Competing Imputation Approaches under Simulated Nonignorable Missingness for Perpetrator Characteristics in the FBI's Supplementary Homicide Reports

G. Lance Couzens[1], Marcus Berzofsky[1]
[1]RTI International, 3040 E Cornwallis Rd, Durham, NC 27709

Abstract
This paper compares two common methods for imputation (Fully Conditional Specification and Weighted Sequential Hot-Deck) under varying levels of simulated non-ignorable missingness in the Federal Bureau of Investigation's (FBI's) Supplementary Homicide Reports (SHRs). The SHR data contain valuable detailed information on victim, perpetrator, and incident characteristics relating to homicides occurring in the U.S., but the utility of these data is limited by the high rates of missingness on perpetrator variables resulting from homicide case insolvency. Particular attention is given to the formulation of the missingness induction mechanism which utilizes information known about victim characteristics in unsolved cases to simulate non-ignorable missingness for known perpetrators. Simulation-based sensitivity results show the methods compare similarly on these data, though neither is able to achieve adequate estimate coverage nor eliminate the directional bias in point estimates given even moderate mechanism strengths.

Key Words: imputation, nonignorable missingness, fully-conditional specification, Supplementary Homicide Reports

1. Introduction

The FBI's Supplemental Homicide Report (SHR) data series has been in existence since the early 1960s and is widely used by criminologists and policy makers to track the volume and underlying characteristics of homicide incidents in the U.S. The SHR program is unique in that it serves as the only national[1] source of data on incidents of homicide that includes characteristics of decedents and offenders, as well as circumstances of the incident. However, the utility of the data is tempered by the challenges posed by item-level missingness related to case solvency. Without careful consideration of missing data problems and how they may impact a given analysis, estimation of homicide rates and characteristics tied to case-solvency may be biased and could lead to erroneous conclusions.

1.1 Missing Data in the SHR

The main factor limiting the utility of the SHR is missing data. As in most voluntary data collections, missing data arises in two forms: unit nonresponse and item nonresponse. In the case of the SHR, unit nonresponse resulting from underreporting of homicides is the simpler problem, and can be addressed through weight-calibration to an external data source such as the FBI's aggregate Uniform Crime Reporting program's homicide statistics or the Centers for Disease Control's National Vital Statistics System (NVSS) mortality data. The focus of this paper is item nonresponse, which differs from unit nonresponse in that data records are partially complete rather than missing entirely.

Item nonresponse in the SHR is a somewhat more complex issue relative to a typical collection. This complexity arises naturally as a result of the SHR's incident focus – data related to perpetrators of homicides and their relationships to victims are only known in instances when the criminal case is solved. In 28.2% of cases from 2004 to 2013 one or more core characteristics of the perpetrator is unknown simply because the case is unsolved at the point when SHR data are submitted to the FBI. To deal with partial records, analysts have a number of options, but any approach that utilizes SHR perpetrator or relationship data must rest on assumptions regarding the nature of missingness. Specifically, one assumes that the data are missing completely at random (MCAR) or missing at random (MAR). When data are MCAR, the probability that a given item is missing does not depend on the unreported value or on any other variables in the data set. MAR relaxes that assumption by simply requiring that the probability does not depend on the value itself once other variables in the data set are controlled for. Data that are neither MCAR nor MAR are said to have nonignorable missingness and pose a significant challenge for analysts.

Though missingness that is random (either MCAR or MAR) can be expected to lead to a reduction in estimation precision (greater uncertainty due to a reduction in the number of data points), it does not lead to bias in point estimates. Controlling for what information is available and assuming MAR is common in practice, but the assumption is made more difficult with SHR data due to the explicit nature of the underlying missingness mechanism and the reasonable suspicion that cases that are not solved may be fundamentally different than those that are.

---

[1] Florida does not submit SHR data to the FBI, although data for the state are available directly, though in a different structure.

2. Methods

2.1 Imputation of Perpetrator Characteristics

The aspect of item nonresponse in the SHR that draws the most attention and warrants the most careful consideration is that which is driven by homicide case solvency. While characteristics of victims, and some characteristics of homicide incidents are knowable even when the offender is unknown to law enforcement, those of the perpetrator and his or her relationship to the victim are not. Although several imputation and weighting techniques have been evaluated for the purposes of estimating victim and offender relationship, (e.g., Wadsworth, 2008), often core perpetrator characteristics are simply assumed to be missing at random. However, it is necessary to determine the potential impacts on estimates of perpetrator age, race, and gender if missingness is not MAR.

SHR victim-level records with known perpetrator characteristics formed the base data for the imputation assessment. Using these records, perpetrator characteristic values were deleted according to a non-ignorable missingness mechanism constructed based on what's known about how victim characteristics affect case-solvency. Missingness was simulated for a large ensemble of assessment replicates, each of which was permuted over a range of non-ignorable missingness mechanism strengths. These 1,000 replicates were each imputed with two methods that best represent current and also best practices in imputation methodology: weighted, sequential hot-deck and multiple imputation via chained equations using fully-conditional specification (FCS).

Imputations were assessed on three criteria: (1) at what rate did imputation models converge[2]; (2) at what rate did post-imputation confidence intervals contain the true estimate; (3) what was the distribution of bias in point estimates (estimate with imputation – estimate with true values). Looking at these criteria across base replicates and missingness mechanism strengths allowed for a direct comparison of performance based on the points at which methods broke down when data were not missing at random. Without assuming the induced mechanism to be correct, methods may still be evaluated relative to one another. The extent to which each method may be said to be working well or not on its own is tied to the believability of the missingness induction mechanism. Using the most realistic mechanism possible, based on the empirical effects of victim demographics on solvency provided the best foundation available to make such a judgment.

2.2 Monte Carlo Framework and Simulated Missingness

Since imputation is a method for replacing missing values, it is inherently challenging to assess its performance. When there are no 'true' values to compare the imputed data to, it is impossible to say whether or not the imputation will result in unbiased estimation. To get a sense of how well a given method is performing, or to compare methods to one another, it is helpful to carry out imputation either on entirely simulated data or on real data with simulated missingness. The former approach has the notable benefit of being completely controlled by the researcher, but it can be challenging to replicate the complex relationships between variables encountered in true data sets. Since the SHR contains a mixture of continuous[3] and nominal categorical variables with relationships that are difficult to simulate, the latter approach was favored. To facilitate imputation

---

[2] In the case of hot-deck, this can be interpreted as the rate at which imputation cells must be collapsed.

[3] Though age is collected in integral form, the underlying distribution is continuous.

testing, known values for perpetrator age, race, and gender were ‘deleted’ (the true values were kept in reserve so they could be compared to imputed values) from a combined 2004-2013 SHR data file which was subset to decedent records for which all three primary characteristics for the first perpetrator[4] were non-missing (corresponding victim characteristics were already imputed as part of the weighting process).

Given that perpetrators from unsolved homicides may reasonably be assumed to differ from those in solved cases, simply deleting their characteristics in a purely random fashion would not provide a useful or accurate representation of the challenges researchers face in imputing SHR perpetrator data. The point remains, however, that the true mechanism for case-solvency is unknowable. Nevertheless, to assess candidate imputation methods, missingness should be simulated in the most realistic way possible. To that end, determination of missingness induction propensities was based on logistic models of perpetrator characteristic missingness indicators as functions of characteristics taken from victims. Simply applying this model to estimate propensity scores for missingness induction for solved cases would result in a completely MAR mechanism. In this assessment, candidate methods were stressed by making the missingness induction method non-ignorable (i.e., the missingness induction propensities were dependent on perpetrator characteristics themselves). This was achieved by applying parameter estimates and covariances from these models to known perpetrator values to estimate missingness induction propensities.

This approach resulted in more accurate missingness rates since it was based on the idea that the best information available about how perpetrator age, race, and gender might affect whether or not the perpetrator is known to law enforcement is how the same three characteristics of victims affect solvency. Under this framework, the known effects of the age, race, and gender of the decedent were transferred to induce simulated missingness for the perpetrator, leading to an SHR data set with simulated non-ignorable missingness for perpetrators that could be used to assess how imputation methods performed in overcoming data that were not missing at random. This formulation of the missingness induction mechanism can only be true if offenders and their victims look the same on average (as measured by the three core characteristics). For example, if non-white male victims are more prevalent in unsolved homicides, non-whiteness and maleness of perpetrators will result in higher propensities for missingness induction. Though this clearly cannot be asserted, this formulation was informed by empirical information from incidents and was therefore less arbitrary than a mechanism constructed from scratch. As outlined above, this approach resulted in a single data set with simulated missingness. Though a fixed mechanism for missingness induction was formulated, actual record-level missingness propensities were still random (according to that mechanism). This means that perpetrators with simulated missing values in a given data set represented a sample of cases that was subject to sampling variation. To factor that variation out of the assessment of imputation performance, missingness was simulated one hundred[5] times, resulting in one hundred unique data sets, or replicates. Measuring imputation performance across these data sets simultaneously provided a more complete picture of a given method’s accuracy and precision.

---

[4] This report focuses only on imputation as it pertains to the first perpetrator. In unsolved homicides it is unclear whether multiple offenders committed the crime and therefore whether or not missing values are valid (no perpetrators exist beyond the first) or unreported.
[5] Though this value is at least one order of magnitude lower than the number of replicates often preferred in a Monte Carlo study, processing times for imputations preclude a larger replicate set.

Once sample variation was factored out of the assessment, an additional remaining concern was whether the chosen mechanism for missingness induction was representative of the true strength of the underlying solvency mechanism at work in the SHR. This is an unanswerable question, though different imputation approaches may be assessed along a gradient of nonignorability strengths ranging from what has been inferred from victim characteristics all the way down to a nearly MCAR mechanism. This was achieved by multiplying parameter estimates from the logistic models described above by deflation factors[6] ranging from 1.0 down to 0.1 in increments of 0.1. Corresponding covariance matrices were also deflated by multiplication by the square of the deflation factor. This resulted in ten permutations of the original 100 replicates along a gradient of non-random missingness ranging from 10% to 100% of the original strength, meaning the full assessment measured performance across 1,000 unique data replicates. Finally, each replicate, though based on a number of SHR collection years, was randomly subsampled to contain approximately the number of decedents encountered in a single year. Working with data sets of this size provided the best representation of what may be encountered for a researcher working with a given single year of SHR data.

2.3 Imputation via Weighted Sequential Hot-Deck

The modern practice of imputation is carried out through a wide array of methods and across a broad spectrum of subject matter domains. Although the core issues to be overcome by any researcher dealing with a missing data problem are universal, certain data and subject types can lend themselves to one or more particular methods on the basis of practicality more than pure theoretical justifications. Such is often the case in studies dealing with human populations where variable types and the relationships between variables can be very difficult to model with imputation techniques that are often favored in more academic settings. For example, imputation based on expectation maximization (EM), as used by Regoeczi and Riedel (2003) for of victim/offender relationship data, is based on the assumption that to-be imputed data arise from a multivariate normal distribution. Clearly this is not the case for sets of variables like age, race, and gender, and using such a method requires data contortions and use of ad-hoc rules for selection of imputed values. So, while this method and others such as Markov Chain Monte Carlo (MCMC) imputation are well-established in imputation literature, they do not apply naturally to data sets such as the SHR.

In a practical sense, hot-deck imputation is appealing in that the structure of the response data itself, including mixed variable types and complex inter-variable relationships, is implicitly used to fill in missing values without requiring an explicit model for these distributions and relationships. In certain situations where missing data patterns are complex, hot-deck can be difficult to implement. In the case of SHR perpetrator data,

---

[6] By diluting the missingness mechanism in this way, the magnitude of missingness is also affected. As the deflation factor approaches zero, so does $\vec{\boldsymbol{\beta}}$, and as a result, the predicted propensity $\hat{p} = 1/[1 + e^{-\vec{\beta}}]$ approaches 0.5. To keep the missingness magnitude constant, the following factor is added to $\vec{\boldsymbol{\beta}}$:

$$(1 - d) * \ln\left[\frac{c}{1 - c}\right]$$

Where d is the deflation factor and c is the desired level of missingness (ranging from roughly 30-35%, depending on the perpetrator characteristic in question).

however, the core variables of interest are missing simultaneously in an overwhelming majority of cases. This makes sense in that the largest driver of perpetrator characteristic missingness is case solvency. When the perpetrator is unknown, all characteristics are missing – in 84.6% of SHR victim records with one or more of perpetrator age, race, and gender missing, all are missing. This relatively simple data structure (though it arises from a complex missingness mechanism) combined with the appeal of having a single, complete data file for analysis makes hot-deck imputation very appealing for the SHR. Weighted, sequential hot-deck was used as one candidate method for imputation of perpetrator characteristics in this assessment. This approach utilizes information from nonresponse-adjustment weights as well as imputation classes and intraclass sorting to improve on basic hot-deck. For records missing all three perpetrator characteristics, a single donor could be used for imputation, though this approach would ignore potentially important information in non-missing variables for the 15.4% of cases where only one or two characteristics are missing. For this reason, perpetrator age, race, and gender were imputed sequentially and conditional on imputed (or reported) values from prior steps in the imputation sequence, with the sequence itself determined by the magnitude of missingness.

Within the context of the Monte Carlo framework, WSHD imputations were carried out automatically with imputation cells defined by the cross-classification of victim gender, victim race, and geographic division[7]. Cells were sorted by victim age and year. Imputations were sequential, in the order of perpetrator gender, race, and age. For the imputation of race and age, imputed gender was added to the imputation cell definition. Additionally, when imputing age, imputed race was added to the cell definition.

For each replicate, imputation with the above specifications was attempted for the three core perpetrator characteristics. The number of imputations for each variable that could be successfully completed with the full specification was tracked across replicates. This provided an estimate for the proportion of time which WSHD can be expected to succeed without collapsing imputation cells, thereby potentially reducing its ability to control for the underlying missingness mechanism. Additionally, once imputation was complete for a given replicate, each perpetrator characteristic was estimated along with corresponding 95% confidence intervals. Across replicates, the post-imputation bias (estimate with imputation minus the true estimate) was measured along with the proportion of convergent replicates where an estimate's confidence interval contained the true estimate. Taken together, these measurements were used to assess how well WSHD performs at controlling bias both in perpetrator characteristics themselves and in their variances.

2.4 Multiple Imputation via Fully-Conditional Specification

Fully-Conditional Specification, or FCS, is a method for multiple imputation in which a set of variables is imputed in an iterative fashion based on univariate models conditioned on previous imputations of the other variables in the set (plus other complete predictors). This method, also commonly referred to as MICE (van Buuren, Boshuizen, and Knook, 1999), is in many ways similar to the MCMC method, though it does not rely on the assumption of an explicit multivariate density for the set of imputation variables. As noted previously, the MCMC method is based on a multivariate normal assumption. For

[7] Geographic division contains the following categories: Possessions, New England states, Middle Atlantic States, East North Central States, West North Central States, South Atlantic States, East South Central States, West South Central States, Mountain States, and Pacific States.

this reason, FCS is a popular choice for practitioners dealing with clearly non-normal data such as in the case of the SHR.

Within the context of the Monte Carlo replication framework, FSC imputations were specified in three separate equations. For each equation, the set of complete predictors $\vec{x}$ contained victim age, victim race, victim gender, geographic division, year, and year squared. Perpetrator age was modeled via truncated regression (lower limit=10; upper limit=100) as a function of $\vec{x}$ and perpetrator race and gender. Perpetrator race was modeled as a multinomial logistic function of $\vec{x}$ and perpetrator age and gender. Perpetrator gender was modeled as a logistic function of $\vec{x}$ and perpetrator age and race. This ensemble of models was estimated iteratively using imputed values from the prior iteration to enhance variable prediction and agreement with correlates. For each missing value in a given replicate, ten imputations were estimated. Variation across imputations was used to measure and incorporate uncertainty resulting from imputation in standard errors.

In a similar fashion to what is described above for WSHD, performance metrics were tracked across replicates. For each replicate, bias in point estimates, coverage of true estimate by confidence intervals, and model convergence were tracked. These metrics were compared to those resulting from WSHD across missingness mechanism strengths.

3. Results

3.1 Perpetrator Imputation Model Convergence Rates: FCS vs. WSHD

To compare the WSHD and FCS methods, the rate at which each can successfully complete imputation for a given variable without modification was measured across replicates and missingness mechanism strengths. For WSHD that modification took the form of collapsing or removing variables to form imputation classes, and for FCS that meant collapsing or removing predictors. Rather than making those adjustments, the proportions of replicates in which they would have proved necessary were tracked. In these instances, imputation would still be possible, but the resulting models would utilize less predictive information and might be less successful at controlling nonrandom missingness mechanisms. Since WSHD is sequential, the convergence rate is variable-specific. For FCS the rate is constant across imputation variables because the process is cyclical. If one component model did not converge, the same was said to be true for the entire process. Interestingly, for WSHD the only variable where convergence was not universally obtained was perpetrator age. This is directly related to the fact that age has the most missingness (~35%) and was imputed last, using imputed gender and race for the formulation of imputation cells. Race specifically proved to be problematic and in practice may need to be collapsed.

Figure 1 shows a 100% convergence rate for WSHD, whereas FCS achieves convergence in around 70% of replicates, declining only slightly as the strength of the missingness mechanism increases (NMAR stands for Not Missing At Random). The situation is the same for perpetrator race.

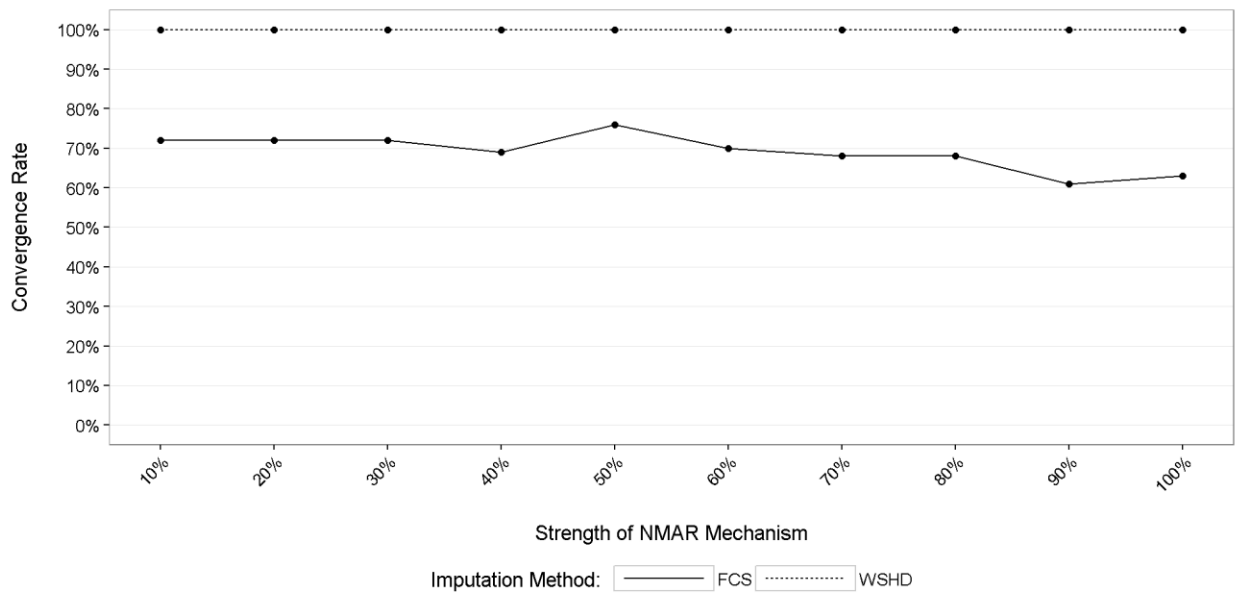

Figure 1: Convergence Rates for Imputation of Perpetrator Gender

Figure 2 shows a convergence trendline for FCS identical to the one shown in Figure 1 (in the case of FCS, the convergence rate was constant across imputation variables). As noted previously, due to the use of imputed race in the construction of imputation cells for perpetrator age, the convergence rate for WSHD was not 100%. Interestingly, the rate trended toward improvement as the strength of the NMAR mechanism increased – it's not clear what's driving this phenomenon. Nevertheless, WSHD was clearly superior to FCS in terms of convergence over the full range of variables and NMAR strengths.

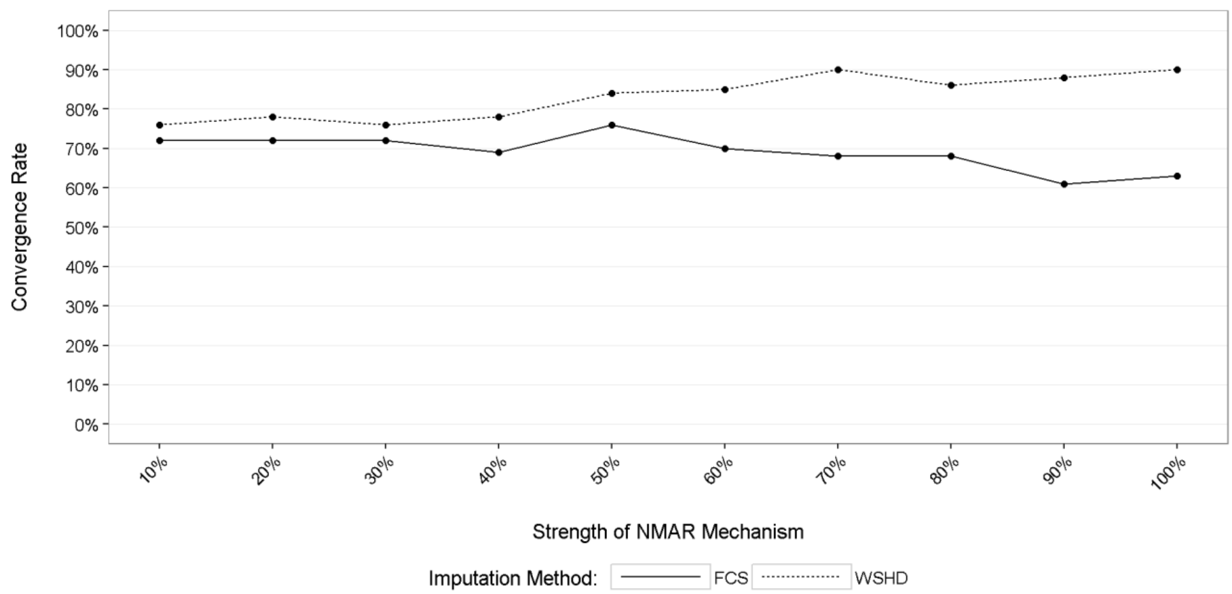

Figure 2: Convergence Rates for Perpetrator Age

3.2 Perpetrator Estimate Coverage Rates: FCS vs. WSHD

To assess how well the two methods performed at generating confidence intervals that captured true estimates of perpetrator characteristics, proportions (gender and race levels) and means (age) estimated prior to missingness induction were compared with upper and lower 95% confidence limits derived from post-imputation data. Since FCS was implemented within a multiple imputation framework, it is expected that it has the potential to be superior on this metric given the incorporation of imputation uncertainty in variance estimation (all else being equal, confidence intervals should be wider with FCS). This advantage can be erased, however, if FCS results in more directional bias than WSHD. Figures 3 through 6 show the coverage rate (the percentage of replicates with confidence intervals that captured the true estimate) for all characteristics of interest and over the full range of nonrandom mechanism strengths.

Figure 3 shows that FCS was superior to WSHD in the sense that the coverage rate broke down more slowly. This is in line with expectations owing to the conservative variance estimators used for FCS. That both methods captured 0% of the true estimates by 60% strength is notable, however. As discussed previously, the method for inducing missingness was based on the best available information as to how the demographics of one of the parties to a homicide incident – the victim – affect the probability of perpetrator missingness, but it is unknowable whether or not the magnitude or direction of the effects or their covariances approximate reality.

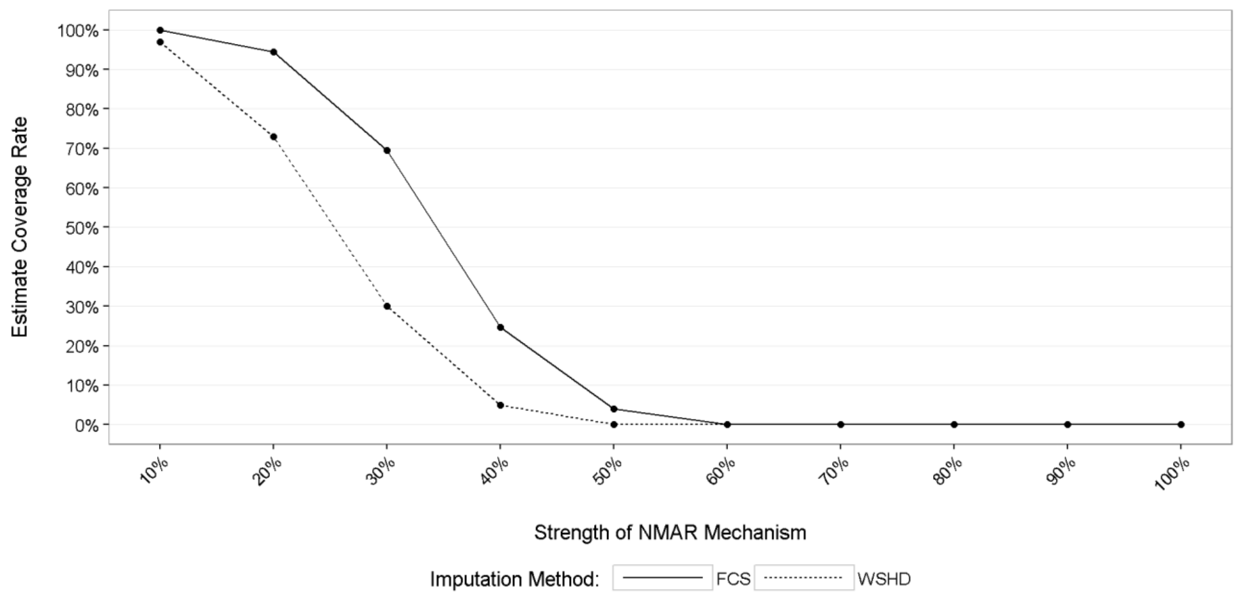

Figure 3: Coverage Rate Comparison for Perpetrator Gender (Percent Male)

Figure 4 again shows FCS breaking down more slowly than WSHD. In this case, both methods failed completely at capturing true estimates by 70% NMAR mechanism strength.

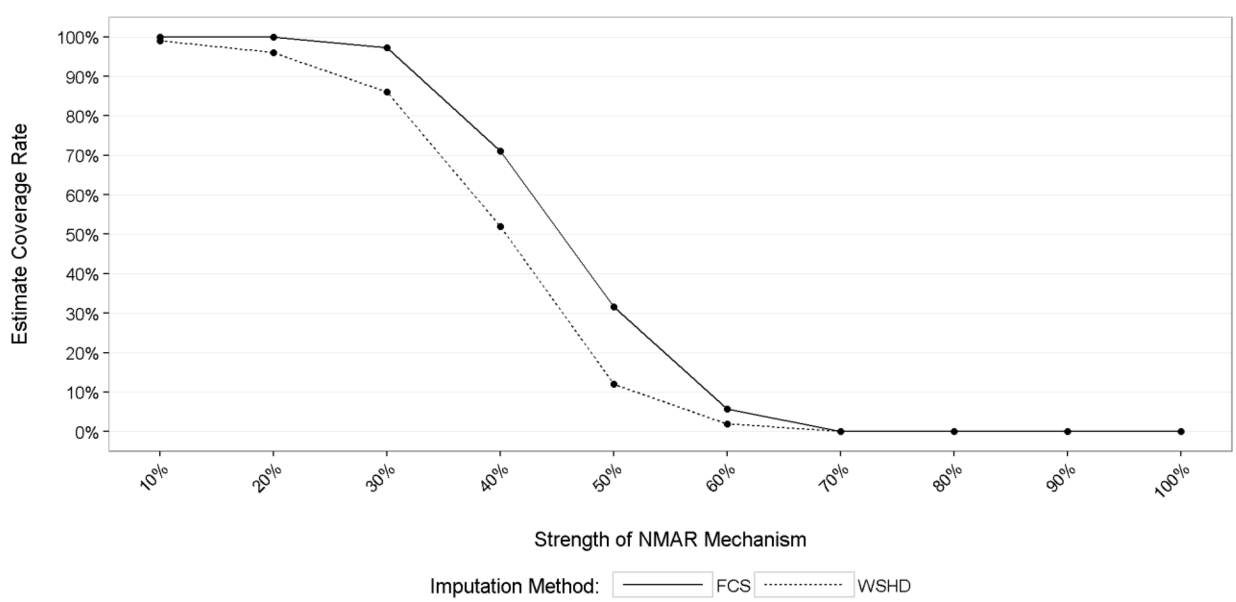

Figure 4: Coverage Rate Comparison for Perpetrator Race (Percent White)

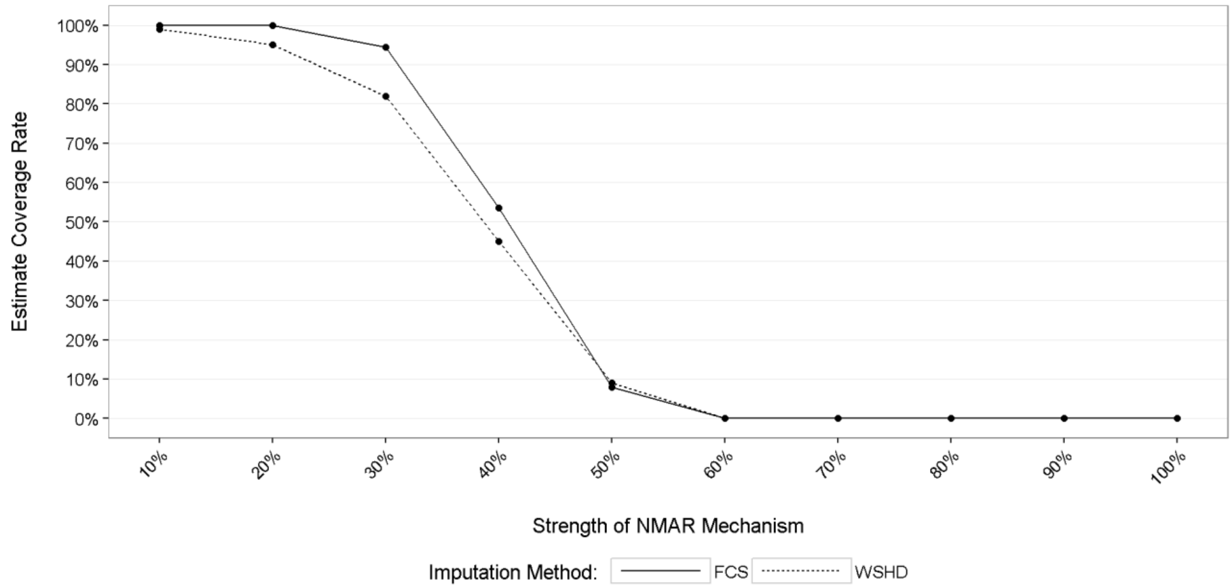

Figure 5: Coverage Rate Comparison for Perpetrator Race (Percent Black)

The strong performance for these estimates was due to two factors: the relative unimportance of those characteristics in the missingness mechanism and small sample sizes (especially in the case of other race). Both methods performed well at capturing perpetrator age (Figure 6). This is likely attributable to low levels of directional bias in the estimation of perpetrator age (similar to the case with percent Hispanic and percent other race).

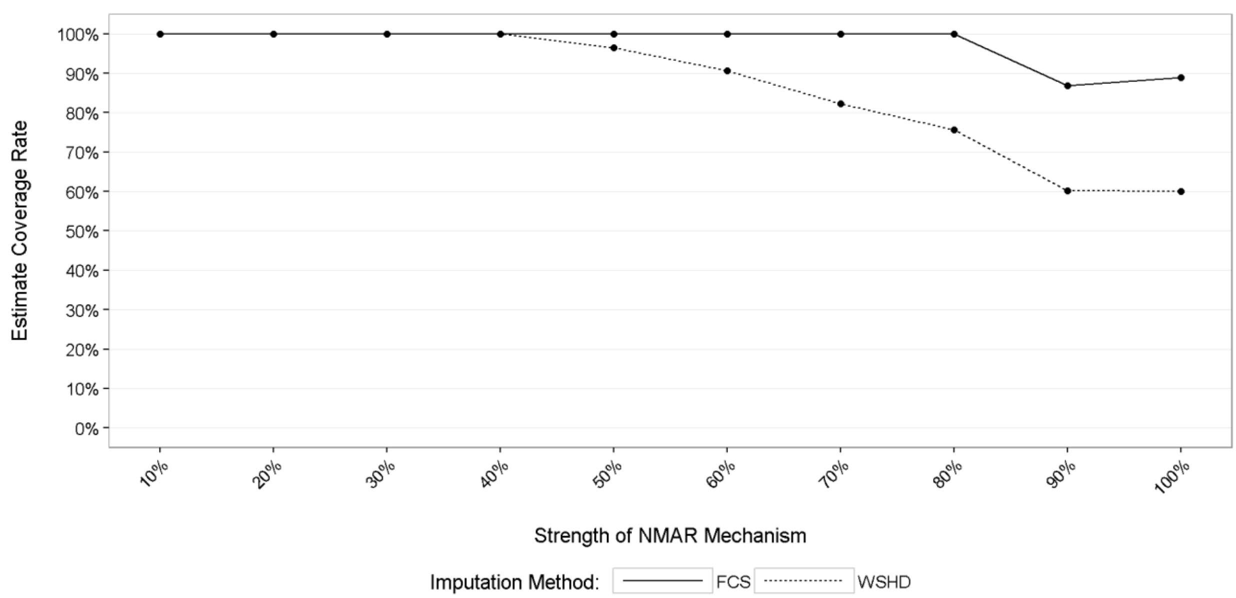

Figure 6: Coverage Rate Comparison for Perpetrator Age

By the coverage metric, FCS imputation was clearly superior to WSHD. What remains somewhat unclear is whether FCS performed well on its own. If it is assumed that the strength of the NMAR mechanism employed here is realistic (without assuming anything about direction or correlation), it can at least be said that neither method performed well in the face of a strong mechanism for any characteristic that was a driver of that mechanism. What this analysis cannot reveal is the direction of failure when it occurred. In other words, it can’t be said based on this alone whether, when the methods failed to capture the true estimate, if it was always in the same direction. This is the question addressed with the analysis of directional bias covered in the next section.

3.3 Perpetrator Estimate Directional Bias: FCS vs. WSHD

In the previous section it was showed that both methods often failed to capture true estimates under even a moderately strong NMAR mechanism, but this doesn’t establish the direction and magnitude of bias in point estimates. Figures 7 through 12 show how well WSHD and FCS imputation methods supported unbiased point estimates in the face of non-ignorable missingness.

Figures 7 and 8 show that FCS and WSHD performed similarly (and not well) at supporting unbiased point estimates for perpetrator race. Both performed badly for non-Hispanic white and non-Hispanic black and reasonably well for non-Hispanic other and Hispanic. WSHD appears to have performed slightly better for the latter two race categories, but not drastically so.

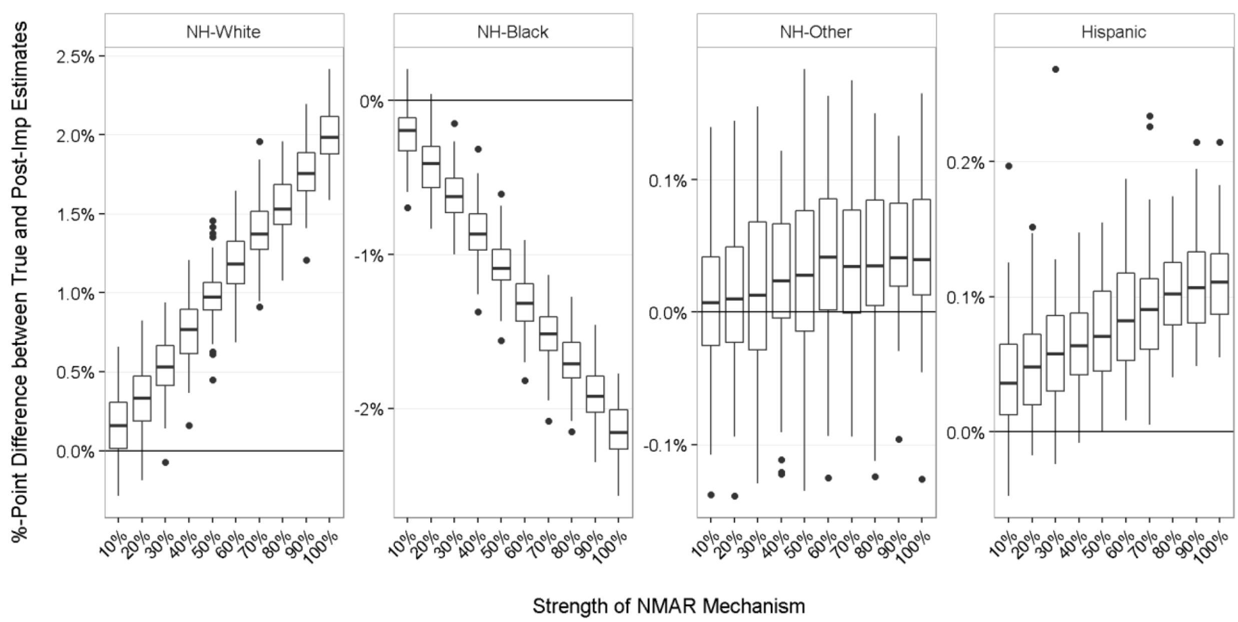

Figure 7: Race Estimate Difference (Post-Imp. — True) Distributions across Replicates and NMAR Strengths (FCS)

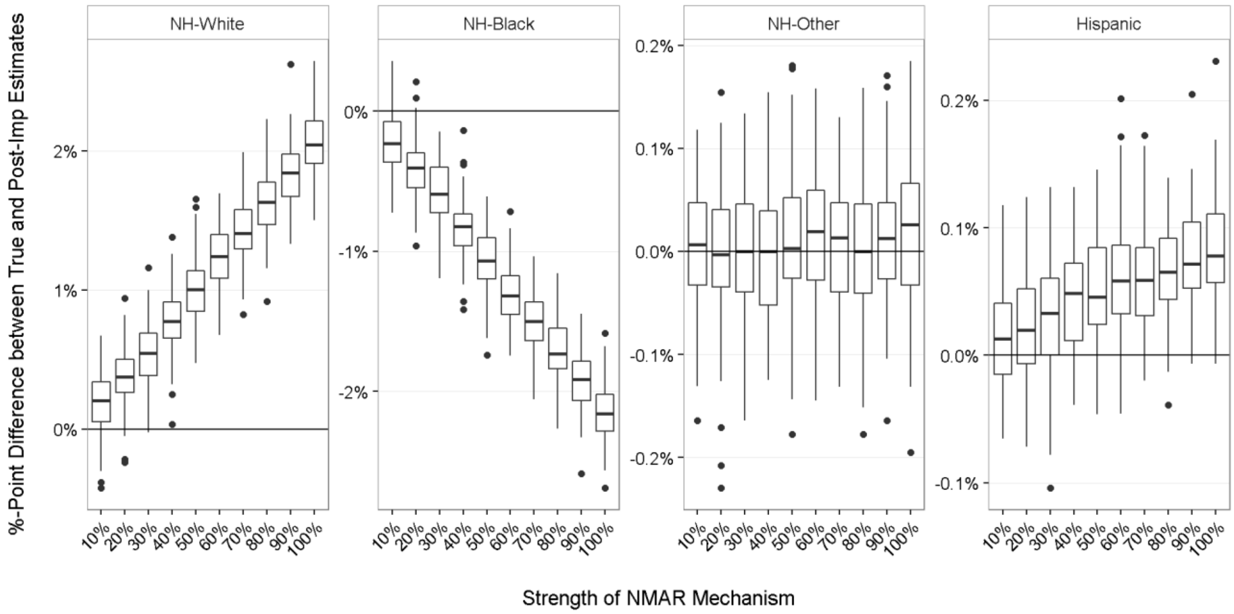

Figure 8: Race Estimate Difference (Post-Imp. — True) Distributions across Replicates and NMAR Strengths (WSHD)

Both methods showed significant directional bias in the estimation of perpetrator gender (% male), as shown in Figures 9 and 10, with no apparent performance differences between them.

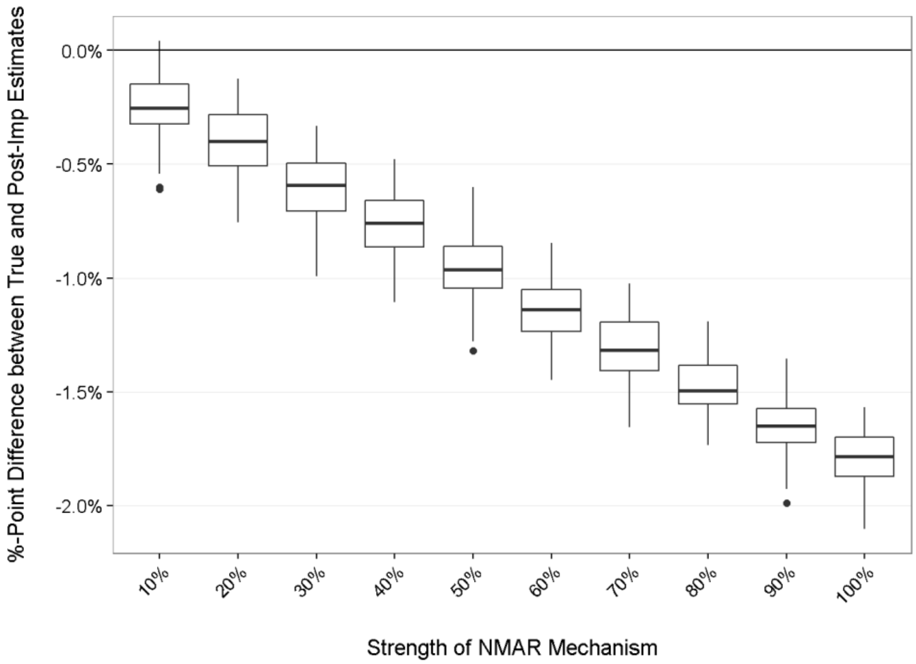

Figure 9: Sex Estimate Difference (Post-Imp. — True) Distributions across Replicates and NMAR Strengths (FCS)

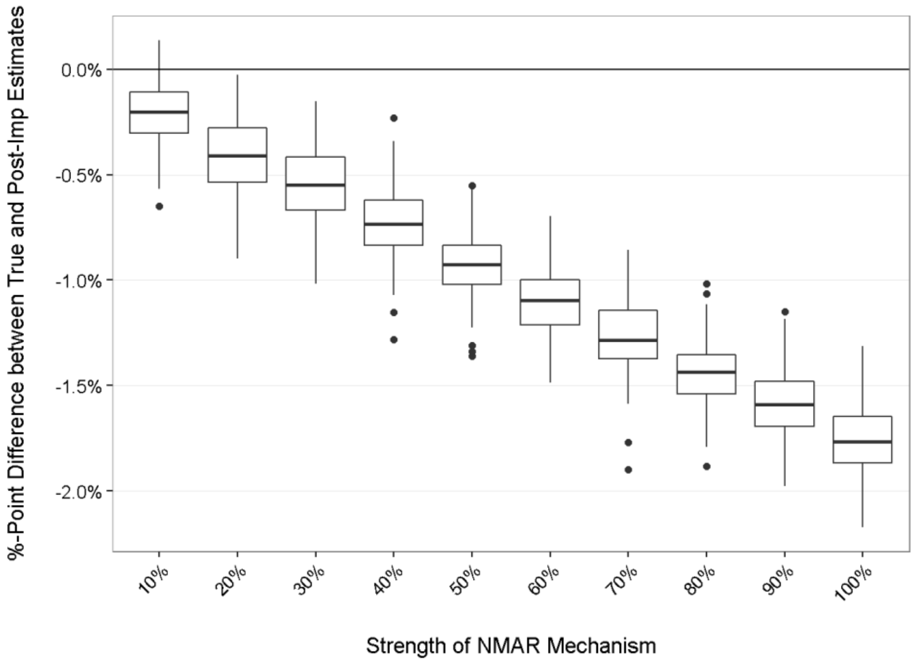

Figure 10: Sex Estimate Difference (Post-Imp. — True) Distributions across Replicates and NMAR Strengths (WSHD)

Figures 11 and 12 show both methods performing well at estimating perpetrator age, with nothing to distinguish one method from the other. All told, the bias exhibited in all three perpetrator characteristics could be problematic for a given analysis. Again, bias results are fundamentally based on the reliability of the missingness induction mechanism. Whether or not perpetrator characteristics themselves have as strong an impact on their missingness as do those of victims is an open question, though, if they do, it is clear that simply accounting for geography, year, and victim characteristics may not be sufficient to control the true underlying mechanism.

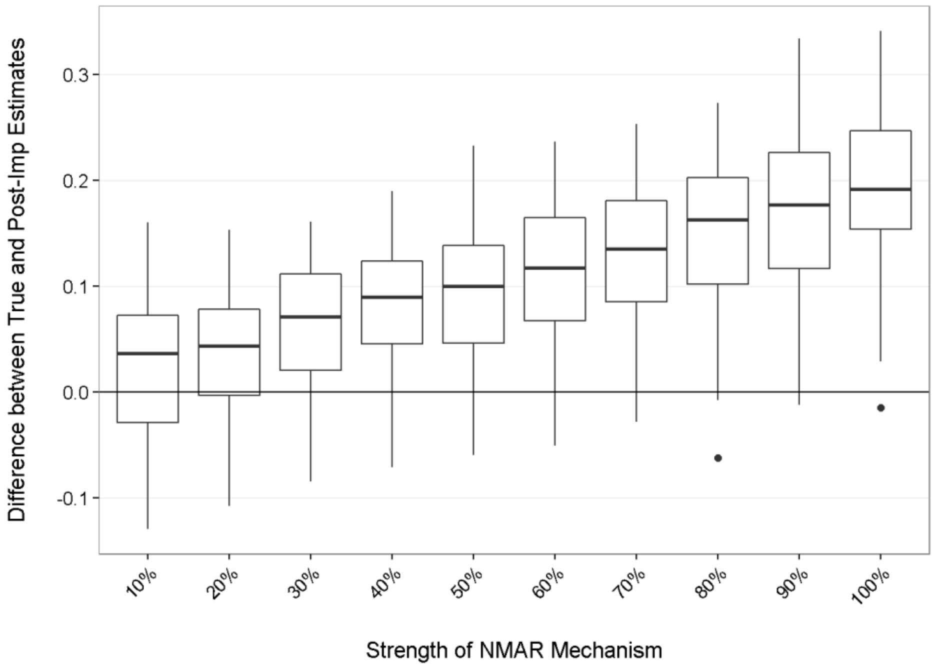

Figure 11: Age Estimate Difference (Post-Imp. — True) Distributions across Replicates and NMAR Strengths (FCS)

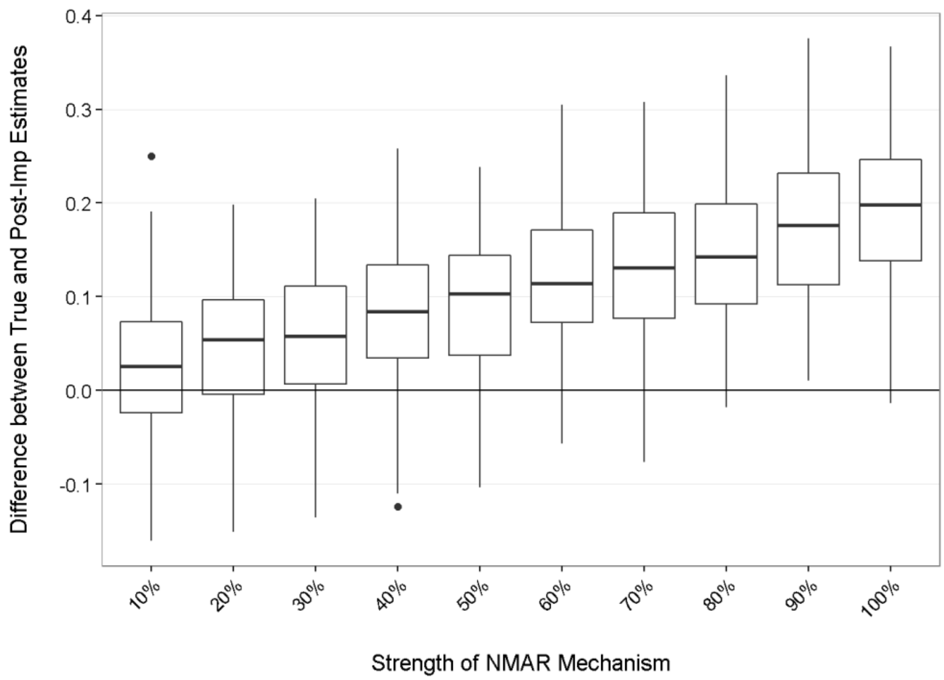

Figure 12: Age Estimate Difference (Post-Imp. — True) Distributions across Replicates and NMAR Strengths (WSHD)

4. Discussion

This report has focused on the primary issue relating to data quality in the SHR: item nonresponse for perpetrator characteristics, which is primarily driven by case-insolvency. Options for addressing item nonresponse are varied. Fox's (2004) weighting method (and by implication its hot-deck analogue) has shown to be one of the more adept methods for dealing with missing SHR data, and its simplicity has resulted in its use for past BJS analytic products. However, this and other methods are based on the assumption that the underlying missingness mechanism is conditionally random, or MAR. If this assumption does not hold, imputed perpetrator characteristics could lead to bias in point estimates and their estimated variances.

To assess the extent to which nonrandom missingness could prove problematic, and to establish whether a traditional hot-deck approach or a more modern, iterative multiple imputation approach is better equipped to combat nonrandom missingness, a Monte Carlo analysis was undertaken. In this analysis, weighted sequential hot-deck and fully-conditional specification techniques were implemented over a series of data replicates with simulated missingness induced according to a purposeful and nonrandom mechanism. This mechanism applied the effects of victim demographics on perpetrator missingness to the known values of comparable perpetrator characteristics in complete data. This mechanism was applied at full strength (as observed empirically in the 2005-2011 SHR data files) and at reduced strengths along a gradient down to the point of completely random missingness.

Using available victim, geographic, and temporal data from each incident, each method was used to fill the missing data across all replicates. The results of perpetrator characteristic estimation post-imputation showed that the candidate methods performed similarly in their abilities to converge for a given dataset, capture true values in estimated confidence intervals, and manage bias in point estimates. A slight edge was given to FCS, as its multiple imputation implementation makes possible the capture of imputation uncertainty in variance estimates, leading to somewhat better true estimate coverage. Both methods were unable to overcome a nonrandom missingness mechanism of even moderate strength – a concern for data users in the context of analyses focused on perpetrator data.

References

Regoeczi, Wendy C and Riedel, Marc. "The Application of Missing Data Estimation Models to the Problem of Unknown Victim/Offender Relationships in Homicide Cases." (2003). *Journal of Quantitative Criminology*, 19(2), 155-183

van Buuren, S and Boshuizen, Hendriek and L. Knook, D. "Multiple imputation of missing blood pressure covariates in survival analysis." (1999). *Statistics in Medicine*. 18. 681-694.

Fox, A. "Missing Data Problems in the SHR: Imputing Offender and Relationship Characteristics." (2004). *Homicide Studies.* 8:214

Wadsworth, T. and Roberts, J. M. "When Missing Data are not Missing: A New Approach to Evaluating Supplemental Homicide Report Imputation Strategies." (2008). *Criminology.* 46: 841-870.