previously in the section, has been concerned with the item imputation problem in which one item is imputed at a time. Wave nonresponse presents a set of missing items which could be compensated simultaneously through the imputation of an entire wave of data. The functional form can be generalized in an obvious way by considering the imputed value to be a vector of values $\dot{y}_i$. Item imputation methods can be generalized to wave imputation methods in a straightforward way.

For example, consider imputation of a class mean for missing items. The wave imputation alternative would assign a vector of class means for the missing items on an entire wave. Similarly, hot deck wave imputation matches a donor to a recipient using auxiliary information from the sample design or another wave and imputes the donor's data for the missing wave to the recipient. Even regression imputation in its various forms can be generalized through multivariate regression methods to handle wave imputation. The simplest form of wave regression imputation is the wave imputation of the version of the carry-over method: all data items that are repeated on each wave are carried over to the missing wave. The method can be adjusted for known changes that might occur from one wave to the next, such as a change in the length of the reference period.

An important advantage for wave imputation compared to repeated item imputations for wave missing data is, in some forms of wave imputation, the ability to handle several types of measures simultaneously. For instance, the same form of regression imputation is not suitable for categorical items and for continuous items, while two different continuous items may require different regression imputation models. Limited measures require a separate imputation for assigning the limiting value and another for assigning nonlimiting values if the limiting value was not assigned in the first imputation. Similarly, hot-deck imputation for each item on a wave would also be cumbersome unless several related items are imputed at the same time. Such imputation must be designed almost on an item by item basis, an approach that may not be appropriate for one time (as opposed to rotating) panels with complex interviews.

4. Weighting or Imputation?

Weighting and imputation have been described as two separate methods for compensating for missing data, but they are in fact closely related when analysis of a single item is considered (Kalton 1983; Oh and Scheuren 1984). The relationship between the two approaches is illustrated by considering a simple hot-deck imputation procedure. A sample of respondents and nonrespondents is divided into imputation classes using auxiliary variables available for both, and a nonrespondent within the class is assigned the value for the missing item from a respondent in the same cell. For analysis of this single item, this imputation scheme is equivalent to a weighting scheme in which the weight of the nonrespondent is added to the weight of the matched respondent who in the hot-deck imputation donated the imputed value. The mean and variance of the item being imputed are the same under either missing data compensation scheme.

Weighting does not typically add nonrespondent weights to individual respondent weights in order to avoid large increases in the precision of estimates due to increased variability in the weights. Rather, the respondent weights within the class are all increased proportionately spreading the adjustment across class members. Imputation can thus introduce larger increases in the variance of estimates than weighting adjustments made within classes, although this increase can be reduced somewhat by appropriate selection of donors (Kalton and Kish 1984) or multiple imputations (Rubin 1979).

Weighting has several important advantages over imputation. It can be applied as a global strategy to all variables simultaneously, thus making it a practical method for missing data compensation. Individual item imputation of many items can be an expensive and time consuming operation. For practical reasons, full advantage of high correlations between survey items over time is not taken advantage of. Thus, separate hot-deck imputations are not performed for each item requiring imputation because it would be prohibitively expensive. Rather, blocks of similar variables are imputed on the same hot-deck run to reduce the cost of imputation.

Weighting also has the advantage of preserving the observed realtionships among survey variables, provided large amounts of data are not eliminated in the weighting process to simplify weight development and use. Increasing the weight of a respondent record effectively reproduces the observed relationships for the proportionate increase in the weight. Imputation manufactures data that can attenuate the covariances among survey variables (Santos 1981; Kalton and Kasprzyk 1982). The manufactured data can even be inconsistent with other variables on the record. Much survey analysis is concerned with relationships among variables examined through cross-tabulations and regression analysis. Imputation can seriously attenuate the strength of observed covariances, and thus reduces the ability to detect important relationships in the data (Lepkowski, Stehouwer and Landis, 1984, provide an example). This attenuation extends to the effects item imputation can have on measures of change over time. Heeringa and Lepkowski (1986) report on a comparison of a simple cross-wave carry-over imputation to a within wave hot-deck imputation that between wave change is seriously distorted by both procedures.

Imputation has an important disadvantage: it fabricates data. The completed nature of an imputed data set is a seductive feature of imputation which can lead to inappropriate analysis. Analysts will tend to treat the imputed values as real values in variance estimation, for instance, and will have greater confidence in survey results than may be warranted because the precision of survey estimates has been overstated by the presence of imputed values.

Of course, weighting is not without its disadvantages. The number of different weights needed for multiple analytic objectives in a panel survey can be sizeable, especially as the number of wave nonresponse patterns increases. Reducing the number of patterns can be accomplished through elimination of data or alteration of data collection procedures to avoid following panel units that fail to respond for a single wave. But both of these alternatives are unattractive. Weighting also does not use the full strength of a correlation between an auxiliary variable and an item with missing data, because the auxiliary variable is often categorized in the formation of weighting classes.

Kalton and Miller (1986) have investigated empirically the quality of alternative imputation and weighting strategies for compensating for wave nonresponse. They simulated missing waves of data among three wave respondents to the first three waves of the SIPP 1984 panel, and then made imputation and weighting compensations that could be compared to the actual values that were simulated to be unknown.

A SEARCH analysis (Sonquist, Baker, and Morgan 1973) was conducted to identify predictors for four wave nonresponse patterns in the data, XXX, XXO, XOO, and XOX. The SEARCH analysis identified a detailed and complex prediction model for the wave nonresponse patterns which defined a set of 41 subgroups with response rates ranging from 61.6 to 98.6 percent XXX respondents. A simulation data set was then formed from the three wave respondents which had the same distribution across the 41 groups by randomly sampling 61.6 percent of the three wave respondents in each group. The resulting sample of 18,481 three wave respondents were then randomly assigned a simulated wave nonresponse pattern within each of the 41 SEARCH groups that corresponded to the total sample distribution of wave nonresponse patterns for that group.

Two methods for compensating for wave nonresponse were then applied to the simulation data set. First, a simple carry-over imputation was used to complete items on missing waves. Second, weights were assigned to 16,635 simulated three wave nonrespondents to compensate for missing data from 1,846 simulated wave nonrespondents (*i.e.*, wave nonresponse patterns XXO, XOO, and XOX). Weighting classes were formed using the survey variables age, sex, household income, race, education level, whether receiving certain types of welfare, whether in the labor force, and whether unemployed. The classes were collapsed until all classes contained a minimum of 20 three-wave respondents. The wave nonresponse adjustment weights ranged from 1.0 to 1.5.

Several analyses were conducted to compare the quality of estimates computed using the imputed data and using the weights and three wave nonrespondents. Comparison of overall survey estimates would not be particularly sensitive to the effects of imputation or weighting for the 10 percent of the sample which had wave nonresponses on at least one wave. The imputed data can be compared directly to the actual data for the wave nonrespondents alone to assess the quality of the imputations. Since the weighted data does not contain the wave nonrespondents, a direct comparison of weighted and actual estimates for wave nonrespondents is not possible. However, noting that the increases in weights for three wave nonresponses reflect the adjusment to the three wave respondents for missing data due to the 1,846 wave nonrespondents, the factor $(w_i - 1)$, where w_i is the weight for the *i th* three wave respondent, is the increase in weight assigned to compensate for wave nonresponse. Thus, weighted estimates using the weighting factor $(w - 1)$ are compared to estimates computed using the actual and imputed data for wave nonrespondents.

Table 4 presents the distributions across the three waves of "Yes" responses for two survey items for the actual and imputed data for wave nonrespondents and for three wave respondents weighted by the added weight assigned to adjust for wave nonresponse. The YNY and NYN patterns for both items do not have any imputed values because with the carry-over imputation and the wave nonresponse patterns occurring in the data these patterns cannot occur. In addition, the patterns YYN and NNY occur rarely since with the carry-over imputation they can only occur in the pattern XOX, an infrequent wave nonresponse pattern.

As a result of the poor representation of these four patterns in the imputed data, the joint distribution for wave nonrespondents with imputed data does not correspond to the actual distribution well. The patterns which represent change (*i.e.*, YYN, YNY, YNN, NYY, NYN, and NNY) are only 4.9 and 5.1 percent of the patterns in the imputed

Table 4

Percent Distribution of Responses Across Waves for Two Items for Wave Nonrespondents and for Weighting Adjustment for Wave Nonrespondents

Response Pattern	Wave Nonrespondents: Actual	Wave Nonrespondents: Imputed	Weighted Adjustment
		Having a Job	
YYY	58.1	63.3	57.4
YYN	2.4	0.4	2.4
YNY	2.5	–	2.5
YNN	3.2	2.6	3.1
NYY	2.5	1.5	2.6
NYN	0.7	–	0.7
NNY	2.7	0.4	2.7
NNN	27.8	31.8	28.6
		Having Savings Accounts	
YYY	45.1	49.9	48.9
YYN	2.4	0.7	2.7
YNY	1.2	–	1.2
YNN	4.4	2.4	3.3
NYY	2.8	1.3	2.7
NYN	0.2	–	0.4
NNY	2.4	0.8	2.3
NNN	41.5	44.9	38.5
Number of persons	1846	1846	16,635

data, compared to 14.0 and 13.4 percent for the actual data. The carry-over imputations are forcing more cross-wave stability in the data than is correct.

In contrast, the weighted distribution agrees fairly closely with the actual distribution for both items. For "Having a Job" the weighted distribution overestimates the NNN pattern and underestimates the YYY pattern somewhat, while for "Having Savings Accounts" there is a more serious underestimation of the NNN pattern and overestimation of the YYY pattern. Nonetheless, the weighting adjustment estimates virtually the same percent of change patterns as the actual distribution, 14.0 and 13.6 percent for the two items, respectively.

The deficit of change patterns for the carry-over imputations is even more apparent when gross change is examined. From wave 1 to 2, actual gross change is 8.9 percent (*i.e.*, patterns YNY, YNN, NYY, and NYN), but the carry-over imputations have only 4.1 percent. But the gross change from wave 1 to 2 for carry-over imputations is composed entirely of the YNN and NYY patterns, most of which is change from a pattern with no missing data on waves 1 and 2 (*i.e.*, XXO). At the same time, gross change from wave 2 to 3 is 8.2 percent, but the carry-over imputation estimates only 0.8 percent wave 2 to 3 gross change, attributable solely to the XOX pattern.

Despite the attenuation of change for these two items by the carry-over imputations, the percent of "Yes" responses for each wave for the carry-over imputations agrees fairly closely with the actual distribution. Table 5 presents the marginal distributions of "Yes" responses for each wave for both items. The percent "Yes" for the first wave carry-over imputations must be identical to the actual percent since there is no wave nonresponse for wave 1 in the simulation data set. The weighting adjustment does depart from the actual percent since the responses for wave nonrespondents on waves two and three are deleted from the data set and must be compensated for.

Although the marginal wave distributions do not demonstrate the dramatic attenuation of change shown for the joint distributions for the carry-over imputation, the effects are still apparent in the change from first to second wave percentages. The actual percentages decrease 2.5 and 2.6 percent from wave 1 to 2 for the two items, respectively. The carry-over imputations only decrease 1.0 and 1.2 percent, respectively, reflecting a dampening of the actual cross-wave change. In contrast, the weighting adjustments decrease 2.3 and 1.4 percent, respectively. Although still underestimates of wave 1 to 2 change, the weighting adjustments do not distort the distribution of change as much as the carry-over imputations.

Kalton and Miller also compared carry-over and weighting adjustments for a continuous item, Social Security income. Social Security income was reported on a

Table 5

Marginal Distribution of "Yes" Responses for Each Wave for Two Items for Wave Nonrespondents and for Weighting Adjustment for Wave Nonrespondents

	Wave Nonrespondents		Weighted Adjustment
	Actual	Imputed	
		Having a Job	
Wave 1	66.2	66.2	65.4
Wave 2	63.7	65.2	63.1
Wave 3	65.8	65.2	63.1
		Having Savings Accounts	
Wave 1	53.1	53.1	56.1
Wave 2	50.5	51.9	54.7
Wave 3	51.5	52.0	55.1

monthly basis, four months per wave, totaling 12 monthly reports for the three panel waves. The carry-over imputation assigned the amount from the latest available month for each missing month. In actual cross-wave imputation, the recipiency of Social Security income would have to be imputed first, followed by the imputation of the amount (*i.e.*, a limited measure requiring a two stage imputation). In the simulated data set, if recipiency were imputed to a person who did not receive Social Security, there would be no actual value to compare an imputed income to. Direct comparison of actual and imputed values cannot then be made, although a comparison of monthly means can be made.

Since one part of their investigation involved direct comparisons of actual and imputed values, Kalton and Miller imputed across waves only when the simulated wave nonresponse was known to be a recipient in the missing wave. Thus, their results concern recipients of Social Security income in the nonmissing and the missing month. In addition, they excluded records which had received Bureau of the Census hot-deck imputations for Social Security income, and months with extremely large Social Security incomes or large changes from month to month were also excluded.

Table 6 presents monthly mean Social Security income for the 12 months of the three wave reference period. As before, the first wave did not have missing data, and the imputed and actual mean monthly incomes are identical. The weighting adjustment estimates tend to be smaller than the actual means for the first wave.

Table 6

Mean Monthly Social Security Income for Wave Nonrespondents Receiving Social Security Income and for Weighting Adjustment for Wave Nonrespondents

Month	Wave Nonrespondents			Weighted Adjustment
	Actual	Imputed	Adjusted Imputed*	
1	388	388	388	386
2	395	395	395	386
3	389	389	389	385
4	387	387	387	386
5	381	382	382	389
6	383	382	384	390
7	387	386	390	394
8	390	387	393	398
9	400	391	396	399
10	395	391	396	400
11	398	391	396	399
12	399	391	396	401

*Carry-over imputed values adjusted for a January, 1984, cost of living increase of 3.5 percent.

In the second and third waves (*i.e.,* months 5 to 12), the monthly means for the carry-over imputations consistently underestimate the actual means. In January, 1984, Social Security recipients received a 3.5 percent cost of living increase. Carry-over imputations from a month before January to January or a later month will consistently underestimate the monthly Social Security income because they will not properly account for this increase. To compensate for this known increase, and to attempt to bring the means for carry-over imputations closer to the actual means, all carry-over imputations from a month before January, 1984 to January or later were increased by 3.5 percent.

The monthly means computed after this adjustment are given in the "Adjusted Imputed" column in Table 6. The underestimation of the carry-over imputation is now corrected for the second wave, although the adjusted imputation means still tend to be smaller than the actual means for the third wave. Kalton and Miller indicate that this deficit is related to the restrictions imposed by imputing only to persons known to be Social Security recipients on the missing wave. In addition, in the second wave almost one-half of the wave nonrespondents still have an actual value for the four months in the wave, while the proportion is lower for the third month. Thus, there is more carry-over imputation for the third than for the second wave.

Kalton and Miller note that it is difficult to draw general conclusions from such limited investigations. Nonetheless, they do find the quality of the weighting adjustment for wave nonresponse to be comparable to, if not better than, the simple carry-over imputation. The weighting adjustment preserves the cross-wave relationships observed in the actual data, and although the means and percents do not agree with the actual means and percents, the departure is not large. On the other hand, the weighting adjustment does discard 10 percent of the records decreasing sample sizes and the precision of weighted estimates. The loss of precision is small in this instance with only three waves and 6 wave nonresponse patterns discarded. However, if the full 8 or 9 waves of the SIPP 1984 panel had been available for their investigation, the amount of data discarded would have been larger. It is unlikely that weighting complete wave respondents would be a preferred strategy with larger amounts of discarded data.

5. Weighting and Imputation

Given the at times complementary strengths and weaknesses of the weighting and imputation strategies, it is natural to consider whether combinations of the two approaches would be suitable. For example, imputation could be used to compensate for some wave nonresponse patterns, and weighting could be used to compensate for the others. There are many possible combinations, the choice among them depending on such characteristics as the panel design, analytic purposes, and types of data collected.

For example, for the three wave panel design, imputation could be used to complete the wave nonresponses for those missing only one wave (*i.e.*, patterns XXO, XOX, and OXX). The remaining patterns would be deleted, and a single weight developed to compensate for them. Thus, records with two waves of data are retained for analysis, and imputation is used for at most one missing wave on a record. Cross-wave imputations would have to be made across only one wave, using the higher correlations of data closer in time. For the ISDP 1979 panel, 11.7 percent of the persons responded on at least two of the first three waves, and they would receive imputation for the missing wave's data. Another 7.9 percent had two waves of missing data, and they would be compensated (together with the unit nonrespondents) by weighting. If the imputations do not involve cross-wave imputation for the same unit, imputations could be done with or without the deleted records. If a hot-deck form of imputation is being used, it seems preferable to retain the deleted records in order to increase the number of donors available for matching.

When there are more than three waves, or the panel waves vary in length, the decision about which patterns should be imputed and which deleted is more difficult. For example, in a six wave panel patterns with one or perhaps two missing waves are candidates for imputation, and four or five missing wave patterns are perhaps most appropriately weighted. But the three wave patterns might be handled either way. The wave pattern XOXOXO might be a good candidate for imputation, while the pattern XXOOOX might be better compensated by weighting.

Variable length reference periods and variation in the number of scheduled interviews that a panel member is scheduled to receive will influence the amount of missing data present due to wave nonresponse as well. For example, because of field operations scheduling, some panel respondents have only three months between interviews while others have four or five. In such a case, the decision about whether to handle wave nonresponse with imputation or weighting might be based on the proportion of the cumulative reference period that is missing. Records with missing data for more than a fixed fraction (*e.g.*, one-half, one-third) of the reference period would be handled by weighting, and the others by imputation.

To reduce the amount of data fabrication, imputation could be used to convert some nonnested patterns into nested patterns. The cross-section and longitudinal weighting suggested by Little and David could then be used to compensate for the remaining wave nonresponse. For example, for the three wave panel, the nonnested patterns OXX and XOX could be completed by imputation, the nested patterns XXO and XOO remain nonimputed, and the nonnested patterns OXO and OOX deleted. This combination retains more of the data than does the imputation to complete records, since the pattern XOO that was deleted for the previous combination is retained. In addition, less data is fabricated under this combination than under the previous one since the XXO pattern, which is likely to be a common wave nonresponse pattern, is not imputed. On the other hand, this combination of completing nested patterns and then weighting requires more complicated weighting than the previous combination. Completing nested patterns and weighting is therefore less attractive than the previous combination which yields a single weight.

Another type of combination is to use imputation for those data which can be imputed well (*e.g.*, the same item across waves, highly correlated across time), and use weighting for other items (*e.g.*, topical items asked only one time). However, presumably there will be analytic interest in examining relationships between items asked on every wave and topical items. It is not clear how one should handle the analysis of items which have different weights within the same record.

Finally, the need for different weights for different analytic needs can be avoided to some extent by providing data sets suitable for different types of analysis. For example, the SIPP is producing individual data sets for each wave with weights suitable for conducting cross-sectional analysis. It is also producing a longitudinal file composed of complete 8 or 9 wave respondents with a weight compensating for the wave nonrespondents deleted from the longitudinal file. Cross-sectional results obtained from the longitudinal file will not be consistent with those from the cross-sectional data sets, nor will they be comparable to longitudinal analyses based on linked cross-sectional data sets. This particular approach illustrates that the treatment of wave nonresponse does not need to be limited to the production of a single data set to be used for multiple purposes.

6. Selecting a Wave Nonresponse Compensation Strategy

The complexity of panel survey design and analysis precludes the recommendation of a single strategy for compensating missing data from wave nonresponse. The specific strategy must be developed with a consideration of such factors as the major survey design objectives, the panel design, the distribution of patterns of wave nonresponse, and the survey data collection organization's capabilities. For the purpose of guiding the choice of a wave nonresponse adjustment strategy, the relative strengths of adjustment strategies described in previous sections are illustrated by comparing their ability to meet several criteria. The comparison is summarized in Figure 2.

Three weighting, three imputation, and two combined strategies are examined. In *complete wave nonrespondents* weighting, the units that responded on all panel waves are weighted to account for all wave nonrespondents. *Nested pattern* weighting uses only the nested wave nonresponse patterns to compensate for all other units, while *nonnested pattern* weighting matches various patterns of wave nonresponse prior to weighting. The nested and nonnested pattern weighting leads to multiple weights, one of which must be chosen as most appropriate for each analysis.

The simplest imputation strategy is to *carry-over* data from a responding wave to a nonresponding wave for the same unit. The *cross-wave hot deck* matches donor and recipient using data from a responding wave, and then transfers donor data on the nonresponding wave to the recipient. Both of these procedures can impute entire

Figure 2

Relative Strengths of Wave Nonresponse Compensation Strategies for Five General Criteria

Compensation Strategy	Practicality	Flexibility	Accuracy	Precision	Preservation of Relationships
<u>Weighting</u>					
Complete Wave Nonrespondents	++	++	+	– –	+
Nested Patterns	o	+	+	o	+
Nonnested Patterns	o	+	+	+	++
<u>Wave Imputation</u>					
Carry–over	+	o	+/–	+	–
Cross–wave Hot Deck	+	+	o	+	– –
<u>Item Imputation</u>					
Regression	– –	– –	++/–	+	–
<u>Combined Strategies</u>					
Imputation to Complete Wave Nonrespondents	o	o	o	o	+
Imputation to Complete Nested Patterns	o	o	o	+	+

waves of data at one time. The *regression* imputation uses a regression model with such auxiliary data as the response to the same item on another wave to predict the value of a missing item. The regression imputation procedure must be done one item at a time.

Imputation to complete wave nonresponse patterns can be combined with weighting for patterns which are not completed by imputation as another strategy. An alternative strategy begins with *imputation to complete nested patterns*, which reduces the amount of fabricated data but introduces the need for multiple weights with a nested weighting strategy to compensate for nonnested patterns that are deleted.

Many criteria could be used to compare these various criteria, but five broad ones are shown in Figure 2. *Practicality* refers to the ease of implementation of the strategy. There are no widely available general purpose computer programs to implement the strategies described here. For a continuing survey operation, special purpose software can be developed which can be used repeatedly across multiple surveys. For a one-time panel survey, extensive software development or analysis to implement one of these approaches is infeasible.

The *complete wave nonrespondents* weighting is likely to be the easiest to implement and use for one-time survey operations, while the other weighting strategies will be more difficult to implement and use. *Nested and nonnested pattern* weighting have a series of large probit or logistic regression models to estimate, plus a choice of weights each time an analysis is made. The wave imputation strategies may be more or less difficult to implement than the weighting strategies depending on how much of each wave is a repetition of items from a previous wave. The regression imputation strategies will be the most difficult of all the strategies to implement because of the need to develop regression models. On the other hand, all imputation strategies will be easier to use than nested and nonnested weighting strategies since at most only one set of weights need be considered in analysis. The combined strategy beginning with *imputation to complete nonrespondents* has the analytic advantage of only one set of weights. However, both combined strategies require considerable effort to implement since both imputation and weighting must be done.

The *flexibility* criterion refers to the ability of the procedure to handle multiple data types in the data record. Clearly, the global weighting strategies are the most flexible procedures. Wave imputation procedures can handle multiple data types easily as well, although to the extent that there are items on a wave not appearing on another wave, wave imputation will not cover all data items, and other imputations must be made. Regression imputation is clearly at a disadvantage when there are multiple data types. The combined strategies are adequate provided that some type of wave imputation can be used prior to weighting.

Accuracy refers to the ability of the compensation procedure to correctly predict the missing value. Kalton and Miller and others (Cox and Cohen 1985; Ghangurde and Mulvihill 1980) have found little difference in the accuracy of weighting and imputation for cross-sectional estimation, but some forms of imputation are inaccurate for longitudinal purposes. Regression imputation can be highly accurate provided the model is correctly specified. The combined strategies should be no more accurate than the weighting strategy overall.

The *precision* of estimates depends on the sample size available. Weighting strategies which delete records will necessarily produce less precise estimates than the other approaches. The more records deleted, the less precise the estimates. The combined strategies will be less affected by deletion of wave nonrespondent records since some wave nonresponses have been completed by imputation.

The final criteria in Figure 2, *preservation of relationships*, is clearly a feature of weighting strategies. The nonnested pattern weighting preserves more of the relationships that the other two strategies since more of the data are retained under this strategy than the other two. The combined strategies will also be strong on this criteria since the imputation that is made will be limited and unlikely to attentuate the strength of relationships in major ways.

On the whole, the weighting strategies appear, on these limited criteria, to be preferable procedures for compensating for wave nonresponse than the others. They are weakest on the precision criteria but stronger on the other criteria. The combined stategies retain some of the strengths of the weighting strategies, but generally appear to be an intermediate choice between weighting and imputation.

References

Bonham, G. (1983), "Procedures and Questionnaires of the National Medical Care Utilization and Expenditure Survey," *National Medical Care Utilization and Expenditure Survey*, Series A, Methodology Report No. 1, DHHS Publication No. 83-20001, Public Health Service, Washington: U. S. Government Printing Office, March, 1983.

Cox, B., and Cohen, S. (1985), *Methodogical Issues for Health Care Surveys*, New York: Marcel Dekker.

Duncan, G., and Kalton, G. (1985), "Issues of Design and Analysis of Surveys Across Time," *Bulletin of the International Statistical Institute*, **51**(1), 14.1, 1-16.

Heeringa, S., and Lepkowski, J. (1986), "Longitudinal Imputation for the SIPP," *Proceedings of the Survey Research Methods Section, American Statistical Association*, (forthcoming).

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

Kalton, G. (1986), "Handling Wave Nonresponse in Panel Surveys," *Journal of Official Statistics*, **3**, 303-314.

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

Kalton, G., and Kish, L. (1984), "Some Efficient Random Imputation Methods," *Communications in Statistics, Theory and Methods*, 13, 1919-1939.

Kalton, G., and Lepkowski, J. (1983), "Cross-wave Item Imputation," in *Lessons of the Income Survey Development Program (ISDP)*, M. David, ed., pp. 171-198, New York: Social Science Research Council.

Kalton, G., Lepkowski, J., and Lin, T. (1985), "Compensating for Wave Nonresponse in the 1979 ISDP Research Panel," *Proceedings of the Survey Research Methods Section, American Statistical Association*, pp. 372-377.

Kalton, G., and Miller, M. (1986), "Effects of Adjustments for Wave Nonresponse on Panel Survey Estimates," *Proceedings of the Survey Research Methods Section, American Statistical Association*, (forthcoming).

Lepkowski, J., Stehouwer, S., and Landis, J. (1984), "Strategies for the Analysis of Imputed Data in a Sample Survey," *Proceedings of the Survey Research Methods Section, American Statistical Association*, pp.622-627.

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

Little, R., and David, M. (1983), "Weighting Adjustments for Non-Response in Panel Surveys," U.S. Bureau of the Census working paper.

Mulvihill, J., and Lawes, M. (1980), "Imputation Procedures for LFS Longitudinal Files," Statistics Canada internal memorandum.

Oh, H., and Scheuren, F. (1983), "Weighting Adjustment for Unit Nonresponse," in *Incomplete Data in Sample Surveys, Volume 2, Theory and Bibliographies*, W. Madow, I. Olkin, and D. Rubin, eds., pp. 143–184, New York: Academic Press.

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

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

Sonquist, J., Baker, E., and Morgan, J. (1973), *Searching for Structure*, Ann Arbor: Institute for Social Research.

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