

8-2004

Missing Data in Homicide Research

Marc Riedel
Southeastern Louisiana University

Wendy C. Regoeczi
Cleveland State University, w.regoeczi@csuohio.edu

Follow this and additional works at: https://engagedscholarship.csuohio.edu/clsoc_crim_facpub



Part of the [Criminology Commons](#)

[How does access to this work benefit you? Let us know!](#)

Repository Citation

Riedel, Marc and Regoeczi, Wendy C., "Missing Data in Homicide Research" (2004). *Sociology & Criminology Faculty Publications*. 117.

https://engagedscholarship.csuohio.edu/clsoc_crim_facpub/117

This Article is brought to you for free and open access by the Sociology & Criminology Department at EngagedScholarship@CSU. It has been accepted for inclusion in Sociology & Criminology Faculty Publications by an authorized administrator of EngagedScholarship@CSU. For more information, please contact library.es@csuohio.edu.

Missing Data in Homicide Research

MARC RIEDEL

Southeastern Louisiana University

WENDY C. REGOECZI

Cleveland State University

This article is an introduction to the special issue of Homicide Studies on missing data. The first section is an overview of the status of missing data approaches in homicide research. It begins by describing the importance of missing data estimation in homicide. This is followed by a discussion of missing data mechanisms, complete case analysis, imputation and weighting, and model-based procedures. The second section is a brief description of each of the articles in this issue. The conclusion describes the myth associated with imputing missing data, the use of missing data approaches in public records, the Supreme Court case that found hot-deck imputation acceptable for the census, and guidelines for handling missing data published by the American Psychological Association. This section concludes by describing the kinds of research that need to be done.

Keywords: missing data; ad-hoc procedures; weighting; imputation; expectation-maximization; listwise deletion; pairwise deletion; legal decisions; merging data sets; hot-deck imputation; MAR; MCAR

It is important to introduce the topic of this special issue by assuring readers that the statistical process of exploring solutions for missing data is no substitute for data collection that results in no missing values. Allison (2002) summarizes the problem:

The upshot is that although some missing data methods are clearly better than others, none of them really can be described as good. The only really good solution to the missing data problem is not to have any. So in the design and execution of research projects, it is essential to put great effort into minimizing the occurrence of missing data. Statistical adjustments can never make up for sloppy research. (pp. 2-3)

What makes the situation critical for homicide researchers is that the task of minimizing the occurrence of missing data is largely beyond their control. First, homicides are statistically rare events, which means that researchers rely on official records provided by police departments or official statistics provided by the Federal Bureau of Investigation (FBI; Riedel, 1999, 2000).

Second, participation in the Uniform Crime Reporting (UCR) program of the FBI is voluntary with respect to Supplementary Homicide Reports (SHR) as well as the National Incident-Based Reporting System (NIBRS). No federal legislation mandates state or local jurisdictions to report crime data to the FBI. The FBI has developed requirements for state UCR collection programs; currently, 44 states have met these requirements and submit crime data to the FBI (Maltz, 1999). However, as Lynn Addington notes in her article in this issue, only 25 states legally require law enforcement agencies to report to the centralized data collection center. In other words, not every state that has a collection center legally requires reporting by local agencies.

Third, checking for errors in the SHR appears to be limited to omissions or glaring anomalies. The extent to which state agencies have auditing programs is unknown; if an omission or anomaly is discovered by the FBI, the state agency is asked to follow up (Maltz, 1999).

Fourth, because homicide is reported locally and nationally by law enforcement and medical personnel, SHR data should be comparable to death certificates filed by county medical examiners/coroners and published by the National Center for Health Statistics (NCHS). This does not seem to be the case. Riedel and Regoeczi (2002) found in reviewing the literature that the larger the unit of analysis, like a nation, the greater the amount of agreement (Hindelang, 1974). Conversely, the smaller the units, like counties, the greater the disagreement (Wiersema, Loftin, & McDowall, 2000).

Finally, the SHR has large amounts of missing data because of declining arrest clearances for homicides. According to the *Uniform Crime Reports* (FBI, 2002), only 62.4% of murders and nonnegligent manslaughters were cleared by arrest in 2001. Typically, uncleared cases have no information on offender age, race/ethnicity, and gender. Because no offender is arrested, there are also large amounts of information missing on victim/offender

relationships as well as circumstances. The decline in arrest clearances extends from at least 1965, when the percentage of homicides cleared by arrest was 90.5% (Riedel, 2002).

To put it bluntly, homicide researchers use data they neither collect nor can assure themselves of their validity or reliability. Given the circumstances, missing data techniques at least offer a way of modeling our uncertainty. In some instances, as the following sections will show, missing data adjustments have been made by agencies, but these are ad hoc procedures that have been subject to little evaluation or comparison with other procedures.

Where Are We Now?

Although homicide researchers like to think they have no involvement with missing data approaches in their analyses, in fact, they use a very common technique: They simply exclude missing data. This technique is called listwise or casewise deletion or complete case analysis and is discussed subsequently. But ignoring missing data does not make it go away; not incorporating data into analyses simply runs the risk of distorting results. One traditional way to avoid any question of bias was to focus research on homicide victims rather than offenders because the former data were more complete. Thus, we have come to know much more about homicide victims than offenders.

There are historical reasons for the widespread use of listwise deletion. First, for most of us, our research training probably never covered the topic of what should be done with missing data. Second, until Rubin's conceptualization in the late 70s and 80s, missing data techniques other than listwise or pairwise deletion, if they were used at all, were largely ad hoc procedures. Finally, the model-based procedures that are now becoming well known were too computationally intensive to be used without modern computers and software, the latter being either absent or difficult to use.

Although there is no consensus on the most suitable approach to handling missing data, the subject has been given a solid conceptual foundation in the work of Donald Rubin and his colleagues (Little & Rubin, 1989; Rubin, 1976, 1987). What makes their perspective important is that statisticians now talk about "modeling our uncertainty." This arcane expression means that

model assumptions underlying the imputations can be displayed and evaluated (Little & Rubin, 2002). Describing the scientific perspective and results in sufficient detail to enable evaluation and replication is, after all, one of the central tenets of the scientific method.

Given the constraints on homicide data described earlier, it seems inevitable to us that missing data techniques will have to occupy a more important position. The quality of homicide data may eventually improve, but what are we to do in the meantime? It appears to us that focusing on how homicide data are missing and what imputation procedures are most useful is not only an incentive to law enforcement to correct problems but is fundamental to assessing and improving the reliability and validity of our findings.

What Lies Ahead?

The purpose of this article is to integrate the contributions in this special issue into the broader context of the literature on missing data. The difficulty is that there is very little consensus as to how that literature should be organized. There are a multitude of distinctions and classifications that serve particular purposes, but none that seem to have widespread consensus.

For example, there are two types of incomplete data. Item nonresponse occurs if the item is not present or if it is unusable (Madow, Nisselson, & Rubin, 1983). A persistent view in homicide research is that the “unknowns” or item nonresponses in the victim/offender relationship variable are stranger homicides because this type of homicide is more difficult to clear by arrest than those in which victims knew their offenders. A second type of incomplete data are unit nonresponse. Unit nonresponse occurs if the unit is relevant to the sample but is not included (Madow et al., 1983). Thus, because homicides are more completely reported on Return A “Crimes Known to the Police” than on the SHR, the information from the former is used to adjust the number of homicides on the latter. Although the distinction between item and unit nonresponse is important, it is not very useful as a classification of missing data techniques because some techniques can be used for both item and unit nonresponses.

This article is divided into two major sections. The first section provides a brief overview of missing data techniques using a classification offered by Little and Rubin (1987). According to these authors, there are four categories of missing data approaches. We have attempted to provide an instance of each type of missing data technique, keeping in mind that there are a number of variations and types within each classification category. We also include a section summarizing the small number of studies that adjust for missing data on victim/offender relationships. This first section serves as an important lead-in to the second section where we provide a brief description of the articles included in this issue, because some of them make use of concepts, practices, and terminology described herein.

Procedures Based on Completely Recorded Units refers to data analysis where incomplete data are simply discarded. In this essay, we discuss listwise and pairwise deletion. Not only does most homicide research use listwise and pairwise deletion, with the former being more common than the latter, our review indicates that it is a common practice in other disciplines as well. Furthermore, the practice of not including cases with missing data, as will be seen, does not necessarily distort results.

Imputation-Based Procedures include a number of techniques that use existing data to impute missing data. Imputation methods involve the substitution of some reasonable guess for missing values (Allison, 2002). The section on imputation-based procedures includes a description and discussion of the strengths and weaknesses of mean substitution, regression imputation, and hot-deck imputation.

Weighting Procedures are widely used as a means of accounting for missing data. To illustrate weighting procedures, we discuss one type, cell weighting. Cell weighting involves adjusting the sample weights so that the sample totals conform to the population totals on a cell-by-cell basis.

Model-Based Procedures make use of a defined model and base inferences on procedures such as maximum likelihood. In this section, we describe the expectation-maximization (EM) method as well as multiple imputation. Whereas there is one application of the former in homicide research, we are not aware of any applications of multiple imputation in studies of homicide.

MISSING DATA TECHNIQUES: AN OVERVIEW

Missing Data Mechanisms

Rubin and his colleagues (Little & Rubin, 2002; Rubin, 1976, 1987, 1996) have defined the different kinds of assumptions that can be made concerning patterns of missing data. The first of these assumptions is that data are Missing Completely at Random (MCAR). The data are said to be MCAR if “the probability of missing data on Y is unrelated to the value of Y itself or to the values of any other variables in the data set” (Allison, 2002, p. 4). When this is true for all variables in the data set, the data are a random sample of the original subset of observations.

One way to test whether the data are MCAR is to compare a statistic like mean victim age by dividing the data into cases missing and not missing information on another variable such as victim/offender relationships and calculating the mean for each group. If, for example, there are no significant differences using a *t* test, the data are observed at random. However, if there are significant differences, the data may be MAR (Regoeczi & Riedel, 2003).

A somewhat weaker assumption is that the data are MAR. If the data are MAR, “the probability that an observation is missing can depend on the values of observed items but not on the value of the missing item itself” (Heitjan, 1997, p. 549). In other words, the probability of missing data on a particular variable must be unrelated to the value of that variable after controlling for other variables in the analysis (Allison, 2002). With respect to homicide, this would mean that to meet the MAR assumption, missing data on victim/offender relationship could depend on the homicide motive, but within each motive category, the probability of missing the victim/offender relationship is unrelated to the victim/offender relationship (Regoeczi & Riedel, 2003). Unfortunately, unlike the condition of MCAR, there are no statistical tests of the MAR assumption.

While not based on likelihood inference and involving only a single predictor variable, the early research by Williams and Flewelling (1987) illustrates this thinking. Missingness is not related to victim/offender relationships but is related to circumstances: Robbery homicides are more likely to involve strangers. Therefore, circumstances can be used to statistically adjust for

missing data on victim/offender relationships, which results in a larger proportion of stranger homicides.

If the data are MAR and the parameters that govern the missing data process are unrelated to the parameters to be estimated, the missing data mechanism is said to be ignorable. However, Allison (2002) treats MAR and ignorability as equivalent conditions and suggests that methods that assume ignorability work “just fine, but you could do better by modeling the missing data mechanism” (p. 5).

If the data are neither MCAR nor MAR and the missingness cannot be predicted from the other variables present in the data set, the missing data mechanism is nonignorable. For example, if participants in a weight-loss study do not attend a weigh-in because of concerns about their weight, missing data for these cases are nonignorable. Schafer (1997) indicates that ignorability is closely related to the richness of the data and the complexity of the data model.

Fichman and Cummings (2003) give an example:

Suppose survey respondents that are heavy computer users are more likely to answer questions about their e-mail usage. Further, suppose that other variables in the survey do not let us predict who is likely to be a heavy computer user. If this holds, then missing data cannot be imputed with the data in hand: the missingness mechanism is nonignorable. (p. 284)

Allison (2002) notes that in the nonignorable case, the missing data mechanism must be modeled. However, he advises extreme caution because of the difficulty in choosing the appropriate model.

PROCEDURES BASED ON COMPLETELY RECORDED UNITS

Among the most common techniques used to deal with item nonresponses are listwise and pairwise deletion. Table 1 is a hypothetical data matrix of seven cases, one dependent variable (Y), and six independent variables ($X_1 - X_6$) that can be used to illustrate listwise and pairwise deletion.

TABLE 1
Data Matrix

Case	Variables						
	Y	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆
1	0	1	1	1	1	1	1
2	1	0	1	1	1	1	1
3	1	1	0	1	1	1	1
4	1	1	1	0	1	1	1
5	1	1	1	1	0	1	1
6	1	1	1	1	1	0	1
7	1	1	1	1	1	1	0

SOURCE: Taken from Riedel (2000, p. 114).

NOTE: Y = dependent variable; X = independent variable; 1 = reported value; 0 = missing value.

In listwise deletion, if any variable for a case is missing information, the entire case is omitted from the analysis. If a homicide researcher were to do a multiple regression on the data in Table 1, for example, he or she would have no cases because each case has one missing value.

Like listwise deletion, pairwise deletion uses only complete case data. The difference is that in pairwise deletion, only the cases with nonmissing values for the two variables under comparison are considered. Thus, in Table 1, a statistical test would compare (Y) to (X₁), then compare (Y) to (X₂), (Y) to (X₃), and so on. This makes a critical difference in terms of missing cases. Whereas listwise deletion would eliminate all cases, only two cases per pairwise comparison would be lost.

Strengths and Weaknesses

The major advantage to listwise and pairwise deletion is their ease of use, particularly with respect to listwise deletion. It is the default method of handling item nonresponses for commonly used statistical software packages such as SAS, SPSS, and Stata. It is an acceptable method of handling missing data if the data are MCAR. In other words, the MCAR assumption is that the data set with no missing values is a random subset of all the observations, in which case, listwise and pairwise deletion do not introduce bias.

On the downside, listwise deletion can result in a large percentage of lost cases. It is possible that item nonresponses may be

concentrated among a few cases, which would reduce the loss, but typically in SHR data, and secondary data in general, item nonresponses are scattered throughout the data set.

When the data are not MCAR, listwise deletion also results in a loss of error degrees of freedom and a loss of power to reject null hypotheses. The resulting larger standard errors provide less precision in estimating population parameters such as correlation and regression coefficients (Cohen & Cohen, 1983).

Pairwise deletion poses a different sort of problem. The correlation in Table 1 between Y and X_1 uses a different subset of cases from Y and X_2 . The correlation between Y and X_1 excludes cases 1 and 2, whereas the correlation between Y and X_2 excludes cases 1 and 3. In general, correlations are not necessarily based on the same subjects or the same number of subjects. Not only may the correlations not be comparable to one another, the means and standard deviations of variables are not referable to the same population. Furthermore, pairwise deletion can lead to statistical results that are mathematically impossible, such as correlations greater than 1.00 (Cohen & Cohen, 1983).

Frequency of Use

Listwise deletion is clearly the most popular method for dealing with incomplete data. King, Honaker, Joseph, and Scheve (2001) surveyed the use of methods for handling missing data in three political science journals—*American Political Science Review*, *American Journal of Political Science*, and *British Journal of Political Science*—for the years 1993 to 1997. Although they could gather information on only 77% of the articles, they concluded that approximately 94% of the survey articles used listwise deletion, losing about one third of their data, on average.

Roth (1994) randomly selected 45 articles from the *Journal of Applied Psychology* and 30 articles from *Personnel Psychology* for the years 1989 to 1991. For the former journal, only 41.7% of the surveyed articles discussed the problem of item nonresponse. With regard to techniques for handling missing data, 37.2% of the studies stated that none were needed, 25.6% used listwise deletion, 11.6% used pairwise deletion, and none used any other techniques. For 25.6% of the articles, the author could not tell which missing data technique was used.

For the journal, *Personnel Psychology*, only 23.3% of the survey articles discussed missing data, 23.3% of the studies claimed no missing data technique was needed, 25.5% used listwise deletion, 23.2% used pairwise deletion, and 27.9% fell into the “could not tell” category. No other missing data technique was used.

Peng, Harwell, Liou, and Ehman (2003) reviewed 1,087 studies in 11 educational journals from 1998 through 2002. Of the 1,087 studies, 28% did not report any missing data problems, 54% exhibited evidence of missing data, and the remaining 18% did not provide sufficient information. Of the 587 studies that showed evidence of missing data, 569 reported dealing with the problem. Of those 569 studies, 89.5% used listwise deletion and 7.6% used pairwise deletion. Only 2.9% used other methods of handling missing data.

Peng et al. (2003) examined the statistics for studies published in 2000, which was 1 year after the American Psychological Association Task Force on Statistical Inference published guidelines for dealing with missing data (Wilkinson & Task Force on Statistical Inference, 1999). Only four studies used the newer and more principled method of EM.

IMPUTATION-BASED PROCEDURES

Mean Substitution

In this widely used method first suggested by Wilks (1932), the mean is substituted for missing values in the distribution, with the purpose of restoring the data set to its original size. The use of mean substitution for estimating coefficients for incompletely observed variables may be effective when the variables are weakly correlated and there is a high proportion of missing data. However, the procedure is inclined to be biased where there are strong correlations among the explanatory variables, even in large samples with random patterns of missing data (Donner, 1982). This method also results in a distorted distribution in the sense that it produces a spike around the mean where all imputed cases are located (Landerman, Land, & Pieper, 1997). The assumption in mean substitution is that missing cases are missing completely at random, which may be problematic. Furthermore, this

technique does not make use of other variables, which could improve the accuracy of imputations (Landerman et al., 1997). Mean substitution has the undesirable effect of substantially decreasing measures of variance and covariance (Riedel, 1991). This is a consequence of the fact that substituted mean values contribute no additional units to the difference between the mean and the score but increase the number of cases, or the denominator, of the variance formula (Riedel, 1991). The variance will be underestimated by roughly the percentage of sample cases with data missing, leading to biased standard errors for regression coefficients and other statistics (Landerman et al., 1997). That the imputed values are a constant and are consequently not related to other variables of interest means an attenuation of correlations with these variables, in turn biasing estimated regression coefficients. In general, this method of imputing missing data should be avoided.

Regression Techniques

The general approach here is that a regression equation is computed for a given variable (Y), which contains missing data for some cases, treating it as an outcome. Cases that are not missing data on this variable are regressed on all other relevant variables, which are treated as predictors. The estimated regression equation is then used to generate predicted values for cases missing a value for Y and these values are inserted to replace item nonresponses (see Allison, 2002).

Obviously, regression imputation does not result in the loss of cases that occurs with listwise and pairwise deletion. High intercorrelations among variables may indicate that regression strategies will work well (Roth, 1994). However, research comparing listwise deletion and mean substitution to linear regression found that for less highly correlated variables ($r = .3$ or less), regression techniques do not add much accuracy over mean substitution. Finally, the sample must be large enough to generate stable regression weights (Donner, 1982). Peng et al. (2003) note that regression estimation is difficult to use when more than one variable has missing values and predicted values may exceed the logical range of scores for the missing data.

Using deterministic regression, where the residuals are set to zero, regression imputation of the dependent variable results in serious overinflation of the explanatory power of the model because of the lack of adjustment in prediction errors: Those cases with missing values on the dependent variable will be perfectly predicted. Similarly, when regression techniques are used to impute values on independent variables, they may contribute to multicollinearity because the imputed values will be perfectly correlated with the rest of the variables in the model (Acock, 1997).

To overcome this limitation, stochastic regression uses a regression estimation method in which imputed values have a random error added to them (Beale & Little, 1975). The latter approach leads to a consideration of multiple imputation, which will be discussed in a subsequent section.

Hot-Deck Imputation

Hot-deck imputation is a strategy in which a missing value (recipient) is replaced with an actual score from a similar case (donor) in the same data set. When the data set that will yield the imputed score is used, it is a "hot-deck" imputation. When data are used from another similar data set, it is called "cold-deck" imputation (Roth, 1994; Sande, 1983).

In general, a hot-deck procedure involves a duplication process: When a value is missing, a reported value is used to represent the missing value. To reduce bias, a value is chosen from a classification group that is designed to be homogeneous and as similar as possible to the group containing the missing value. Thus, it is assumed that within each classification group, the nonrespondents follow the same distribution as respondents.

For most hot-deck procedures, there is very little theory and few articles have been published in statistical journals on the procedure or definitions. Most hot-deck procedures rely on a "common-sense approach" rather than a rigorous theoretical development. For example, although variance estimators are discussed for simple cases, there are no variance formulas for more sophisticated procedures found in actual applications (Ford, 1983).

Although more sophisticated hot-deck procedures are currently used, the basic hot-deck procedure used for Current Population Surveys (CPS) illustrates the process. For item nonresponse, the Bureau of the Census divides the sample of 47,000 households into 20 cells to duplicate values for missing labor force items. The cells are formed by cross-classifying two sex groups, two race groups, and five age groups. Thus, the hot-deck procedure assumes that these three classification variables are highly correlated with labor force items and are always present.

The procedure is sequential in which the processing occurs by nearest geographic location. When the computer hits a missing labor force item, the hot-deck value of the most recently processed record of an age-race-sex cell without missing values is imputed. Where the first record has a missing value, cold-deck values are inserted from previous CPS surveys.

An alternative is to impute a randomly chosen, rather than a sequential, value from the classification group. If the records arranged in geographic order are similar, then duplicating the last reported value is better than duplicating a random one (Ford, 1983).

The major problem involves the choice of classification variables. Unless the researcher can statistically justify a limited number of variables, the inclusion of more and more variables quickly makes the process unmanageable. Thus, as indicated in the previous example with CPS, the use of three groups, two of which are dichotomies, results in 20 cells. Hot-deck imputation for categorical variables avoids a difficulty faced when using data sets containing continuous variables. Where continuous rather than categorical variables are used, collapsing into smaller categories not only discards information but introduces an element of arbitrariness (Roth, 1994; Sande, 1983).

Imputation Approaches in Existing Homicide Studies

The small number of homicide studies to address the issue of missing data has focused almost exclusively on missing values for victim/offender relationships. This research has been generally driven by the view that missing data are primarily stranger

homicides because the decline in arrest clearances is due to the inability of law enforcement to clear offenses where offenders are unknown and/or witnesses are unavailable (Reiss, 1971; Riedel, 1993; Riedel & Jarvis, 1998). Williams and Flewelling (1987) produced a weighting and adjustment procedure that resulted in an increased proportion of stranger homicides. This was followed by a study conducted by Pampel and Williams (2000) using multinomial regression that showed a disproportionate number of homicides by acquaintances.

The preceding research implicitly assumed that the missing data were MAR. Messner, Deane, and Beaulieu (2002) take the view that prior relationships between victims and offenders are interactional in nature and cannot be understood exclusively by the characteristics of one or the other. If the data were MAR, they could be imputed by victim characteristics. If the missing data mechanism is nonignorable, methods that assume MAR are suspect and imputation must take into account both victim and offender characteristics. To avoid the MAR assumption, Messner et al. (2002) use a log-multiplicative model known as the heterogeneous column RC(L) model. Their imputation method results in a greater proportion of unknown victim/offender relationships being allocated to the stranger category (which increased from 17% to 24%) than the methods used by Williams and his colleagues.

WEIGHTING PROCEDURES

Weighting for nonresponse is a very general and frequently used set of techniques. For example, weighting techniques are used by the U.S. Census Bureau (2003) in the *Survey of Income and Program Participation* (SIPP) as well as the Centers for Disease Control and Prevention's (2003) *Behavior Risk Surveillance System*. Madow, Nisselson, and Rubin (1983) describe several other government and business data sets that use weighting.

Weights are commonly assigned to records to make the weighted records represent the population of inferences as much as possible. The first step is to assign a base weight that is either the inverse of the element's selection probability or proportional to that inverse (Kalton & Flores-Cervantes, 2003). For example, in

the U.S. Census SIPP survey, the base weight for each sample person per month is the inverse of the probability of a sample person's address being selected.

The second stage of weighting is to attempt to compensate for unit or total nonresponse. This involves adjusting the base weights to compensate for the nonresponding unit. The general strategy is to identify respondents who are similar to nonrespondents in terms of auxiliary information available to both responding and nonresponding units. The base weights of respondents are then increased so that they resemble nonrespondents.

A third stage involves adjusting the weights to make the resultant weighted sample estimates conform to known population values for some key variables. Thus, a common form of adjustment forces the sample joint distribution (such as a sex by age-group distribution) to match the known joint distribution of the population.

Kalton and Flores-Cervantes (2003) provide a simple example using a sample of 1,000 and a population of 1,500. There are two auxiliary variables, "A" and "B," both of which are categorical. Variable A has four categories (A1-A4) and variable B has three categories (B1-B3). The initially weighted sample joint distribution and the population joint distribution of A and B are given in the left and center sections of Table 2.

The weighting adjustment for Cell A1/B1 is $80/20 = 4.00$ and Cell A1/B2 = 1.00. The cell weighting adjustments for sample elements in each cell are shown in the right section of Table 2.

One disadvantage of cell weighting is that it can result in large variability in the distribution of weighting adjustments, which can inflate the variances of the survey estimate. This increase in variability can reduce the precision of the survey estimates. A useful measure of the loss of precision is given by $F = 1 + CV(w_i)^2$. The expression $CV(w_i)$ is the coefficient of variation of the weights. Thus, in Table 2, where $F = 1.24$, there is a four-fold variation from 1.00 in Cell A1/B2 to 4.00 in Cells A1/B1 and A3/B3, representing a substantial decrease in effective sample size. Kalton and Flores-Cervantes (2003) also discuss additional methods of weighting including raking, generalized regression estimation, logistic regression weighting, mixtures of methods, and methods for restricting the range of resultant adjustments.

TABLE 2
An Example of Cell Weighting

	<i>Sample</i>				<i>Population</i>				<i>Weights</i>		
	<i>B1</i>	<i>B2</i>	<i>B3</i>	<i>Total</i>	<i>B1</i>	<i>B2</i>	<i>B3</i>	<i>Total</i>	<i>B1</i>	<i>B2</i>	<i>B3</i>
A1	20	40	40	100	80	40	55	175	4.00	1.00	1.38
A2	50	140	310	500	60	150	340	550	1.20	1.07	1.10
A3	100	50	50	200	170	60	200	430	1.70	1.20	4.00
A4	30	100	70	200	55	165	125	345	1.83	1.65	1.79
Total	200	330	470	1000	365	415	720	1500	F = 1.24		

SOURCE: Reproduced from Kalton and Flores-Cervantes (2003, pp. 84-85).

MODEL-BASED PROCEDURES

There are a number of procedures that define a model for the missing data and base inferences on the likelihood under that model. Parameters can be estimated by such procedures as maximum likelihood. Two advantages of the model-based procedures are flexibility and avoidance of ad hoc methods because model assumptions underlying the methods can be displayed and evaluated (Little & Rubin, 1987). This section discusses the methods of EM and multiple imputation.

Expectation-Maximization

The EM algorithm is a very general and very old maximum likelihood (ML) approach to missing data problems using the MAR assumption. Applications of this approach date back as far as McKendrick (1926), who used it on medical data. The best known work on EM is by Dempster, Laird, and Rubin (1977; Little & Schenker, 1995).

Roth (1994) summarizes the process:

The approach iterates through a process of estimating missing data then estimating parameters. Typically, the first iteration involves estimating missing data and then estimating parameters using ML estimation based on the actual data as well as the missing data estimates. The second iteration would require re-estimating the missing data based on the new parameter estimates and then recalculating the new parameters based on actual and re-estimated missing data (Little & Rubin, 1987). The approach continues until there is convergence in the parameter estimates. (p. 545)

There are several advantages to using EM. First, software is readily available in SPSS, SAS, BMDP, and SYSTAT. The MI procedure in SAS automates all of the steps (Allison, 2002). Second, it can be shown that the EM algorithm converges reliably in that under general conditions, each iteration increases the log likelihood and converges to a stationary value (Little & Rubin, 1987). The major disadvantage is that the rate of convergence can be painfully slow if there is a great deal of missing data. Acock (1997) describes one application on a very large data set with a very large number of variables and substantial missing data that ran for 20 hours. Nevertheless, his review of a large number of missing data estimation models led to the conclusion that EM was the best general solution to missing data problems.

The only instance known to the authors of EM applied to homicide data was done by Regoeczi and Riedel (2003). The purpose of that study was to determine to what extent missing data on victim/offender relationships involved stranger homicides. On one hand, some previous research described in earlier sections indicated that missing data disproportionately involved strangers. On the other hand, Decker (1993), after careful coding of St. Louis data that reduced unknown victim/offender relationships to 4%, found that stranger homicides do not account for the majority of homicides classified as unknown relationships; indeed, they may be distributed among uncleared cases in the same proportions as they are among cleared homicide cases.

None of the seven EM analyses using various combinations of victim, offender, and offense predictors and different classifications of victim/offender relationships showed any appreciable increase in stranger homicides in relation to nonimputed data. However, imputation using victim and offense characteristics, a fourfold classification of victim/offender relationships (intimate partner, other family, friend/acquaintance, and stranger), and clearance status resulted in a modest 5.5% increase in stranger homicides over nonimputed data.

Multiple Imputation

Multiple imputation was designed by Rubin (1996) to address real, practical issues. Many investigators have problems similar to homicide researchers. Homicide researchers use data sets like the SHR for a wide variety of purposes and problems. Rubin's

concern was that these users had varying degrees of statistical expertise and access to computing power, as well as showing wide variation in their ability to take account of missing data. Furthermore, widely used publicly available data sets, like the SHR, invariably contained missing values that were typically “not of any nice neat type” (p. 474). Variation in statistical expertise meant that the response of investigators to missing data would range from treating it as a nuisance to using a wide variety of adjustments and imputations that, minimally, would make comparisons of results difficult. The solution, Rubin thought, was a method of imputation that would allow any statistical technique to be applied to any data set in which missing data have been imputed and are therefore part of a complete data set. Thus, other than the usual collection of statistical techniques available in such packages as SPSS, SAS, or Stata, only software used in imputation would be needed.

The reason the latter was an “achievable basic objective” was because listwise, pairwise, and mean substitution were used in the same way, which accounted for their appeal. In other words, when one of the preceding methods of accounting for missing data was routinely applied, conventional statistical programs could be used. The problem, of course, was that listwise, pairwise, and mean substitution can yield statistically invalid answers.

Among the appeals of techniques such as multiple imputation and EM are that these methods reflect the fact that there is a level of uncertainty in the missing data estimates and, thus, address the problem of underestimated variances for variables with missing data, albeit in different ways. Whereas EM uses residual variances and covariances as a way of correcting conventional formulas, in multiple imputation, random draws are made from the residual distribution of each imputed variable and that random number is subsequently added to the imputed value (Allison, 2002; Iverson, 1984).

In multiple imputation, m values are imputed for each missing item and m data sets are created. These complete data sets will have the same observed values, but the missing values are filled in with different imputations to reflect uncertainty (King et al., 2001; Little & Rubin, 1987; Rubin, 1987; Schafer, 1997). Thus, compared with imputing a single value, multiple imputation is better able to represent the uncertainty surrounding imputed values

because each datum is replaced with two or more values drawn from an appropriate distribution for the missing values (Little & Schenker, 1995). Schafer (1999) suggests that when the rate of missing data is less than 5%, single imputation may be accurate. As a general rule, no more than 5 to 10 data sets are needed.

Each of the resulting data sets can be analyzed using conventionally available statistical methods. SAS 8.2 has a program "PROC MI" that will do the imputations and will also do statistical analyses of all data sets in a single step. The desired analysis is performed on each data set using standard complete-data methods such as multiple, logistic, or probit regression, factor analysis, variance components estimation, and so on. The parameter estimates are then averaged across the m samples to produce a single-point estimate (Allison, 2000, 2002; Little & Schenker, 1995; Rubin, 1987; Schafer, 1997, 1999). The program "MIANALYZE" in SAS 8.2 will accomplish this task. The variation in parameter estimates across imputed data sets can be used to adjust the standard errors, which would otherwise typically be too low.

We are not aware of any application of multiple imputation to homicide data. A major strength of multiple imputation is that it can be used with any kind of data and any kind of model, and it allows for the use of a wide variety of statistical techniques for analysis. In addition, it meets statistical standards that include producing estimates "that are consistent, asymptotically efficient, and asymptotically normal when the data are MAR" (Allison, 2002, p. 27). Multiple imputation can also be easily adapted to deal with suspected nonignorability (Heitjan, 1997). The imputation models can include variables that are not used in the analysis model, allowing predictions of missing data to be based on the maximum amount of available information. This is highly significant given that, in principle, imputations should be based on the predictive distribution of the missing values given the observed values and therefore should take into account all observed items (Little, 1988; Little & Schenker, 1995).

Its major disadvantage is that it can be cumbersome to implement, although SAS software has made the task much easier. A more fundamental problem is that multiple imputation will produce slightly different estimates every time it is used. The reason a unique set of numbers is not produced is because randomization is introduced in the imputation process.

CURRENT STRATEGIES FOR ADDRESSING MISSING DATA IN HOMICIDE RESEARCH

The articles in this special issue tackle the problem of missing data in homicide research from several angles. Lynn Addington's study provides a well-needed comparison of the extent of item nonresponse in the SHR and NIBRS. James Alan Fox and Robert Flewelling both contribute articles that describe the application of various weighting procedures and hot-deck imputation methods to missing homicide data. Finally, articles by Ineke Haen Marshall and Carolyn Rebecca Block, and Jason Van Court and Roger Trent, explore solutions to the problem of missing data through the merging of data sets.

Comparing Data Sources

Lynn A. Addington

Lynn Addington's article compares 1999 NIBRS data with 1999 SHR data taken from the Fox data file available at NACJD. Two groups of cases were used: One is the full SHR, whereas the other is a stratified sample to minimize small agency bias.

It is not surprising that little information is missing for victim characteristics in any of the three data sets. For offender characteristics, NIBRS had much less missing data. The major differences were between the SHR and NIBRS with respect to information on offenders, victim/offender relationships, weapons, and circumstances. The addition of an updating function is new with NIBRS, but Addington found that agencies are not updating because instructions do not provide sufficient guidance or because data edits do not trigger an update.

Weighting Procedures and Hot-Deck Imputation

James Alan Fox

Fox describes a weighting procedure applied to SHR data archived at the National Archives of Criminal Justice Data (<http://www.icpsr.umich.edu/NACJD/>), which initially used

seven categories for victim age, three victim race groups, two gender groups, 26 year categories, and 51 state classifications to produce a five-dimensional cross-tabulation with thousands of cases. The base weight was established by tallying the total number of offenders and the number that have complete offender data; the weight was the inverse proportion of complete cases. These adjustment cell weights are applied to offender records based on their cell membership and whether the offender information is complete. This original version is then compared with a revised weighting procedure that uses the characteristics of urbanness, region, gun use, and circumstances along with victim age, race, and sex in creating adjustment cells. The revised weights are much less skewed, with none exceeding 8.08; their application increases the number of youthful offenders and the percentage of Black offenders and decreases the percentage of male offenders.

In one of the few instances of hot-deck imputation with homicide data, Fox notes that the procedure is particularly well suited for categorical variables that make up most homicide data. Guided by research reported earlier in the article, Fox constructs donor pools using victim age, race, sex, urbanness, and circumstances. Fox finds that the hot-deck approach increases the percentage of young assailants, Blacks, and males.

Finally, Fox describes an ad hoc procedure for estimating the number of intimate partner homicides based on assumptions concerning what proportion of unsolved cases in various victim subgroups are likely intimate homicides. Subgroups of victims are created by stratifying by year, seven age categories, gender, five levels of urbanness, four regions, and three circumstance categories. Using this procedure, the number of intimate homicides was increased.

Robert L. Flewelling

Originally submitted as a research note, Flewelling's article conceptualizes a procedure that expands on the number of both victim and offender variables that are used to adjust for missing data beyond those used by Williams and Flewelling (1987). A record for each offender in the SHR file is created, including those where characteristics of the offender are unknown. Victim and offender characteristics, relationships, and circumstances are col-

lapsed into small numbers of categories. Their cross-classification results in 90 possible combinations of offender age, gender, race/ethnicity, and relationship with victim. For cases with known offender characteristics, a value of 1 is assigned to the dummy variable associated with that combination; the other 89 dummy variables are given a zero. For cases with missing data on offender characteristics and victim-offender relationships, the dummy variables are assigned probabilities that sum to 1.0 across the 90 dummy variables. The article provides a description of how these probabilities can be empirically derived from cases with valid data on offender characteristics and victim/offender relationships. As Flewelling notes, imputation classes may be formed and desired counts may be obtained by aggregating across categories.

Merging Data Sets

The final two articles in this issue bring the topic of missing data close to the admonition given by Allison (2002) at the outset: "The only really good solution to the missing data problem is not to have any" (p. 2). In both the Marshall and Block and the Van Court and Trent articles, the goal is to take account of missing homicide data by combining data sets that report on the same event for a more comprehensive data set.

Ineke Haen Marshall and Carolyn Rebecca Block

As any homicide researcher knows, comparative or cross-national research on homicide presents some of the most challenging problems because of differences in the nature and quality of data. Marshall and Block provide an important solution to the problem of missing cases by using two well-established methodological practices: multiple data sources and averaging.

To take account of missing cases, Marshall and Block describe an effort to create an International Homicide Index by averaging together values from multiple data sources, including homicides reported in UN Surveys for 1990 and 1994, World Health Organization death certificate data for 1992, Interpol homicide rates for 1994, and crude homicide rates for 23 countries obtained in response to a survey by the Centers for Disease Control in 1996.

They compute a rank order for each of the five data sources used, then calculate a Standardized Rank. The Standardized Rank is calculated by dividing the country's rank on a given data source by the number of countries for which data are available from that particular source and multiplying the result by 100. The composite measure, the International Homicide Index, was obtained by averaging the Standardized Ranks. Marshall and Block examine the construct validity of their index by examining its correlations with robbery victimization, feeling unsafe, perceptions of harassment and violence in the workplace, and unemployment.

Jason Van Court and Roger B. Trent

Although there have been a number of studies examining the agreement between SHR and death certificates available as vital statistics, the initial research by the Epidemiology and Prevention for Injury Control Branch of the California Department of Health Services pursued an innovative approach: They made use of a probabilistic matching procedure to match 34,542 SHR homicide cases with 170,011 death records with injury codes. They succeeded in matching 32,122 (93%) of the SHR cases with information from the Vital Statistics Death Records. An initial description of the process is given in Van Court and Trent (2001). Complete information and data are available at the National Archives of Criminal Justice Data Web site (<http://www.icpsr.umich.edu:8080/NACJD-STUDY/03482.xml>) as well as at EPICenter California Data Online (http://www.applications.dhs.ca.gov/epicdata/content/TB_linked.htm).

The goal of their article in this issue is to reduce the number of unlinked SHR cases ($N = 2,420$, or 7% of all SHRs). They were able to increase the match rate from 93.0% to 94.2% by employing a broader definition of the kinds of deaths in Vital Statistics Death Records that could be possibly linked with SHR reports. It appears that the missing cases were heavily influenced by child abuse cases. These cases were specified as child abuse homicide cases by law enforcement but not by coroners or medical examiners. The latter often classified them as "other unknown or unspecified cause."

CONCLUSION

Perhaps one of the largest obstacles to moving beyond listwise and pairwise deletion in homicide research is the myth that missing data techniques result in “made-up” data. Imputed data somehow do not have the same standing as observed data. Fichman and Cummings (2003) address this issue:

All the data we work with has some degree of uncertainty. If we collect some y_i at two points in time, estimate a test-retest reliability and determine that the two estimates of y_i are different though our model or theory of the world says they are the same, we have uncertainty in our estimate of y_i . Any score we assign someone for y_i is in some sense “made up” as when we take the average of the two scores or the sum of the two scores as our best estimate of y_i . When we use such an estimate of y_i to predict standing on some other variable z_i , the predicted score is “made-up;” computed from observations on y_i as might be done in a regression equation. . . . What is important is to represent properly the degree of uncertainty in our estimate of y_i . (pp. 302-303)

What homicide researchers have overlooked is that government agencies have been using some version of accounting for missing data for some time in making available official statistics. For example, although the FBI does not impute SHR data, it does impute crime and arrest data for underreporting agencies. Although the FBI does not publish data including imputations below the state level, NACJD publishes imputed data at the county level (Maltz, 1999).

National studies of homicide and drugs may involve data on blood alcohol tests that have been imputed. The National Center for Statistics and Analysis of the National Highway Traffic Safety Administration (2002) tried several approaches to correct for missing blood alcohol test results in the Fatality Analysis Reporting System. Whereas they formerly used discriminant function analysis, they are currently using multiple imputation.

Not only do many agencies, including the Census, use imputation methods, the U.S. Supreme Court has recently ruled that one method of imputation is legally acceptable. A recent U.S. Supreme Court case recognized that hot-deck imputation was an acceptable census procedure even when it could result in a change in the number of congressional representatives available to a state. In

Utah v. Evans (2002), Justice Breyer delivered the opinion of the court:

The question before us is whether the Census Bureau's use in the year 2000 census of a methodology called "hot-deck imputation" either (1) violates a statutory provision forbidding use of "the statistical method known as sampling" or (2) is inconsistent with the Constitution's statement that an "actual Enumeration" be made. 13 U.S.C. 195; U.S. Const., Art. I, 2, cl. 3. We conclude that use of "hot-deck imputation," violates neither the statute nor the Constitution.

After analyzing the 2000 Census data, Utah brought the lawsuit against the Secretary of Commerce and the acting director of the Bureau of the Census because, using hot-deck imputation, Utah claimed it had lost population and therefore a congressional representative. North Carolina benefited from the same method and gained a congressional representative. This would not have happened, Utah claimed, if the Census had recorded the relevant information as zero.

Although the court has prohibited sampling with respect to apportionment, it accepted the government's argument of a distinction between sampling and imputation. The government argued that imputation is not sampling (*Utah v. Evans*, 2002):

Imagine a librarian who wishes to determine the total number of books in a library. If the librarian finds a statistically sound way to select a sample (e.g., the books contained on every 10th shelf) and if the librarian then uses a statistically sound method of extrapolating from the part to the whole (e.g., multiplying by 10), then the librarian has determined the total number of books by using the statistical method known as "sampling." If, however, the librarian tries to count every book one by one, the librarian has not used sampling. Nor does the latter process suddenly become "sampling" simply because the librarian, finding empty shelf spaces, "imputes" to that empty shelf space the number of books (currently in use) that likely filled them—not even if the librarian goes about the imputation process in a rather technical way, say by measuring the size of nearby books and dividing the length of each empty shelf space by a number representing the average size of nearby books on the same shelf.

In addition, the professional association of one discipline is moving in the direction of providing publishing guidelines that

make explicit recommendations about missing data. The American Psychological Association's Board of Scientific Affairs established a Task Force on Statistical Inference that published a report on the use of statistical methods in psychology journals (Wilkinson & Task Force on Statistical Inference, 1999). The Task Force made several recommendations on how missing data should be handled. Perhaps the most prominent was the advice that listwise and pairwise were among the worst methods available for practical applications.

Finally, the events of September 11, 2001, and the enormous expenditure of funds to combat terrorism are providing indications as of this writing that government grants for providing various kinds of technology to combat terrorism are going to be increased, whereas funds for social science research are going to be decreased or eliminated. Because the pressure within academic institutions to publish is not going to lessen, homicide researchers are going to have to depend more than they already do on data sources that are accessible and inexpensive (Riedel, 2000). Homicide researchers need to be prepared to understand missing data procedures, if they are used, and how to apply them if they are not.

WHAT SHOULD BE DONE?

Based on a review of the literature as well as the articles that appear in this issue, we have some suggestions for future research. Perhaps the most important suggestion is more research directed at improving our understanding of how and why data are missing. The articles in this issue provide an important first step. Addington's study of item nonresponses in NIBRS in comparison with the SHR is an example of needed research. Although it is done in the context of describing imputation techniques, the extensive description of missing data in the SHR by Fox and the strengths and limitations of cross-national data sets by Marshall and Block are sources of extremely valuable information.

Discussion and research comparing different imputation techniques are sorely lacking. To some extent, this issue is addressed

by Fox and Flewelling, but research using the same data set and comparing different imputation techniques is needed. The innovative probabilistic matching described by Van Court and Trent has led to an extremely useful data set. We could obtain a better understanding of missing data if the same technique were applied in other states.

The integration of the types of research mentioned in the preceding paragraphs would, of course, lead to more reliable and valid data for research as well as policy and practice. It seems unlikely that much progress will be made on appropriate imputation of homicide data until we know much more about not only how and why data are missing but which kinds of missing data techniques are most appropriate. One of the most interesting questions is how a widely used method of imputation would change the results of research designed to test theories. Until that occurs, the consequences of imputations on theory testing are simply a matter of speculation.

In addition, established methods of imputation will have to be transparent and accessible, a valuable point made by Flewelling in an earlier draft of the article available here. It seems to us essential that any imputed data set should include nonimputed data so as to encourage additional research on imputation. Although Rubin's (1996) acknowledged purpose was to develop an imputation method that could be used generally, his terminology and model are difficult to understand. Furthermore, the software to do model-based imputation is just becoming available in standard packages.

Finally, a system of accountability and sanctions will ultimately have to be used to reduce the problem of missing data in homicide to a minimum. It seems incongruous to us that, on one hand, most of us have to provide truthful and complete information about social, demographic, and personal characteristics on forms that threaten us with negative consequences for incomplete or misleading information. On the other hand, no such procedure exists for information about crime recorded by the police. Although data are certainly missing for reasons that are beyond the control of law enforcement, much missing data appears to us to be a case of the rules not being applied to rule enforcers.

ACKNOWLEDGMENT

The editors for this issue would like to thank the following two people for their help in selecting, evaluating, and editing the articles. The external reviewers were Steven F. Messner, distinguished teaching professor, Department of Sociology, the University at Albany, SUNY, and Kirk R. Williams, professor of sociology, Sociology and Family Research Lab, University of New Hampshire.

REFERENCES

- Acock, A. C. (1997). Working with missing data. *Family Science Review*, 10, 76-102.
- Allison, P. (2000). Multiple imputation for missing data: A cautionary tale. *Sociological Methods and Research*, 28, 301-309.
- Allison, P. D. (2002). *Missing data*. Thousand Oaks, CA: Sage.
- Beale, E.M.L., & Little, R.J.A. (1975). Missing values in multivariate analysis. *Journal of the Royal Statistical Society*, B37, 129-146.
- Centers for Disease Control and Prevention. (2003). *Behavioral risk surveillance system: Frequently asked questions*. Retrieved September 2003 from <http://www.cdc.gov/brfss/faqs.htm#3>
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Decker, S. H. (1993). Exploring victim-offender relationships in homicide: The role of individual and event characteristics. *Justice Quarterly*, 10, 585-612.
- 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*, B39, 1-38.
- Donner, A. (1982). The relative effectiveness of procedures commonly used in multiple regression analysis for dealing with missing values. *American Statistician*, 36, 378-381.
- Federal Bureau of Investigation. (2002). *Crime in the United States Uniform Crime Reports*. Retrieved July 2003 from http://www.fbi.gov/ucr/cius_01/01crime3.pdf
- Fichman, M., & Cummings, J. N. (2003). Multiple imputation for missing data: Making the most of what you know. *Organizational Research Methods*, 6, 282-308.
- Ford, B. F. (1983). An overview of hot-deck procedures. In W. G. Madow, I. Olkin, & D. B. Rubin (Eds.), *Incomplete data in sample surveys: Theory and bibliographies* (Vol. 2, pp. 185-207). New York: Academic Press.
- Heitjan, D. F. (1997). Annotation: What can be done about missing data? Approaches to imputation. *American Journal of Public Health*, 87, 548-550.
- Hindelang, M. J. (1974). The Uniform Crime Reports revisited. *Journal of Criminal Justice*, 2, 1-17.
- Iverson, G. R. (1984). *Bayesian statistical inference*. Newbury Park, CA: Sage.
- Kalton, G., & Flores-Cervantes, I. (2003). Weighting methods. *Journal of Official Statistics*, 19, 81-97.
- King, G., Honaker, J., Joseph, A., & Scheve, K. (2001). Analyzing incomplete political science data: An alternative algorithm for multiple imputation. *American Political Science Review*, 95, 49-69.

- Landerman, L. R., Land, K. C., & Pieper, C. F. (1997). An empirical evaluation of the predictive mean matching method for imputing missing values. *Sociological Methods & Research*, 26, 3-33.
- Little, R.J.A. (1988). Missing-data adjustments in large surveys. *Journal of Business and Economic Statistics*, 6, 287-301.
- Little, R.J.A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley.
- Little, R.J.A., & Rubin, D. B. (1989). The analysis of social science data with missing values. *Sociological Methods and Research*, 18, 292-326.
- Little, R.J.A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). New York: Wiley.
- Little, R.J.A., & Schenker, N. (1995). Missing data. In G. Arminger, C. Clogg, & M. Sobel (Eds.), *Handbook of statistical modeling for the social and behavioral sciences* (pp. 39-75). New York: Plenum Press.
- Madow, W. G., Nisselson, H., & Rubin, D. B. (Eds.). (1983). *Incomplete data in sample surveys: Report and case studies* (Vol. 1). New York: Academic Press.
- Maltz, M. D. (1999). *Bridging gaps in police crime data* (No. NCJ 176365). Washington, DC: Bureau of Justice Statistics.
- McKendrick, A. G. (1926). Applications of mathematics to medical problems. *Proceedings of the Edinburgh Mathematics Society*, 44, 98-130.
- Messner, S. F., Deane, G., & Beaulieu, M. (2002). A log-multiplicative association model for allocating homicides with unknown victim-offender relationships. *Criminology*, 40, 457-480.
- National Highway Traffic Safety Administration. (2002). *Transitioning to multiple imputation—A new method to impute missing blood alcohol concentration (BAC) values in FARS*. Retrieved September 2003 from <http://www-nrd.nhtsa.dot.gov/pdf/nrd-30/NCSA/Rpts/2002/809-403.pdf>
- Pampel, F. C., & Williams, K. R. (2000). Intimacy and homicide: Compensating for missing data in the SHR. *Criminology*, 38, 661-680.
- Peng, C.Y.J., Harwell, M., Liou, S. M., & Ehman, L. H. (2003). *Advances in missing data methods and implications for education research*. Unpublished manuscript.
- Regoeczi, W. C., & Riedel, M. (2003). The application of missing data estimation models to the problem of unknown victim/offender relationships in homicide cases. *Journal of Quantitative Criminology*, 19, 155-183.
- Reiss, A. J. (1971). *The police and the public*. New Haven, CT: Yale University Press.
- Riedel, M. (1991). *Estimating missing data: Traditional approaches*. Unpublished manuscript.
- Riedel, M. (1993). *Stranger violence: A theoretical inquiry*. New York: Garland Publishing.
- Riedel, M. (1999). Sources of homicide data: A review and comparison. In M. D. Smith & M. A. Zahn (Eds.), *Homicide: A sourcebook of social research* (pp. 75-95). Thousand Oaks, CA: Sage.
- Riedel, M. (2000). *Research strategies for secondary data: A perspective for criminology and criminal justice*. Thousand Oaks, CA: Sage.
- Riedel, M. (2002). Arrest clearances. In D. Levinson (Ed.), *Encyclopedia of crime and punishment* (pp. 55-57). Thousand Oaks, CA: Sage.
- Riedel, M., & Jarvis, J. (1998). The decline of arrest clearances for criminal homicide: Causes, correlates, and third parties. *Criminal Justice Policy Review*, 9, 279-305.
- Riedel, M., & Regoeczi, W. C. (2002, June). *The California linked homicide file: Exploring a new data source*. Paper presented at the Homicide Research Working Group, St. Louis, MO.
- Roth, P. (1994). Missing data: A conceptual review for applied psychologists. *Personnel Psychology*, 47, 537-560.
- Rubin, D. (1976). Inference and missing data. *Biometrika*, 63, 581-592.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: John Wiley & Sons.

- Rubin, D. B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association*, 91, 473-489.
- Sande, I. G. (1983). Hot-deck imputation procedures. In W. G. Madow & I. Olkin (Eds.), *Incomplete data in sample surveys: Proceedings of the symposium* (Vol. 3, pp. 339-349). New York: Academic Press.
- Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. Boca Raton, FL: Chapman & Hall/CRC.
- Schafer, J. L. (1999). Multiple imputation: A primer. *Statistical Methods in Medical Research*, 8, 3-15.
- U.S. Census Bureau. (2003). *Survey of income and program participation: SIPP weighting*. Retrieved September 2003 from <http://www.sipp.census.gov/sipp/weights.html>
- Utah v. Evans. 536 U.S. 1 (2002).
- Van Court, J., & Trent, R. B. (2001). *Matching homicide reports and death records in California*. Paper presented at the Homicide Research Working Group, Orlando, FL.
- Wiersema, B., Loftin, C., & McDowall, D. (2000). A comparison of Supplementary Homicide Reports and national vital statistics system homicides estimates for U.S. counties. *Homicide Studies*, 4, 317-340.
- Wilkinson, L., & Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54, 594-604.
- Wilks, S. S. (1932). Moments and distributions of estimates of population parameters from fragmentary samples. *Annals of Mathematical Statistics*, 3, 163-195.
- Williams, K., & Flewelling, R. L. (1987). Family, acquaintance, and stranger homicide: Alternative procedures for rate calculations. *Criminology*, 25, 543-560.