ALAN C. ACOCK *Oregon State University*

Working With Missing Values

*Less than optimum strategies for missing values can produce biased estimates, distorted statistical power, and invalid conclusions. After reviewing traditional approaches (listwise, pairwise, and mean substitution), selected alternatives are covered including single imputation, multiple imputation, and full information maximum likelihood estimation. The effects of missing values are illustrated for a linear model, and a series of recommendations is provided. When missing values cannot be avoided, multiple imputation and full information methods offer substantial improvements over traditional approaches. Selected results using SPSS, NORM, Stata (mvis/micombine), and M*plus *are included as is a table of available software and an appendix with examples of programs for Stata and M*plus.

Traditional approaches for working with missing values can lead to biased estimates and may either reduce or exaggerate statistical power. Each of these distortions can lead to invalid conclusions. Missing values are endemic across the social sciences (Juster & Smith, 1998), and family studies is no exception. King, Hopnaker, Joseph, and Scheve (2001) found that about 50% of the participants in political survey data have missing values, and family research often approximates this level of missing values. Many of the major data sets that are utilized in articles appearing in family journals have serious problems with missing values. This is true even for large public use data sets such as the National Survey of Families and Households, the National Longitudinal Survey of Youth, the General Social Survey, the Panel Study of Income Dynamics, and the Survey of Income and Program Participants. I address several questions in this article: Why are missing values a concern? How good are traditional approaches? Is it better to drop cases than to "make up" values? What new strategies are available? What are their strengths and pitfalls?

There are many types of studies for which missing values are an issue. I focus on survey analysis, but missing values are a problem for experimental designs and administrative data as well. Specifically, this article covers options that are available when a person agrees to participate, but then does not complete all the items. Methods for the case where a person misses a wave of a panel study, or drops out of the study prior to completion, are noted but not developed here. Much has been written regarding missing values in the statistical literature. Over the past decade, numerous strategies were introduced to family scholars that are innovative improvements over traditional approaches (Allison, 2002; Little & Rubin, 1987, 2002; Royston, 2005; Schafer, 1997; van Buuren, Brand, Groothuis-Oudshoorn, & Rubin, in press). For technical treatments of the topic, consult Little and Rubin (2002), Schafer, or van Buuren et al. Here, I focus on practical issues, applications, and guidelines for family scholars.

I begin with a discussion of types of missing values and when they are a problem. After noting the importance of missing values analysis, I present a critical review of traditional approaches to working with missing values (i.e., listwise

Department of Human Development and Family Sciences, Oregon State University, 325B Milam Hall, Corvallis, OR 97331 (alan.acock@oregonstate.edu).

Key Words: MAR, MCAR, missing data, missing values, multiple imputation.

deletion, pairwise deletion, indicator variable, and mean substitution). A relatively nontechnical review of the newer solutions that involve single imputation, multiple imputation, and full information maximum likelihood estimation follows. Finally, I illustrate several approaches with different kinds of missing values and I conclude with a series of recommendations.

Types of Missing Values

There are several classifications of missing values. These classifications influence the optimal strategy for working with missing values. This section covers data that are missing by definition of the subpopulation, missing completely at random (MCAR), missing at random (MAR), and nonignorable (NI) missing values.

Missing by Definition of the Subpopulation

Some survey participants are excluded from the analysis because they are not in the subpopulation under investigation. If the researcher is comparing the social networks of married women to those of unmarried, lesbian women, then dropping all men and all unmarried women who are not lesbians is appropriate because they do not fit the focus of the study. These respondents, then, are missing from the analysis by definition. An investigator needs to eliminate them from the data before describing any problems with missing values. An author should note the total sample size and then state the number of participants who fit the definition of the study population, namely, those participants who were married women or unmarried, lesbian women. It is important to distinguish between observations that are deleted by the nature of the subpopulation being studied and observations that should be included but who have missing values. Only the latter are problematic.

Most surveys have several codes for missing values to distinguish participants who should be treated as missing by definition from those for whom it is appropriate to impute values. The codes often distinguish among respondents who (a) refused to answer, (b) answered that they *don't know*, (c) have a valid skip, or (d) were skipped by interviewer error. A researcher should only impute values for participants who are in the subpopulation being investigated. Usually, for example, valid skips should not be imputed, although people who were skipped by interviewer error should be imputed. The distinctions between types of missing values are lost in data sets when only a single code (e.g., −9, dot) is used.

Distinguishing between those who are missing by definition and those who should have an imputed value is sometimes difficult. As noted by Little and Rubin (2002), deciding what to do with people who respond that they *don't know* is especially challenging. In some situations, a *don't know* response may be halfway between *agree* and *disagree*. For example, if asked to rate your marital satisfaction on a *very satisfied* to *very dissatisfied* scale, one researcher may say that *don't know* is halfway between *satisfied* and *dissatisfied* and assign the corresponding value. Another researcher may feel that doing so is not justified. Participants may say *don't know* because they are ambivalent; that is, sometimes *extremely satisfied* and sometimes *extremely dissatisfied*, but never halfway between the two. Giving participants a score that is halfway between *satisfied* and *dissatisfied* or imputing a value for them in some other way may not make sense from this perspective because the response options do not make sense to participants. The *don't know* option is also problematic when answering the question requires special knowledge. If people in the United States were asked to rate the average marital satisfaction of women in the Ukraine, they may answer *don't know* because they are unsure where the Ukraine is, much less know anything about gender and marital satisfaction in that country. This does not mean they are halfway between high and low; it does mean the item is not meaningful to them. Thus, imputing a value for them might be inappropriate. The researcher might define people who do not have an opinion as legitimately not part of the subpopulation being studied.

Defining the subpopulation for the study, eliminating people who do not fit this domain, and imputing values that are missing must be done with great care. These decisions need to be clear to the reader, but the reality is that far too few papers are clear in their decision-making process. A cursory review of major family journals indicates that some authors do little in the way of clarifying the process by which missing values or attrition reduced the sample size, nor do they explicate how these problems introduce potential bias in their findings. Once a data set is reduced to those participants who should

have data, an analysis of missing values and attrition is necessary. Participants who have missing values, whether they skipped individual items or dropped out of a wave of a longitudinal study, should be compared to those who are analyzed. This comparison can be done using a χ^2 test or t tests of variables for which there is information. Suppose 100 fathers are missing in the second wave of a three-wave study. This is an example of attrition. Do the 100 fathers who have 100% missing values at Wave 2 differ on education, race, and so on, using Wave 1 data? Significant differences will alert the reader to potential bias in the findings. Perhaps the people who dropped out had lower education at the time of Wave 1. Thus, the findings underrepresent people with limited education, and this may cause a bias in the result.

When there is a single wave of data, it is common for 30% of participants not to answer questions about their income. Much can be known about participants with missing values on such individual items. As with attrition analysis, the participants with missing values can be compared to those without missing values on available variables.

On the one hand, both missing values and attrition analysis alert readers to potential biases and the limits to the value of the research for generalizing. On the other hand, if there are few statistically or substantively significant differences between those who were dropped and those who were analyzed, then this reassures the reader of the strength of the analysis and of the generalizability of the findings. A summary of these comparisons can be presented in a simple table. It is recognized that such analyses are not a fully adequate test of whether the data are missing randomly.

Missing Completely at Random

The idea of values *missing completely at random* appears in every technical paper on missing values. The term has a precise meaning (Little & Rubin, 1987; Rubin, 1977): Thinking of the data set as a large matrix, the missing values are randomly distributed throughout the matrix. This rarely happens in family studies because it is well established that men, individuals in minority groups, people with high incomes, those with little education, and people who are depressed or anxious are less likely than their counterparts to answer every item in a questionnaire.

MCAR is an unreasonable assumption for many family studies. One exception is when data are *missing by design*. Giving a 100-item interview to children between ages 5 and 9 would create a serious fatigue problem. A researcher might randomly select 20 items for each child and just ask those items. Because 80% of the values for each child will be missing, using *listwise* or *casewise* deletion (see below), it is likely there would be no usable observations. These data, however, would meet the assumption of MCAR because the random process would insure that whether a child answers any one item is unrelated to the child's score on any of the 100 items. Modern approaches to missing values allow researchers to estimate population parameters that are unbiased compared to the results that would have been obtained if each child answered all 100 items and did not get fatigued in the process. The only limitation is that uncertainty is introduced by the imputation process, and this uncertainty reduces statistical power compared to having complete data.

Missing at Random

MAR is a more realistic assumption for family studies. The missing data for a variable are MAR if the likelihood of missing data on the variable is not related to the participant's score on the variable, *after controlling for other variables in the study*. These other variables provide the *mechanism* for explaining missing values. In a study of maternal depression, 10% or more of the mothers may refuse to answer questions about their level of depression. Suppose a study includes poverty status coded as 1 = *poverty status*, 0 = *not in poverty status*. A mother's score on depression is MAR if her missing values on depression do not depend on her level of depression, after controlling for poverty status. If the likelihood of refusing to answer the question is related to poverty status but is unrelated to depression within each level of poverty status, then the missing values are MAR. The issue of MAR is not whether poverty status can predict maternal depression, but whether poverty status is a mechanism to explain whether a mother will or will not report her depression level (*pattern of missingness*).

A variable is a mechanism if it helps to explain whether or not a respondent answers a question (Raghunathan, 2004; Schafer, 1997). Although family studies often have a huge problem with missing values, there is some understanding of the mechanism. Just as importantly, several of these mechanisms are included in most large surveys. Common mechanisms include education, race, age, gender, and indicators of psychological well-being. There may be other mechanisms that cannot be used because they are unmeasured. In addition to variables serving as mechanisms for missingness, there may be other causes such as sampling design. The MAR assumption is valid if it can be assumed that the pattern of missing values is conditionally random, given the observed values in the mechanism variables. These variables that serve as mechanisms explaining missingness may or may not be part of the theoretical model the researcher is using to explain the outcome variable.

NI Missing Values

Data may be missing in ways that are neither MAR nor MCAR, but nevertheless are systematic. In a panel study of college students where an outcome variable is academic performance, there is likely to be attrition because the students who drop out of college and are lost to the study are more likely to have low scores on academic performance. Ways to model NI data are beyond the scope of this paper but are addressed in Muthén and Muthén (2004).

Traditional Approaches to Working With Missing Values

Traditional approaches to working with missing values include listwise deletion, pairwise deletion, mean substitution, and inclusion of an indicator variable. Here, I describe each of these approaches and point to situations in which each approach is problematic.

Listwise or Case Deletion

Listwise or case deletion is the most common solution to missing values. It is so common that it is the default in standard statistical packages. Many researchers comment that this approach is conservative and that they do not want to "make up" data, but listwise deletion typically results in the loss of 20%–50% of the data. Of greater concern, it often addresses missing values in a systematic way. If the assumption of MCAR is met, then listwise deletion is conservative because the smaller sample size will inflate the standard errors and reduce the level of significance. Therefore, in cases where the assumptions of MCAR are met, the conservatism means increasing the risk of a Type II error. Such reduced power may not be as serious with a large sample.

If the data do not meet the assumption of MCAR, listwise deletion may yield biased estimates. Generally, there will be a bias because the complete cases may be unrepresentative of the population; for example, less educated, more mental health problems, and so on (but see Graham & Donaldson, 1993, for a description of special cases where estimates may not be biased under MAR). In the multivariate case, the bias caused by listwise deletion may exaggerate some effects or underestimate others. In either case, the estimates will be incorrect. The bias can even reverse the direction of effects asserting negative relations that are actually positive and positive relations that are actually negative (Acock, 1989; Anderson, Basilevsky, & Hum, 1985; King et al., 2001). King et al. estimate that political science articles, on average, are one standard error farther removed from the truth because of listwise deletion; their point estimators can be either too high or too low. Listwise deletion can give biased estimates of point estimators even when data are MAR (see von Hippel, 2004).

In sum, if the missing values are MCAR, then listwise deletion will give unbiased estimates and the only cost is a reduction in statistical power. If there is a sufficiently large sample, power is not an issue, and the pattern of missing values is completely random, then the listwise solution is a reasonable strategy. If there is not a large sample, however, or the missing values are not MCAR, then listwise deletion is not an optimum strategy.

Pairwise Deletion

Pairwise deletion is rarely used in family studies, although it is available in many software programs. Pairwise deletion uses all available information in the sense that all participants who answered a pair of variables are used to estimate the covariance between those variables

regardless of whether they answered other variables. For example, the covariance between income and depression might be based on 50% of the participants who answer both of the items, but the covariance between number of children and age of oldest child might be based on 99% of the participants who answered both of those items.

One reason pairwise deletion is unpopular is that it can produce a covariance matrix that is impossible for any single sample. Specifically, because each covariance could be based on a different subsample of participants, the covariances do not have the constraints they would have if all covariances were based on the same set of participants. It is possible that the pairwise correlation matrix cannot be inverted, a necessary step for estimating the regression equation and structural equation models. This problem may appear in the program output as a warning that a matrix is not positive definite. This problem can occur even when the data meet the assumption of MCAR.

A final issue with pairwise deletion is that it is difficult to compute the degrees of freedom because different parts of the model have different samples. Selecting the sample size using the correlation that has the most observations would be a mistake and would exaggerate statistical power. Selecting the sample size using the correlation that has the fewest observations would reduce power. The advantage of pairwise deletion over listwise deletion, however, is that pairwise deletion uses all the information observed.

Mean Substitution

Some researchers use mean substitution, and some programs make this option simple to use. For example, SPSS has a box the researcher can check to do it automatically. Use of the mean substitution option may be based on the fact that the mean is a reasonable guess of a value for a randomly selected observation from a normal distribution. With missing values that are not strictly random, however, the mean substitution may be a poor guess. People who are at the middle of distribution on most variables tend to be the most likely to answer questions. People at the extremes more often refuse to answer questions. Bill Gates, a cofounder of Microsoft, would not be likely to give an interviewer his annual income in a telephone survey. If a sample mean of $M = \$45{,}219$ was substituted for his missing income data, it would be a very bad guess. Similarly, people who are very poor may be reluctant to share this information in a telephone interview and estimating their income at $45,219 is also a poor guess. This potential bias applies to many variables studied by family scholars. For instance, people who are very depressed are more likely to skip items measuring depression, so substituting a mean score for these participants would make no sense.

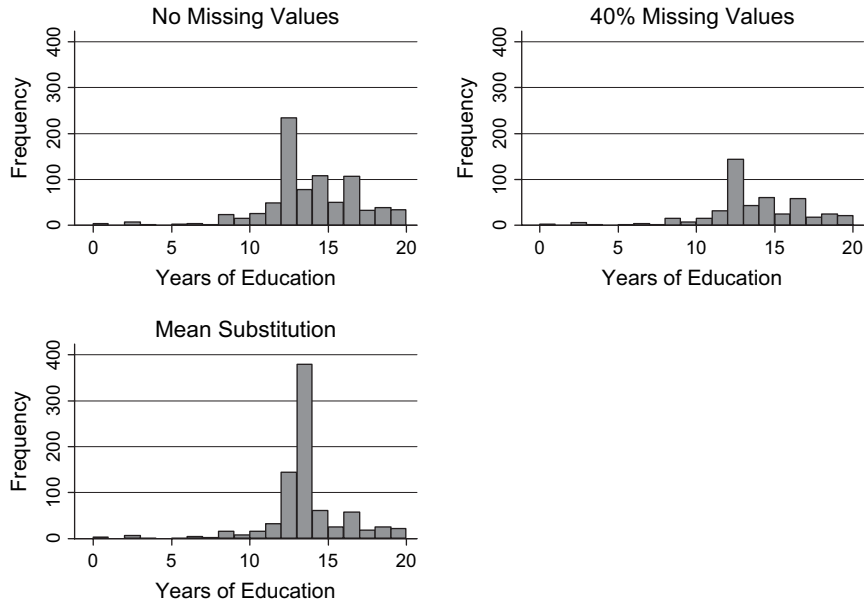
Mean substitution is especially problematic when there are many missing values. If 30% of the people do not report their income and $45,219 is substituted for each of them, then 30% of the sample has zero variance on income, thus greatly attenuating the variance of income. This attenuated variance leads to underestimating the correlation of income with any other variable.

Figure 1 illustrates how the use of mean substitution distorts the distribution and variance of reported years of education. The graph with no missing values has an $SD = 3.10$; it serves as our reference. With 40% of the scores on education MCAR, the standard deviation is similar, $SD = 3.25$. Although the distribution has fewer observations because of the missing values, the shape of the distribution is not distorted. The distribution when the mean is substituted for 40% of observations with missing values is completely distorted and has a greatly reduced standard deviation of just 2.49.

Some speculate that mean substitution will always have a conservative bias, but this is mistaken. The more missing values for the education variable that are given the same value, the greater the attenuation of the variance, and the greater the bias of the parameter estimate for the effect of education toward zero. Because different variables have different amounts of attenuation, some may have their partial effects overestimated and most will have their partial effects underestimated. Thus, the β for education with so many missing values will be underestimated, whereas the β for another variable with no missing values might be overestimated (Acock, 1989).

Mean substitution for subgroups. The mean substitution approach substitutes the mean for subgroups (Acock & Demo, 1994). Instead of substituting the mean household income for people not reporting their household income, a researcher might first categorize the sample by

FIGURE 1. MEAN SUBSTITUTION DISTORTS DISTRIBUTION AND ATTENUATES VARIANCE



marital status and compute the mean for each status. The mean for married people would be higher than the mean for never-married people. Although everybody within each category would have the same mean and this would attenuate the variance, it would both (a) be a better estimate and (b) preserve more variance than giving everyone with a missing value the overall mean.

Indicator/Dummy Variable Adjustment

Cohen and Cohen (1983) and Cohen, Cohen, West, and Aiken (2003) popularized a strategy of creating an indicator variable for missing values. If there is a single predictor, say marital conflict, a binary indicator variable is created that is coded as 1 if the *value for marital conflict is missing* and 0 if the *value for marital conflict is present*. Next, those people who have missing values on marital conflict are assigned the mean for marital conflict. (Actually, any arbitrary value will work as well.) When the model is estimated, the regression estimates will be the same as they were using listwise deletion, and the indicator variable will represent how much those with missing values differ on the mean of the outcome variable. Although this approach yields the same parameter estimates as does listwise deletion, it gives a false sense of statistical power. The additional indicator variables will, of course, use up a few degrees of freedom. For large samples, this loss of degrees of freedom will lead to far less loss of power than the artificially inflated sample size will exaggerate power. The researcher will have the full sample size rather than the listwise sample size, and the inflated sample size does not reflect the uncertainty associated with the missing values.

When researchers have multiple predictors, they can create a missing value indicator variable for each predictor. Because people who skip one item often skip other items, this use of several indicator variables can create a multicollinearity problem. Also, the parameter estimates will not be the same as with listwise deletion when there are multiple predictors. The indicator variable approach therefore can lead to biased estimates (Jones, 1996).

Summary of Traditional Approaches

Although often used, none of the traditional approaches described is an optimal solution for missing values except under specialized circumstances. These approaches can result in serious

biases in a positive or a negative direction, increase Type II errors, and underestimate correlations and β weights. Listwise deletion works reasonably well if values are MCAR and the sample is large. Unfortunately, the MCAR assumption is often unreasonable and it is misleading to call listwise deletion conservative in any sense other than increasing Type II errors. Pairwise deletion works reasonably well when researchers assume missing values are MCAR, but it can produce a covariance matrix that is not positive definite. Pairwise deletion may be less biased than other options when researchers assume MAR and have appropriate mechanisms included as covariates but the appropriate degrees of freedom for tests of significance are ambiguous. Mean substitution may be the worst choice because it attenuates variance and can produce inconsistent bias when there is great inequality in the number of missing values for different variables.

Modern Alternatives for Working With Missing Values

Several newer approaches for dealing with missing values exist, and most software programs now offer options that are more reasonable than the traditional approaches. Note that hot-deck imputation (not discussed here) has long been available and has advantages over other traditional approaches, but it has rarely been used in family studies (see, e.g., Sande, 1983). *Expectation maximization* (EM) as implemented in SPSS can impute a single new data set that has no missing values. *Multiple imputation* improves on this approach by using the consistency (or inconsistency) of estimations derived from multiple imputations as additional information, and it can estimate standard errors that are unbiased. A growing variety of software packages offer slightly different implementations of this approach. Structural equation modeling software and some multilevel software offer a full information maximum likelihood solution to missing values. In this approach, missing values are not imputed, but all observed information is used to produce the maximum likelihood estimation of parameters. Advocates of each approach are typically critics of alternatives, but often the criticisms have little consequence for practical data analysis. I briefly review these approaches by focusing on specific software implementations. References to technical explanations of each strategy are provided. Together, these approaches represent improvements over traditional approaches.

Single Imputation Using EM

EM is a maximum likelihood approach that can be used to create a new data set in which all missing values are imputed with maximum likelihood values. This approach is based on the observed relationships among all the variables and injects a degree of random error to reflect uncertainty of imputation. An explication is available in Dempster, Laird, and Rubin (1977), and a short summary is available at http://www.cs.brown.edu/research/ai/dynamics/tutorial/Documents/ExpectationMaximization.html. I do not give a detailed description of the process here. Values are imputed iteratively until successive iterations are sufficiently similar. Each successive iteration has more information because it utilizes the information from the preceding iteration. This iterative process is continued until the covariance matrix for the next iterations is virtually the same as that for the preceding iteration. This iterative maximum likelihood process usually converges quickly, but if there are many missing values and many variables, it can involve a great deal of computer time.

One way to do single imputation is to use a missing values module that is optional with the SPSS package. This SPSS MVA (missing value analysis) module will impute missing values using a variation of the EM approach. In addition to providing the imputed values, SPSS's implementation of EM provides useful information on patterns of missing data and differences between cases with and without imputed values.

In an article in *The American Statistician*, von Hippel (2004) was critical of the way SPSS implements EM in the MVA module. A key aspect of EM single imputation is that the new data set with no missing values incorporates a random disturbance term for each imputed value to reflect the uncertainty associated with the imputation. von Hippel is critical of how SPSS's MVA module does this. He stated,

> The final method, expectation maximization (EM), produces asymptotically unbiased estimates, but EM's implementation in MVA is limited to point estimates (without standard errors) of means, variances, and covariances. MVA can also impute values using the EM algorithm, but values are imputed without residual variation, so

> analyses that use the imputed values can be biased. (von Hippel, 2004, p. 160)

von Hippel acknowledges that although SPSS does not add the residual variation appropriately, it makes an adjustment later in the process. If a researcher chooses to do single imputation, there are freeware programs available that may be superior to SPSS. An example is Graham's program EMCOV, available at http://methodology.psu.edu/downloads/EMCOV.html. The NORM and Stata programs discussed in the next section also will produce single imputations.

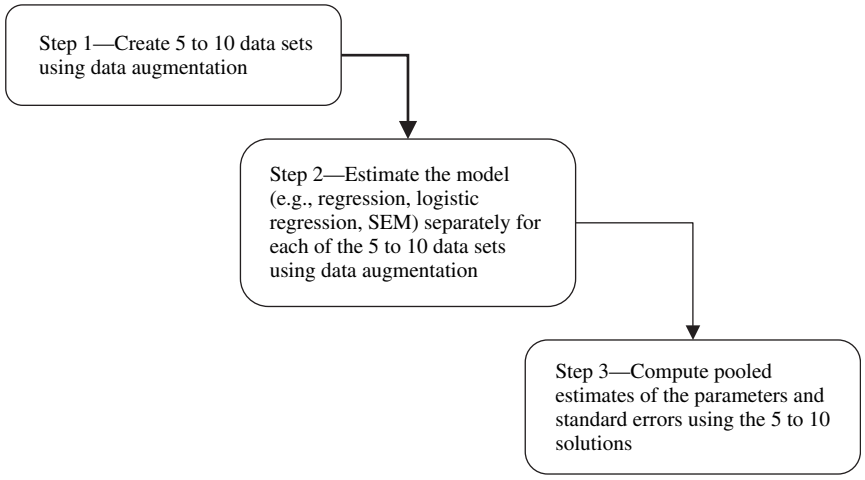
Multiple Imputation

Single imputation using EM is an important advance over traditional approaches, but it has one inherent flaw. Because single imputation omits possible differences between multiple imputations, single imputation will tend to underestimate the standard errors and thus overestimate the level of precision. Thus, single imputation gives the researcher more apparent power than the data justify. Multiple imputation (*m* separate data sets are imputed) allows pooling of the parameter estimates to obtain an improved parameter estimate. Multiple imputations produce a somewhat different solution for each imputation. If these *m* solutions were very similar, this would be evidence supporting the imputation. If these solutions differed markedly, however, then it is important to incorporate this uncertainty into the standard errors. Multiple imputation allows a researcher to incorporate this missing data uncertainty. (For more information, see http://www.stat.psu.edu/~jls/mifaq.html#whatis.)

Multiple imputation involves a three-step process as shown in Figure 2. Depending on the software being used, the process can be tedious. For example, a highly respected multiple imputation freeware package called NORM (Schafer, 1999) requires saving the data set from the standard software (e.g., SAS, SPSS, or Stata) as a plain text file, importing this file into NORM, generating 5–10 imputed data sets, exporting these back to the standard software, running the analysis for each of the 5–10 imputed data sets, recording the parameter estimates and standard errors into an ASCII file, entering these back into NORM, and then pooling the results. Other programs such as M*plus* and Stata are capable of importing the data sets imputed by NORM and combining the results in a single step. SAS (MI and MIANALYZED) and Stata (ice, mvis, and micombine) simplify this further by doing both the multiple imputations and the pooling of estimates without using other programs. M*plus* and HLM cannot generate their own multiple data sets, but they can read the *m* data sets imputed by NORM and pull the estimates.

It is reasonable to expect that all major software packages will incorporate multiple imputation methods over the next few years. If a

FIGURE 2. THREE STEPS TO MULTIPLE IMPUTATION



researcher does not have access to software that can handle multiple imputation in an integrated way, one solution is to do a single imputation for the preliminary analysis with that data set. Then, once the researcher is confident in the model, it is possible to use the multiple imputation only on the final model. The technical advantages of multiple imputations compared to single imputation are unarguable because multiple imputation allows for unbiased standard errors and single imputation does not.

How is the imputation performed? Programs vary in their capabilities. SAS offers three methods: (a) regression model, (b) propensity score method, and (c) a collection of techniques called Markov chain Monte Carlo (MCMC). The MCMC approach is advocated by Schafer (1997), who implements this approach in NORM and Stata. A technical explanation of this approach is provided by van Buuren, Boshuizen, and Knook (1999) and at http://www.multiple-imputation.com.

How are the imputations combined? Each parameter estimate is simply the mean of the m estimates, where m is the number of replications. The standard error, however, incorporates the uncertainty by adding to the mean of the error variances the variance between the solutions. Simulations reported by Schafer (1997) show that with $m = 5$ imputations, when the MAR assumption is correct, multiple imputation is 94% as efficient as if there were no missing values when actually 30% of the values are missing. A similar efficiency is achieved with $m = 10$ imputations, when 50% of the values are missing. These results are somewhat problematic, however, because they depend on the moment structures, missingness process and patterns, and the parameters under consideration. Still, they suggest that a small number of imputations, say 10, will be adequate for most applications.

There is not yet a rule of thumb for the number of imputations to use. von Hippel (2005) indicates that, when assumptions are justified, 10 imputations will produce a standard error that is just 2% larger than an infinite number of imputations when 40% of values are missing (see also Allison, 2002; Hershberger & Fisher, 2003; Rubin, 1987). Programs such as Stata, M*plus*, HLM, and SAS make it very simple to use 10–20 imputations, and a number in this range should be more than sufficient. The specialized freeware programs CAT, PAN, and MIXED are adaptations of NORM designed for categorical, panel, and mixed categorical/quantitative data, respectively. Currently, these require S-Plus software. S-Plus 6 or a later version has a library of missing data functions, but a discussion of them is beyond the scope of this article. (See van Buuren, et al., in press, for a discussion of the most general package, MICE.)

As programs develop better ways of incorporating multiple imputation, it is useful to have a set of standards for evaluating them. A freeware command called ice can be added to Stata (available by the time this article is published; Royston, 2005), and its extensive capabilities (duplicating capabilities in MICE for those using S-Plus) present a high standard. Ice can use a different estimation method for each variable depending on whether the variable is continuous (regression), binary (logistic regression), ordinal (logistic regression), or categorical with three or more categories (multinomial regression). Because some variables are continuous and others are categorical, ice's ability to use a different estimation command for each variable within a single command is an attractive feature. It can work with censored variables (survival analysis) and with interaction terms using transformations and separate equations for each variable. For example, one would not impute X_1 from the interaction term X_1X_2.

Many data sets require weighting, and ice is capable of incorporating a weight variable into the analysis. This approach works with a wide variety of analyses including regression, logistic regression, ordinal regression, multinomial regression, Poisson regression, the general linear model, and many others. Finally, it involves three single-line commands. Other than computer time, there are no marginal costs to using 10–20 imputations.

Full Information Maximum Likelihood Approaches

Structural equation modeling (e.g., SEM) and multilevel (e.g., HLM) software packages have multiple ways of working with missing values. The full information approach is available with all the major packages. This approach implements the algorithm developed by Little and Rubin (L. Muthén, personal communication, February, 2005). It does not actually impute missing values but uses all the available information

to provide a maximum likelihood estimation. I focus here on SEM programs, specifically M*plus*, but the methods apply equally to other SEM programs and to HLM. Many researchers only think of using an SEM program with latent variables, but the programs work with or without latent variables. M*plus* is used for illustration because of the extraordinary generality of its applications to situations for which missing values may be problematic. It provides maximum likelihood estimation for continuous, censored, binary, ordered categorical, categorical with three or more categories, counts, or combinations of these either with or without latent variables.

Advanced users of M*plus* have a warehouse of options available to them for handling missing values. Special approaches can be applied when data are MCAR, MAR, or NI. Robust standard errors and bootstrap standard errors are also possible. In this article, only M*plus*'s most basic approach is used. This implementation of M*plus* is intended for missing values that are MAR and uses full information maximum likelihood estimation.

One limitation of many applications of a structural equation modeling approach is that models typically only include variables that have an explicit role in the analytical model. They typically do not include available variables that may be mechanisms for missingness. It is possible, however, to include all the appropriate variables and to create a covariance matrix that is then analyzed. In M*plus*, this is simple to implement by adding the variables that are mechanisms to the model statement as outcome variables. (Simply ignore regression results where the mechanisms are outcomes; see the Appendix.) Because the inclusion of the mechanisms makes the MAR assumption much more reasonable, their inclusion would seem to be an important consideration, but it has not yet been done by researchers using structural equation modeling programs in family studies. (This approach was suggested in L. Muthén, personal communication, November, 2004). If the variables that explain missingness are excluded, it is more difficult to justify the MAR assumption.

Patterns of Missing Values

Various software packages provide information about the patterns of missing values. This information may indicate how many people missed each possible combination of items. Results are shown for each combination of two items, three items, four items, and so on. SPSS's MVA module may give the most information, even providing *t* tests on differences in the means derived from imputed values and the means derived from observed values. An example of rather basic information provided by M*plus* appears in Table 1.

Table 1 provides a list of patterns of missing data. In this case, there are 10 patterns (columns). The *x* for each variable under the first pattern indicates that these participants have no missing data. The table below the patterns shows that 550 observations have this pattern of answering all of the items. The 2nd pattern is only missing income, and the 10th pattern is missing everything except the number of children and the participant's age.

Examining the patterns of missing values can be helpful. It provides a way to see whether there might be one or two problematic variables. Some programs show the proportion of data present for each pair of variables, but the more complex patterns tell the best way of working with missing values because they pinpoint where missing values are a problem. For example, there may be one triad of variables that have a special problem that is not obvious from the number of missing values on individual variables or pairs of variables. Alternatively, if all or most of the patterns were missing the same variable, say income, dropping it or finding another indicator for the concept might be appropriate. Dropping income and keeping education might result in almost the same explanatory power as keeping both variables and dramatically minimizes the amount of missing values. In this example, Table 1 shows no special problem for income or for any other variable.

Empirical Examples of Approaches

So far, I have summarized the strengths and limitations of various approaches. In actual examples, the potential limitations and biases introduced by different approaches may be more or less problematic. To provide some indication of how well the different approaches perform, I created a data set that will serve as the standard for which there are no missing values. These data were taken from the 2002 General Social Survey and consist of 818 observations that had no missing values on a set of variables. In this example, I regressed health on number of children, general happiness, income in 1998 dollars, age, and education. The solution for this

TABLE 1. SELECTED MISSING VALUE DIAGNOSTICS PROVIDED BY M*plus*

SUMMARY OF MISSING DATA PATTERNS

MISSING DATA PATTERNS

	1	2	3	4	5	6	7	8	9	10
HLTH	x	x	x	x						
CHILDS	x	x	x	x	x	x	x	x	x	x
HAP_GEN	x	x			x	x	x			
INCOME98	x		x		x	x		x	x	
AGE	x	x	x	x	x		x	x		x
EDUC	x	x			x	x	x			

MISSING DATE PATTERN FREQUENCIES

Pattern	Frequency	Pattern	Frequency	Pattern	Frequency
1	550	5	27	9	4
2	81	6	2	10	14
3	77	7	12		
4	30	8	21		

PROPORTION OF DATA PRESENT

	HLTH	CHILDS	HAP_GEN	INCOME98	AGE	EDUC
HLTH	0.902					
CHILDS	0.902	1.000				
HAP_GEN	0.771	0.822	0.822			
INCOME98	0.767	0.833	0.708	0.833		
AGE	0.902	0.993	0.819	0.825	0.993	
EDUC	0.771	0.822	0.822	0.708	0.819	0.822

complete data set is the standard of comparison for the different approaches to missing values. Gender and satisfaction with finances were included as mechanisms on the assumption that men would be less likely than women to answer several of the items and people who were dissatisfied with their finances would be less likely than those who were satisfied to report their income.

The second data set is a 50% random sample from the target population. This data set was created using a random generation process. For this reason, it should match the assumptions of MCAR, and the resulting 410 cases should have unbiased parameter estimates. They should have larger standard errors, however, because the number of observations is reduced by half. This data set is included to compare it to the data sets that do not meet the MCAR assumption. Some methods are expected to work better with this data set than with those that are not MCAR.

A third data set has varying amounts of missing values for different variables and introduces systematic bias. Using Stata, I randomly dropped 40% of the values on general happiness for people who have fewer than 12 years of education. Another 40% of income scores were randomly dropped for people who were low or high on financial satisfaction. This procedure makes financial satisfaction a de facto mechanism for missingness on income. I also randomly dropped 40% of the education scores for those who had fewer than 12 years of education. Finally, I randomly dropped 10% of the health scores and 1% of the age scores. After merging these files together, the third data set has 431 observations using listwise deletion. The number of observed values for each variable varied because of the selection process (number of children had 818 cases, general happiness had 672, income had 681, age had 812, and education had 672). This data set is included to provide a challenging test for all of the approaches to missing values.

Table 2 summarizes the results and average margin of error for β weights, and Table 3

TABLE 2. ESTIMATES OF R^2 AND βS ALONG WITH THEIR MAGNITUDE OF ERRORS AS A PERCENTAGE OF THE ESTIMATE FOR THE COMPLETE SAMPLE

Method	R^2	No. of Children β[a]	General Happiness β	Income β	Age β	Education β	*M* |%| Error on Significant Parameter Estimates
Complete sample ($N = 818$)	.192	.056	−.187	.227	−.231	.139	
50% MCAR ($n = 410$)	.198	.063	−.193	.228	−.233	.112	7
	3%		3%	1%	1%	19%	
Listwise deletion ($n = 431$)	.203	.037	−.220	.233	−.228	.100	13
	6%		18%	3%	1%	28%	
Pairwise deletion	.193	.043	−.201	.203	−.212	.156	9
(n = unknown)	1%		7%	11%	8%	12%	
Mean substitution ($n = 818$)	.164	.036	−.185	.190	−.199	.148	9
	16%		1%	16%	4%	6%	
EM imputation	.229	.042	−.210	.239	−.228	.148	6
(SPSS; $n = 818$)	18%		12%	5%	1%	6%	
Multiple imputation	.193	.043	−.194	.236	−.213	.143	5
(NORM; $n = 818$)	1%		4%	4%	8%	3%	
Multiple imputation Stata	.204	.034	−.191	.233	−.217	.143	4
($n = 818$)	7%		2%	3%	6%	3%	
M*plus* ($n = 818$)	.199	.040	−.200	.221	−.215	.144	6
	4%		7%	3%	7%	3%	
M*plus* with mechanisms	.201	.018	−.197	.224	−.217	.143	4
($n = 818$)	5%		5%	1%	6%	3%	

Note: MCAR = missing completely at random.

[a]Not significant across all estimation methods.

summarizes the results and average margin of error for *t* tests. I focus here on the comparison of the βs. The problem with comparing the *t* tests is that single imputation can inflate the *t* tests because single imputation fails to incorporate the uncertainty introduced by the imputation process. By contrast, multiple imputation tends to reduce the *t* tests slightly because it allows us to incorporate the uncertainty associated with the imputation process. The *t* tests based on multiple imputation should be smaller than those for the sample with no missing data because they have the uncertainty inherent in the imputation process.

It must be emphasized that this is a single illustration and not a systematic simulation. It does, however, provide an idea of what happens in the face of random compared to systematically missing values. The pairwise deletion and the multiple imputation using NORM were within 1% of the actual R^2. As expected, the mean substitution attenuated the estimated R^2 by 16%. The largest bias was from the single imputation using the EM imputation within SPSS MVA. When NORM was used to produce multiple imputations, none of the five imputed data sets were remotely this far off. It must be stressed, however, that this is a single illustration and not a systematic simulation.

The 50% sample ($n = 410$) is MCAR and should be unbiased but has greatly reduced power (conservative bias leading to increased Type II errors). Table 2 shows that the R^2 of .198 is not far off, with the parameter estimates' average being about 7% off in one direction or the other. Not surprisingly, Table 3 shows that the *t* tests are off by a large percentage, 33%, and always underestimated the *t* test values obtained where there were no missing values. This is the nature of the conservative bias that affects the standard error and hence the *t* test rather than the parameter estimates.

The listwise or case deletion has the largest percentage error in parameter estimates, overestimating some and underestimating others. The absolute values of the errors in the parameter estimates are off by an average of 13%. Listwise deletion overestimated the R^2 by 6% and has an expected conservative bias (increased risk of Type II errors) in underestimating the *t* tests by an average of 20%. I emphasize here that the

TABLE 3. ESTIMATES OF *t* TESTS AND MAGNITUDE OF THEIR ERRORS AS A PERCENTAGE OF THE ESTIMATE FOR THE COMPLETE SAMPLE

Method	No. of Children *t*[a]	General Happiness *t*	Income *t*	Age *t*	Education *t*	*M* |%| Error on Significant Parameter Estimates
Complete sample ($N = 818$)	1.57	−5.81	6.45	−6.54	3.96	
50% MCAR ($n = 410$)	1.22	−4.22	4.57	−4.60	2.22	33
		27%	29%	30%	44%	
Listwise deletion ($n = 431$)	.84	−5.57	5.48	−5.26	2.35	20
		4%	15%	20%	41%	
Pairwise deletion (n = unknown)	1.01	−5.22	4.86	−5.95	3.74	13
		10%	25%	9%	6%	
Mean substitution ($n = 818$)	99	−5.67	5.49	−5.53	4.28	10
		2%	15%	15%	8%	
SPSS EM imputation ($n = 818$)	1.19	−6.59	6.79	−6.59	4.23	7
		13%	5%	1%	7%	
Multiple imputation (NORM; $n = 818$)	1.04	−5.16	5.84	−5.12	2.96	
		11%	9%	22%	25%	17
Multiple imputation (Stata; $n = 818$)	93	−5.40	5.88	−5.28	3.64	11
		7%	9%	19%	8%	
M*plus* ($n = 818$)	1.06	−5.49	5.59	−5.69	3.64	9
		6%	13%	12%	8%	
M*plus* with mechanisms ($n = 818$)	.99	−5.43	−5.75	−5.78	3.63	9
		7%	10%	11%	8%	

Note: EM = expectation maximization; MCAR = missing completely at random.

[a]Not significant across all estimation methods.

conservative bias only applies to increased Type II errors, and individual parameter estimates can be either over- or underestimated. Although pairwise deletion is one of the least popular approaches, it does a reasonable job both for the estimates and for the *t* tests. It is virtually perfect for R^2.

The mean substitution approach greatly attenuates the R^2, and this is consistent with the way mean substitution reduces the variance of variables. Its parameter estimates, on a percentage basis, are not too biased in this example, but they are slightly biased toward zero for all but one of the predictors.

The EM single imputation in SPSS MVA was anticipated to do well with the point estimates (R^2 and βs) but to have a problem with *t* tests. This did not materialize in this example because the single imputation overestimates the R^2 by 18%, the worst of the alternatives. The βs were reasonably close. The *t* test values were the closest of the alternatives, but this is not a good thing because they did not adequately incorporate the uncertainty of imputation as is incorporated with multiple imputation (von Hippel, 2004).

The multiple imputation approach using NORM worked well for this particular data set in terms of R^2 and the point estimates. The *t* tests produced by NORM were 17% smaller than the *t* tests for the target data. This was expected. Stata's mvis and micombine commands (Royston, 2004) performed well and were off on the R^2 by just 7% and even closer on the parameter estimates than NORM. Given the simplicity of using the Stata commands (see the Appendix), Stata may be an excellent choice for users who have access to it. When the ice command replaces the mvis command in Stata, the results should be even better.

I used two approaches in M*plus*, the first of which excluded mechanisms and the second of which included them. (The Appendix contains the program that includes the mechanism.) The R^2, point estimates, and *t* tests were reasonably close using M*plus* and slightly better when the mechanisms were included.

In sum, I have illustrated several of the approaches to working with missing values. The full information maximum likelihood approaches used in programs such as M*plus* and HLM as well as

the multiple imputation approach have substantial advantages over the traditional approaches. The next section provides specific recommendations on how researchers should work with missing values. These recommendations range from how to avoid missing values in the first place to best practices.

Recommendations and Conclusions

These recommendations rely on the statistical processes and potential problems rather than on the particular empirical illustration used in this article. There are three sets of recommendations: data management, less than ideal strategies, and strategies to implement.

Data management:

1. The best solution is to minimize missing values when the data are being collected.
2. A researcher should explain how cases are dropped from analysis and the percentage of observations dropped by different approaches to working with missing values.
3. Researchers should keep information on why a person has a missing value. Distinguishing what should and should not be imputed is usually impossible with a single code (e.g., −9) for every type of missing value.
4. If a *don't know* response is interpretable as being somewhere on an underlying continuum between agree and disagree, then assigning or imputing a value may be reasonable. Otherwise, it is problematic.
5. The new techniques for working with missing values are powerful tools that can be misused if researchers impute values for participants who should be excluded from the analysis.

Less than ideal strategies:

1. The mean substitution approach is probably the worst solution to missing values because it attenuates variance and often provides poor imputed values.

Table 4. Selected Software Packages Used in Working With Missing Value[a]

Software Package	Availability	Single Imputation	Multiple Imputation	Full Information Maximum Likelihood Estimation
Freeware				
Amelia	http://gking.harvard.edu/amelia/		yes	
CAT	http://www.stat.psu.edu/~jls/misoftwa.html#aut		yes	
EMCOV	http://methcenter.psu.edu/downloads/EMCOV.html	yes		
NORM	http://www.stat.psu.edu/~jls/misoftwa.html#aut	yes	yes	
MICE	http://www.multiple-imputation.com		yes	
MIXED	Free with R, commercial with S-Plus http://www.stat.psu.edu/~jls/misoftwa.html#aut	yes	yes	
MX	http://www.vcu.edu/mx/			yes
PAN	Free with R, commercial with S-Plus http://www.stat.psu.edu/~jls/misoftwa.html#aut	yes	yes	
Commercial software				
AMOS	http://www.spss.com			yes
EQS	http://www.mvsoft.com/			yes
HLM	http://www.ssicentral.com/hlm/index.html		yes	yes
LISREL	http://www.ssicentral.com/lisrel/mainlis.htm			yes
M*plus*	http://www.statmodel.com		yes	yes
SAS	http://www.sas.com		yes	
SOLAS	http://www.statsol.ie/solas/imputationtechniques.htm	yes	yes	
S-Plus	http://www.stat.psu.edu/~jls/misoftwa.html#aut, pan, cat mixed, plus other options available	yes	yes	
SPSS	http://www.spss.com, optional module	yes		
Stata	http://www.stata.com, installing ice or mvis		yes	

[a]Many of these software packages are being revised. Rather than relying on the capabilities listed in this table, consult their current Web pages.

2. Listwise or casewise deletion is acceptable if the missing values are MCAR, the sample size is sufficient such that power is not an issue, and there are few missing values.
3. Single imputation is not an optimal approach for the final analysis.

Strategies to implement:

1. Include all variables that are potential mechanisms explaining missingness even when these are not included in the analysis step (Meng, 1995; Rubin, 1996).
2. Include all variables (both predictors and outcomes) in the model at the imputation stage. If the dependent variable is related to an independent variable, this relationship should be incorporated in the imputation step. The parameter estimate for an analysis variable that is not included in the imputation step will be biased downward (King et al., 2001; Meng, 1995; Rubin, 1996).
3. It is difficult to know whether multiple imputation or full information maximum likelihood estimation is best, but both are major advances over traditional approaches. Both work best on large samples.

Table 4 provides a list of selected programs and their capabilities for working with missing values at the time of this writing. A researcher whose current choice of statistical software is limited has several options, including choices that are available as freeware. A Web page for each package is provided. Because these programs are being revised constantly, the Web pages should be consulted to learn about current capabilities for working with missing values.

The days that journals tolerate the absence of analysis of the missing values and the use of traditional approaches to missing values should be numbered, except where traditional approaches can be justified. In general, multiple imputation and the approaches available in structural equation modeling software are the best that are currently available.

REFERENCES

Acock, A. C. (1989). Measurement error in secondary data analysis. In K. Namboodiri & R. Corwin (Eds.), *Research in sociology of education and socialization* (Vol. 8, pp. 201–230). Greenwich, CT: Jai Press.

Acock, A. C., & Demo, D. (1994). *Family diversity and well-being*. Thousand Oaks, CA: Sage.

Allison, P. D. (2002). *Missing data*. Thousand Oaks, CA: Sage.

Anderson, A. B., Basilevsky, A., & Hum, D. P. J. (1985). Missing data: A review of the literature. In P. H. Rossi, J. D. Wright, & A. B. Anderson (Eds.), *Handbook of survey research* (pp. 415–494). Burlington, MA: Academic Press.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, J., Cohen, P., West, S., & Aiken, L. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, *39*, 1–39.

Graham, J. W., & Donaldson, S. I. (1993). Evaluating interventions with differential attrition: The importance of nonresponse mechanisms and the use of followup data. *Journal of Applied Psychology*, *78*, 119–128.

Hershberger, S. L., & Fisher, D. G. (2003). A note on determining the number of imputations for missing data. *Structural Equation Modeling*, *10*, 648–650.

Jones, M. P. (1996). Indicator and stratification methods for missing explanatory variables in multiple linear regression. *Journal of the American Statistical Association*, *91*, 222–230.

Juster, F. T., & Smith, J. P. (1998). Improving the quality of economic data: Lessons from the HRS and AHEAD. *Journal of the American Statistical Association*, *92*, 27.

King, G., Hopnaker, J., Joseph, A., & Scheve, K. (2001). Analyzing incomplete political science data: An alternative algorithm for multiple imputation. *American Political Science Review*, *95*, 49–69.

Little, J. R., & Rubin, D. (1987). *Statistical analysis with missing data*. New York: Wiley.

Little, J. R., & Rubin, D. (2002). *Statistical analysis with missing data* (2nd ed.). New York: Wiley.

Meng, X. L. (1995). Multiple imputation inferences with uncongenial sources of input. *Statistical Science*, *10*, 538–573.

Muthén, L., & Muthén, B. (2004). *M*plus *user guide*. Los Angeles: Statmodel.

Raghunathan, T. E. (2004). What do we do with missing data? Some options for analysis of incomplete data. *Annual Review of Public Health*, *25*, 99–117.

Royston, P. (2004). Multiple imputation of missing values. *Stata Journal*, *4*, 227–241.

Royston, P. (2005). Multiple imputation of missing values: Update. *Stata Journal*, *5*, 88–102.

Rubin, D. (1977). Formalizing subjective notions about the effect of nonrespondents in sample surveys. *Journal of the American Statistical Association*, *72*, 538–543.

Rubin, D. B. (1987). *Multiple imputation for survey nonresponse*. New York: Wiley.

Rubin, D. B. (1996). Multiple imputation after 18+ years (with discussion). *Journal of the American Statistical Association*, *91*, 473–489.

Sande, I. G. (1983). *Hot-deck imputation procedures: Incomplete data in sample surveys* (Vol. 3). New York: Academic Press.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. London: Chapman & Hall.

Schafer, J. L. (1999). NORM: *Multiple imputation of incomplete multivariate data under a normal model, version 2. Software for Windows*. Retrieved December 15, 2004, from, http://www.stat.psu.edu/~jls/misoftwa.html

van Buuren, S., Boshuizen, C. H., & Knook, D. L. (1999). Multiple imputation of missing blood pressure covariates in survival analysis. *Statistics in Medicine*, *1*, 681–694.

van Buuren, S., Brand, J. P. L., Groothuis-Oudshoorn, C. G. M., & Rubin, D. B. (in press). Fully conditional specification in multivariate imputation. *Journal of Statistical Computation and Simulation*.

von Hippel, P. T. (2004). Biases in SPSS 12.0 missing value analysis. *American Statistician*, *58*, 160–165.

von Hippel, P. T. (2005). How many imputations are needed? A comment on Hershberger and Fisher (2003). *Structural Equation Modeling*, *12*, 334–335.

APPENDIX
THE STATA PROGRAM

The Stata program has a user-written set of commands (findit ice) that generates *m* data sets, computes *m* solutions, and then combines them. I illustrate the approach using the command mvis, but the command ice will replace it by the time this article is published. The solution is provided for the unstandardized parameter estimates. The $m = 10$ regression procedures shown here are used to obtain the standardized βs and the R^2. The βs and R^2 reported in Table 2 are the means of the corresponding values from these solutions. This procedure and its extensions (miset, miappend, mimerge, misave, mido, mici, mifit, mireset) can be used with a wide variety of analytic strategies other than ordinary least squares (OLS) regression. If only unstandardized estimates are required, individual regression equations are unnecessary.

```
*multimp.do
*Multiple imputation usign mvis and micombine
use "j:\flash\missing\miss_systematic.dta", clear
mvis childs satfin male hap_gen income98 educ hlth age using imputed, \\\
      m(10) cmd(regress) replace seed(111)
use imputed, clear
micombine regress hlth childs hap_gen income98 age educ
regress hlth childs hap_gen income98 age educ in 1/818, beta
regress hlth childs hap_gen income98 age educ in 819/1636, beta
regress hlth childs hap_gen income98 age educ in 1637/2454, beta
regress hlth childs hap_gen income98 age educ in 2455/3272, beta
regress hlth childs hap_gen income98 age educ in 2373/4090, beta
regress hlth childs hap_gen income98 age educ in 4191/4908, beta
regress hlth childs hap_gen income98 age educ in 4909/5726, beta
regress hlth childs hap_gen income98 age educ in 5727/6544, beta
regress hlth childs hap_gen income98 age educ in 6545/7362, beta
regress hlth childs hap_gen income98 age educ in 7363/8180, beta
```

THE M*plus* PROGRAM

The following is the M*plus* input that allows researchers to include additional variables that explain missingness whether or not they are in the explanatory model. In this example, the mechanism variables are *satfin* and *male*. To illustrate the most general and basic case and to parallel treatment in NORM and MVA, this example does not define *male* as a nominal variable. The output in which *satfin* and *male* are outcome or *Y* variables is simply ignored, but this approach allows them to be mechanisms (L. Muthén, personal communication, November, 2004).

```
Title:
  Missing values including mechanisms
Data:
  File is miss_systematic-999.dat ;
Variable:
  Names are
  childs satfin male hap_gen ident income98 educ hlth age;
  Missing are all (-999) ;
  Usevariables are
  hlth childs hap_gen income98 age educ satfin male;
Analysis:
  Type = missing ;
Model:
  hlth on childs hap_gen income98 age educ;
  satfin on childs hap_gen income98 age educ;
  male on childs hap_gen income98 age educ;
Output:
  standardized;
```

M*plus* has other ways of working with missing values. For example, it is possible to use multiple imputations produced by other programs such as NORM or Stata as *m* data files and M*plus* will combine the multiple solutions. M*plus* also has enormous capabilities for working with NI missing data that are beyond the scope of this article.