

## Article

# A Bootstrap Method for a Multiple-Imputation Variance Estimator in Survey Sampling

Lili Yu <sup>1,\*</sup>  and Yichuan Zhao <sup>2</sup>
<sup>1</sup> Department of Biostatistics, Epidemiology and Environmental Health Sciences, JPH College of Public Health, Georgia Southern University, Statesboro, GA 30460, USA

<sup>2</sup> Department of Mathematics and Statistics, Georgia State University, Atlanta, GA 30303, USA

\* Correspondence: lyu@georgiasouthern.edu

**Abstract:** Rubin's variance estimator of the multiple imputation estimator for a domain mean is not asymptotically unbiased. Kim et al. derived the closed-form bias for Rubin's variance estimator. In addition, they proposed an asymptotically unbiased variance estimator for the multiple imputation estimator when the imputed values can be written as a linear function of the observed values. However, this needs the assumption that the covariance of the imputed values in the same imputed dataset is twice that in the different imputed datasets. In this study, we proposed a bootstrap variance estimator that does not need this assumption. Both theoretical argument and simulation studies show that it was unbiased and asymptotically valid. The new method was applied to the Hox pupil popularity data for illustration.

**Keywords:** bootstrap; congeniality; domain mean; Rubin's variance estimator; survey sampling



**Citation:** Yu, L.; Zhao, Y. A Bootstrap Method for a Multiple-Imputation Variance Estimator in Survey Sampling. *Stats* **2022**, *5*, 1231–1241. <https://doi.org/10.3390/stats5040074>

Academic Editor: Wei Zhu

Received: 28 October 2022

Accepted: 26 November 2022

Published: 29 November 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Surveys are popular tools for empirical research in many fields, such as the social sciences, marketing, and public health. One of the goals of survey data analysis is the estimation of subpopulations, which are called domains. The estimator of the domain mean  $\theta$  usually involves weights in survey data analysis [1–5], which is written as

$$\hat{\theta} = \sum_{i \in A} \alpha_i y_i \quad (1)$$

where  $A$  represents the set of indices of the sample elements,  $y_i$  are the observed sample values in  $A$ , and  $\alpha_i$  is the known weight function that is free of  $y_i$ . The  $\alpha_i$  can be used to adjust the non-response items or for calibration purposes.

Imputation is commonly used in survey data analysis to handle the non-response items in survey samples. It replaces the non-response items in the survey sample with substituted values. Single imputation leads to underestimating the variances of the estimators [6]. To overcome this, a multiple imputation method [6,7] was proposed to account for the variances in the analysis when substituted values are used instead of the true observations. The basic idea of multiple imputations is to impute each non-response value  $M$  times to create  $M$  imputed datasets. Then, each imputed dataset is analyzed using standard data analysis techniques. Finally, Rubin's rule is used to obtain Rubin's estimator and Rubin's variance estimator by combining the results from each imputed dataset.

Research on multiple imputations is motivated by a Bayesian framework. Rubin [6] claimed that multiple imputations can provide a valid frequentist inference in many applications, assuming "proper" imputation [6] or congeniality condition [8] for both the imputation model and the analysis model. However, Fay [9,10] and Binder and Sun [11] found that it is difficult to satisfy the conditions for "proper" imputation for general complex sampling schemes. In a parametric model setting, Wang and Robin [12] and Nielsen [13] showed that Rubin's variance estimator is unbiased when the complete-sample estimator is

the maximum likelihood estimator. Then, Rubin and Wang [14] showed the bias of Rubin's variance estimator for nonparametric inference procedures and proposed a new variance estimator. Yang and Kim [15] showed the bias of Rubin's variance estimator for the method of moments estimator. They also proposed a new variance estimator when the complete-sample estimator is the method of moments estimator or maximum likelihood estimator. However, these proposed variance estimators cannot be directly used when the complete-sample estimator is given as (1) for the domain mean. For estimator (1), Kim et al. [16] derived the closed-form bias of Rubin's variance estimator and proposed a new variance estimator when the imputed values can be expressed as a linear function of the response items. However, it requires an assumption that the covariance of the imputed values in the same imputed datasets is twice that of the different imputed datasets. Kim et al. [16] did not discuss how to check this assumption. Our simulation study results presented in Section 3.1 for the domain mean estimation showed that Kim's variance estimator is biased, which indicated that this assumption may not be satisfied.

In this study, we proposed a bootstrap variance estimator without this assumption. Traditional bootstrapping [17] is usually used to estimate the properties of an estimator by sampling from an approximating distribution. For independent and identically distributed sample data with  $n$  observations, it obtains bootstrap samples, each with  $n$  observations, by randomly drawing observations with replacement from the original sample. It was used for variance estimation in the presence of single imputation for missing data analysis [18–25]. Recently, researchers [26–31] applied the traditional bootstrap for multiple imputations. They applied multiple imputations for each bootstrap sample to obtain the parameter estimator. Then, the variance of the estimator is estimated by the sample variance of the estimators from the bootstrap samples. Therefore, they need to refit the model  $M \times B$  times, where  $M$  is the number of imputed samples and  $B$  is the number of bootstrap samples. Obviously, the traditional bootstrap methods are computationally intensive. In this study, we proposed a computationally efficient approach for a bootstrap procedure. We used the bootstrap samples with single imputation and multiple imputation samples to estimate the variance. Therefore, it only needed to refit the model  $M + B$  times to estimate the variance of the multiple imputation estimator with the complete-sample estimator in the form of (1).

## 2. Methods

### 2.1. Rubin's Method

We consider data generated from a superpopulation model, i.e., a random sample from an infinite population. The sampling mechanism and the response mechanism are assumed to be ignorable under the superpopulation model in the sense of Rubin [6]. Let  $y_1, \dots, y_n$  be the complete sample for the random variable  $Y$ . Without loss of generality, assume that the first  $r$  observations are observed and the remaining  $n - r$  observations are missing, i.e.,  $\mathbf{Y}_{\text{obs}} = \{y_1, \dots, y_r\}$  and  $\mathbf{Y}_{\text{mis}} = \{y_{r+1}, \dots, y_n\}$ . Denote  $R_i$  as the indicator of whether the  $i$ th observation is observed. If  $R_i = 1$ ,  $y_i$  is observed and if  $R_i = 0$ ,  $y_i$  is missing. The multiple imputation method uses the following three steps to estimate the parameter domain mean  $\theta$  for the complete-sample estimator in the form of (1).

Step 1. Create  $m = 1, \dots, M$ , imputed datasets, where the observations  $y_i^*, i = 1, \dots, n$ , are imputed only for the missing values; otherwise, they remain as observed.

Step 2. Apply the complete-sample estimation procedure to each imputed dataset. Let  $\hat{\theta}_n^m$  be the imputed-sample estimator of  $\theta$  for the  $m$ th imputed dataset and  $\hat{V}_n^m$  be the imputed-sample variance estimator of  $\hat{\theta}_n^m$ .

Step 3. Combine the results from the imputed datasets using Rubin's rule [6]. Rubin's estimator of  $\theta$  is  $\hat{\theta}_{MI} = M^{-1} \sum_{m=1}^M \hat{\theta}_n^m$ , and Rubin's variance estimator is

$$\hat{V}_{MI}(\hat{\theta}_{MI}) = W_M + (1 + M^{-1})B_M \quad (2)$$

where  $W_M = M^{-1} \sum_{m=1}^M \hat{V}_n^m$  is the within-imputation variance and  $B_M = (M - 1)^{-1} \sum_{m=1}^M (\hat{\theta}_n^m - \hat{\theta}_{MI})^2$  is the between-imputation variance.

Rubin's variance estimator is based on the decomposition

$$\text{var}(\hat{\theta}_{MI}) = \text{var}(\hat{\theta}_n) + \text{var}(\hat{\theta}_{MI} - \hat{\theta}_n) + 2\text{cov}(\hat{\theta}_{MI} - \hat{\theta}_n, \hat{\theta}_n) \quad (3)$$

where  $\hat{\theta}_n$  is the complete-sample estimator of  $\theta$ . In Rubin's variance estimator (2),  $W_M$  estimates the first term of (3) and  $(1 + M^{-1})B_M$  estimates the second term of (3). Kim et al. [16] showed that for estimators in the form of (1),  $(1 + M^{-1})B_M$  is an asymptotically unbiased estimator of the second term of (3) based on the covariance assumption, that is, the covariance between  $y_i^{*m_1}$  and  $y_j^{*m_1}$  is twice that of the covariance between  $y_i^{*m_1}$  and  $y_j^{*m_2}$  for  $i \neq j$  and  $m_1 \neq m_2$ . Therefore, if  $W_M$  is an unbiased estimator of the first term of (3), the asymptotic bias of Rubin's variance estimator is

$$\text{bias}(\hat{V}_{MI}) = -2\text{cov}(\hat{\theta}_{MI} - \hat{\theta}_n, \hat{\theta}_n). \quad (4)$$

Kim et al. [16] derived the closed-form bias of (4) for the estimators in the form of (1) when the imputed values can be expressed as a linear function of the observed items. We can see that Kim's method [16] requires the covariance assumption. In this study, we proposed a bootstrap variance estimator for  $\hat{\theta}_{MI}$  that does not need this assumption, and hence, this new method has a much wider range of applications.

## 2.2. The Bootstrap Estimation Method

### 2.2.1. Assumptions

We employed the regularity assumptions for the data and the imputation mechanism in Kim et al. [16] for the newly proposed method.

(C.1) Both the complete-sample point estimator and variance estimator are asymptotically unbiased, i.e.,

$E(\hat{\theta}_n) = \theta + O(n^{-1})$  and  $E(\hat{V}_n) = \text{var}(\hat{\theta}_n) + O(n^{-1})$ , where  $\hat{V}_n$  is the complete-sample variance estimator of  $\hat{\theta}_n$ .

(C.2) The conditional expectations of the imputed values are the same as those of the real values, i.e.,

$E(y_i^{*m} | A, A_R) = E(y_i | A, A_R)$ ,  $m = 1, \dots, M$ , where  $A_R$  is the set of indices of the observed items.

(C.3) The conditional variances of the imputed values are asymptotically the same as those of the real values, i.e.,

$$\max_{i,j} \left| \text{cov}(y_i^{*m}, y_j^{*m} | A, A_R) - \text{cov}(y_i, y_j | A, A_R) \right| = o(1), \quad m = 1, \dots, M.$$

(C.4) Conditional on  $A$ ,  $A_R$  and  $Y_{obs}$ , the imputed values  $y_j^{*m}$  of  $y_j$ ,  $j \in A_k$  are conditionally independent and identically distributed, where  $A_k$  is the set of indices of the missing items.

Assumptions (C.1) and (C.2) ensure that the imputed-sample estimators of the form (1) are approximately unbiased. Assumption (C.3) ensures that the covariances of the imputed values are asymptotically the same as those of the original values. Therefore, we can treat the imputed values similarly to the original values in the inference procedure. Assumption (C.4) ensures both the estimators  $\hat{\theta}_n^m$ ,  $m = 1, \dots, M$ , and the naive variance estimators  $\hat{V}_n^m$ ,  $m = 1, \dots, M$ , are identically distributed. Therefore, the imputed values and the imputed samples can be treated equally in the inference procedure.

### 2.2.2. Variance Decomposition

Instead of the decomposition (3), we decomposed the variance as

$$\text{var}(\hat{\theta}_{MI}) = V_1 + V_2 + V_3 + V_4, \quad (5)$$

where  $V_1 = \text{var}(\sum_{i \in A_R} \alpha_i y_i)$ ,  $V_2 = M^{-1} E\{\text{var}(\sum_{j \in A_k} \alpha_j y_j^{*m} | \mathbf{Y}_{obs})\}$ ,  $V_3 = \text{var}\{E(\sum_{j \in A_k} \alpha_j y_j^{*m} | \mathbf{Y}_{obs})\}$ , and  $V_4 = 2\text{cov}(\sum_{j \in A_k} \alpha_j y_j^{*m}, \sum_{i \in A_R} \alpha_i y_i)$ .  $V_1$  is the sum of the variances for the observed observations. The sum of  $V_2$  and  $V_3$  is the sum of the variances for the imputed observations based on the law of total variance.  $V_4$  is the sum of the covariances of the observed observations and the imputed observations.

### 2.2.3. Bootstrap Variance Estimator and Its Properties

To estimate (5), we generated bootstrap data as follows. For  $b = 1, \dots, B$ :

Step 1: Take a random sample of size  $n$  with a replacement from the original sample, which consists of independent and identical observations. Denote the bootstrap sample as  $(y_1^b, R_1^b), \dots, (y_n^b, R_n^b)$ . Denote the observed items as  $\mathbf{Y}_{obs}^b = \{y_1^b, \dots, y_{r'}^b\}$  and the missing items as  $\mathbf{Y}_{mis}^b = \{y_{r'+1}^b, \dots, y_n^b\}$  in the bootstrap method. Let  $A^b$ ,  $A_R^b$ , and  $A_k^b$  represent the set of indices of the sample elements, the set of indices of the observed items, and the set of indices of the missing items in the  $b$ th bootstrap sample, respectively.

Step 2: With the resultant bootstrap sample, calculate the imputed values  $y_j^{*b}$  for missing items  $y_j^b$  using a single imputation method, which is the same as the multiple imputation method with  $M = 1$ . The imputed-sample estimator in the  $b$ th bootstrap sample is  $\hat{\theta}^b = \sum_{i \in A^b} \alpha_i^b y_i^{*b}$ .

We estimated  $V_1$  and  $V_2$  using multiple imputed samples and estimated  $V_3$  and  $V_4$  using bootstrap samples.  $V_1$  can be estimated from the imputed-sample output based on condition (C.1), i.e.,  $\hat{V}_1 = r(nM)^{-1} \sum_{m=1}^M \hat{V}_n^m$ .  $V_2$  is estimated using the empirical conditional variance, which is  $\hat{V}_2 = M^{-1}(M-1)^{-1} \sum_{m=1}^M (\sum_{j \in A_k} \alpha_j y_j^{*m} - H_1)^2$ , where  $H_1 = M^{-1} \sum_{m=1}^M \sum_{j \in A_k} \alpha_j y_j^{*m}$ .  $V_3$  is estimated using  $\hat{V}_3 = (B-1)^{-1} \sum_{b=1}^B (\sum_{j \in A_k} \alpha_j^b \eta_j^b - H_2)^2$ , where  $\eta_j^b \equiv E(y_j^{*b} | \mathbf{Y}_{obs}^b)$  and  $H_2 = B^{-1} \sum_{b=1}^B \sum_{j \in A_k} \alpha_j^b \eta_j^b$ .  $V_4$  is estimated using  $\hat{V}_4 = 2 \sum_{j \in A_k} \alpha_j \sum_{i \in A_R} \alpha_i (B-1)^{-1} \sum_{b=1}^B (\bar{\eta}_j^b - H_3)(\bar{Y}_{r'}^b - H_4)$ , where  $\bar{\eta}_j^b = (n-r')^{-1} \sum_{i \in A_k^b} \eta_j^b$ ,  $\bar{Y}_{r'}^b = (r')^{-1} \sum_{i \in A_R^b} y_i^b$ ,  $H_3 = B^{-1} \sum_{b=1}^B \bar{\eta}_j^b$ , and  $H_4 = B^{-1} \sum_{b=1}^B \bar{Y}_{r'}^b$ . Therefore, the bootstrap variance estimator is  $\hat{V} = \hat{V}_1 + \hat{V}_2 + \hat{V}_3 + \hat{V}_4$ .

The newly proposed bootstrap variance estimator is an asymptotically unbiased estimator for the variance of the multiple imputation estimator. Under the conditions (C.1)–(C.4),  $\hat{V}_1$  is an asymptotically unbiased estimator of  $V_1$ .  $\hat{V}_2$  is an asymptotically unbiased estimator because  $E(\hat{V}_2) = E\{E(\hat{V}_2 | \mathbf{Y}_{obs})\}$ .  $\hat{V}_3$  is an asymptotically unbiased estimator of  $V_3$  because it is based on the Glivenko–Cantelli Theorem:

$$E(\hat{V}_3) = E[E\{(\sum_{j \in A_k} \alpha_j^b \eta_j^b)^2 | \mathbf{Y}_{obs}\} - \{E(\sum_{j \in A_k} \alpha_j^b \eta_j^b | \mathbf{Y}_{obs})\}^2],$$

where  $\eta_j^b = E(y_j^{*b} | \mathbf{Y}_{obs})$ .

For  $\hat{V}_4$ , we derived

$$E\{\text{cov}(\bar{Y}_{r'}^b, \bar{\eta}_j^b | A, A_R, \mathbf{Y}_{obs})\} = \text{cov}(y_i, y_j^{*m} | A, A_R) + O(\frac{1}{r}),$$

and

$$\begin{aligned} & E(\frac{1}{B-1} \sum_{b=1}^B (\bar{\eta}_j^b - H_3)(\bar{Y}_{r'}^b - H_4)) \\ &= E\{\frac{1}{B-1} E(\sum_{b=1}^B ((\bar{\eta}_j^b - \mu_1) + (\mu_1 - H_3))((\bar{Y}_{r'}^b - \mu_2) + (\mu_2 - H_4)) | \mathbf{Y}_{obs})\} \\ &= E\{\text{cov}(\bar{\eta}_j^b, \bar{Y}_{r'}^b | \mathbf{Y}_{obs})\}, \end{aligned}$$

where  $\mu_1 \equiv E(\bar{\eta}_j^b | \mathbf{Y}_{obs})$  and  $\mu_2 \equiv E(\bar{Y}_{r'}^b | \mathbf{Y}_{obs})$ . Thus,  $\hat{V}_4$  is an asymptotically unbiased estimator of  $V_4$ .

To account for the uncertainty in the variance estimator with a small-to-moderate imputation size and bootstrap replicates, a  $t$ -distribution is used to calculate  $100(1 - \alpha)\%$

confidence intervals for  $\theta$ , i.e.,  $\hat{\theta}_{MI} \pm t_{v, 1-\alpha/2} \hat{V}^{1/2}$ . The degrees of freedom  $v$  is an approximate number along the lines of the Satterthwaite [32] method, which is derived as

$$v = \frac{\hat{V}^2}{(M-1)^{-1} \hat{V}_2^2 + (B-1)^{-1} [\hat{V}_3^2 + \hat{V}_5^2 + \hat{V}_6^2 + \hat{V}_7^2 + 2(1-\pi^2) \hat{V}_5(\hat{V}_3 - \hat{V}_7) - 2\hat{V}_3\hat{V}_7 - 2\hat{V}_5\hat{V}_6\pi^2]},$$

in which  $\pi$  is the response rate of the interested domain,  $\hat{V}_5 = (B-1)^{-1} \sum_{b=1}^B (H_5 + H_6)^2$ ,  $\hat{V}_6 = (B-1)^{-1} \sum_{b=1}^B H_5^2$ , and  $\hat{V}_7 = (B-1)^{-1} \sum_{b=1}^B H_6^2$ , where  $H_5 = \sum_{i \in A_R} \alpha_i (\bar{Y}_{r'}^b - H_4)$  and  $H_6 = \sum_{j \in A_k} \alpha_j (\bar{\eta}_j^b - H_3)$ . The detailed derivation of  $v$  is provided in the Appendix A.

### 3. Examples and Results

Simulations and a real data analysis were conducted to investigate the performance of the proposed bootstrap variance estimator ( $\hat{V}$ ). We compared it with Rubin's variance estimator ( $\hat{V}_{MI}$ ), Kim's method ( $\hat{V}_k$ ) [16], and the traditional bootstrapping method ( $\hat{V}_B$ ). As in Yang and Kim [15], we report the relative bias (Rbias) of the variance estimators, mean width (mwidth), and coverage probabilities of the 95% confidence intervals (95cov) for  $\theta$ . The relative bias was calculated using  $(E(\tilde{V}) - \text{var}(\hat{\theta}_{MI})) / \text{var}(\hat{\theta}_{MI}) \times 100\%$ , where  $\tilde{V}$  is a variance estimator. The 95% confidence intervals were calculated based on a  $t$ -distribution with  $v$  degrees of freedom. We considered different values of  $B$ ,  $M$ , and  $n$  to investigate their effects. These results were based on 5000 Monte Carlo runs for each setting. The R code used for these simulations is available from the first author upon request.

#### 3.1. Simulation 1: Domain Mean Estimation

We simulated a sample  $y_1, \dots, y_n$  from  $N(\mu, \sigma^2)$ , where  $\mu = 2$ ,  $\sigma^2 = 4$ , and  $n = 500$  or  $1000$ . Then non-respondents were generated based on the response rate  $\pi$  and using the *ampute* function in the *mice* package of R. We considered two response rates, which were 0.8 and 0.5. To indicate whether an observation belonged to a domain  $Z$ , we generated  $z_i$  from Bernoulli distribution with probability  $d$ , which was 0.2 or 0.6 in the simulation study. If  $z_i = 1$ , this meant that  $y_i$  was in domain  $Z$  and was 0 otherwise. The  $z_i$ ,  $i = 1, \dots, n$ , were always observed.

Now we would like to estimate the mean  $\theta$  of the domain  $Z$ . Without loss of generality, denote the first  $r$  observations as respondents ( $y_1, \dots, y_r$ ) and the remaining  $(n-r)$  observations as the non-respondents ( $y_{r+1}, \dots, y_n$ ). Then, the complete-sample estimator of  $\theta$  is

$$\hat{\theta}_n = \sum_{i=1}^n \frac{z_i}{\sum_{i=1}^n z_i} y_i.$$

Let  $n_d = \sum_{i=1}^n z_i$  be the size of the domain  $Z$  and  $r_d = \sum_{i=1}^r z_i$  be the number of response items in the domain  $Z$ . It is obvious that  $\alpha_i = z_i / n_d$ .

To handle the missing observations, we used the multiple imputation method with  $M = 10$  or  $30$ . The  $m$ th imputed-sample estimator is

$$\hat{\theta}_n^m = \sum_{i=1}^n \frac{z_i}{\sum_{i=1}^n z_i} y_i^{*m}.$$

If  $y_i$  is missing, the conditional expectation of the imputed value is  $\eta_i \equiv E(y_i^{*m} | y_1, \dots, y_r) = \bar{y}_r$  and  $\bar{y}_r = \sum_{i=1}^r y_i / r$ .

Now we used the newly proposed bootstrap method with  $B = 200$  or  $500$  in Section 3 to estimate the variance of the multiple imputation estimator. Denote the bootstrap sample as  $\{y_1^b, \dots, y_n^b\}$  in which the response items are  $\mathbf{Y}_{obs}^b = \{y_1^b, \dots, y_{r'}^b\}$ , the non-response items are  $\mathbf{Y}_{mis}^b = \{y_{r'+1}^b, \dots, y_n^b\}$ , and the corresponding indicators for domain  $Z$  are  $z_i^b$ ,  $i = 1, \dots, n$ . Let  $n_d^b$  be the size of the domain  $Z$  in the  $b$ th bootstrap sample. Then, we calculate

$$\begin{aligned}\hat{V}_1 &= \frac{r_d}{n_d^2(n-1)M} \sum_{m=1}^M \sum_{i=1}^n (y_i^{*m} - \bar{y}^m)^2, \text{ where } \bar{y}^m = \frac{1}{n} \sum_{i=1}^n y_i^{*m}; \\ \hat{V}_2 &= \frac{1}{M(M-1)} \sum_{m=1}^M (\sum_{j=r+1}^n \frac{z_j}{n_d} y_j^{*m} - H_1)^2, \text{ where } H_1 = \frac{1}{M} \sum_{m=1}^M \sum_{j=r+1}^n \frac{z_j}{n_d} y_j^{*m}; \\ \hat{V}_3 &= \frac{1}{B-1} \sum_{b=1}^B (\sum_{j=r'+1}^n \frac{z_j^b}{n_d} \bar{Y}_{r'}^b - H_2)^2, \text{ where } H_2 = \frac{1}{B} \sum_{b=1}^B \sum_{j=r'+1}^n \frac{z_j^b}{n_d} \bar{Y}_{r'}^b \text{ and } \\ \bar{Y}_{r'}^b &= \frac{\sum_{i=1}^{r'} y_i^b}{r'}; \\ \hat{V}_4 &= 2 \sum_{j=r+1}^n \frac{z_j}{n_d} \sum_{i=1}^r \frac{z_i}{n_d} (B-1)^{-1} \sum_{b=1}^B (\bar{Y}_{r'}^b - H_3)(\bar{Y}_z^b - H_4), \text{ where } \bar{Y}_z^b = \frac{\sum_{i=1}^{r'} z_i^b y_i^b}{\sum_{i=1}^{r'} z_i^b}, \\ H_3 &= B^{-1} \sum_{b=1}^B \bar{Y}_{r'}^b \text{ and } H_4 = B^{-1} \sum_{b=1}^B \bar{Y}_z^b.\end{aligned}$$

To compare the newly proposed method with the method proposed by Kim et al. [16], we used the following formula derived by them [16]:

$$\hat{V}_k = W_M + (1 + M^{-1})B_M + 2n_d^{-2}m_d(1 - r^{-1}r_d)\hat{\sigma}^2, \text{ where } m_d = n_d - r_d \text{ and } \hat{\sigma}^2 = (n-1)^{-1}M^{-1} \sum_{m=1}^M \sum_{i=1}^n (y_i^{*m} - \bar{y}^m)^2.$$

The simulation results are summarized in Table 1. They show that the newly proposed method  $\hat{V}$  had a smaller relative bias than both Rubin's method  $\hat{V}_{MI}$  and Kim's method  $\hat{V}_k$ , especially for cases with a small response rate. We notice that the relative bias was larger for the proposed method when the sample size became larger. This may have been due to the smaller value of  $var(\hat{\theta}_{MI})$  for the larger sample size. Other methods have this pattern as well. The width of the confidence interval of Rubin's method was usually larger, while that of Kim's method was usually smaller than that of the newly proposed method. The coverage rate of the newly proposed method was close to the nominal level. However, Rubin's method had a larger coverage rate and Kim's method had a smaller coverage rate than the nominal level. These results indicated that the covariance assumption required in Kim's method may not be satisfied, and hence,  $(1 + M^{-1})B_M$  may not unbiasedly estimate  $var(\hat{\theta}_{MI} - \hat{\theta}_n)$  for the domain mean estimation. The newly proposed method  $\hat{V}$  had a similar performance to the traditional bootstrap method  $\hat{V}_B$ . However, it was much faster than the traditional bootstrap method  $\hat{V}_B$ , especially for cases with large  $n$ ,  $B$ , and  $M$ . For example, for a case with  $n = 1000$ ,  $B = 500$ ,  $M = 30$ , and 5000 Monte Carlo runs, the new method needed about half an hour, while the traditional bootstrap method needed about four hours. In addition, the values of  $B$  considered in these examples had no significant effects on these methods. Larger values of  $n$  and  $M$  resulted in smaller widths of the confidence intervals.

### 3.2. Simulation 2: Linear Regression

The variance estimation method proposed in this study can not only be applied to the domain mean estimation but also to any other estimator that has the form of (1). In this example, we showed its performance for the mean estimation of the response variable in the linear regression model. We considered the model used in Yang and Kim [15] as follows:

$$y_i = \beta x_i + e_i, \quad i = 1, \dots, n,$$

Where  $n = 500$  or  $1000$ ,  $\beta = 0.1$ ,  $x_i \sim \exp(1)$ , and  $e_i \sim N(0, \sigma_e^2)$  with  $\sigma_e^2 = 0.25$ . In this data setup, we assumed that  $x_i$  were observed fully. The missingness in  $y_i$  was controlled by  $\delta_i$ , which followed  $\delta_i \sim \text{Ber}(p_i)$  and  $p_i = 1/\{\exp(-\phi_0 - \phi_1 x_i)\}$ . Two scenarios were considered, namely,  $(\phi_0, \phi_1) = (-1.5, 2)$  and  $(\phi_0, \phi_1) = (3, -3)$ , which yielded the approximate average response rate of 0.5. We were interested in estimating  $\theta = E(y_i)$ . The normal imputation procedure [33] was used to generate the imputed values for the missing observations. Let  $\mathbf{Y}_{obs} = \{y_1, \dots, y_r\}$  be the observed observations and  $\mathbf{Y}_{mis} = \{y_{r+1}, \dots, y_n\}$  be the missing observations. In the simulations, we set  $B = 200$  or  $500$  and  $M = 10$  or  $30$ . Then, the complete-sample estimator was  $\hat{\theta}_n = \sum_{i=1}^n y_i / n$ .



**Table 1.** Simulation results for the domain mean estimation.

<i>n</i>	<i>B</i>	<i>M</i>	$\pi$	<i>d</i>	Rbias (%)				Mwidth ( $\times 10^2$ )				95cov ( $\times 10^2$ )			
					$\hat{V}$	$\hat{V}_{MI}$	$\hat{V}_K$	$\hat{V}_B$	$\hat{V}$	$\hat{V}_{MI}$	$\hat{V}_K$	$\hat{V}_B$	$\hat{V}$	$\hat{V}_{MI}$	$\hat{V}_K$	$\hat{V}_B$
500	500	10	0.8	0.2	0.6	146.2	−12.5	0.5	65	102	58	65	95	99	89	95
					0.6	1.6	21.6	−4.9	1.5	48	53	46	48	95	97	95
					0.5	0.2	−1.3	237.7	−41.1	0.7	61	113	36	61	95	99
		30	0.8	0.2	0.6	−0.7	18.1	−40.5	−0.6	55	61	42	56	95	96	85
					0.6	1.0	156.8	−12.6	0.9	63	100	57	62	95	99	92
					0.6	1.3	21.6	−5.3	1.3	47	52	46	47	95	97	95
		10	0.5	0.2	0.6	−2.2	261.4	−54.3	0.6	58	109	33	57	95	100	63
					0.6	−1.4	17.8	−42.2	−1.4	54	59	41	54	95	97	86
					0.6	−1.3	−42.2	17.8	−1.3	54	59	41	54	95	97	86
	200	10	0.8	0.2	0.5	146.2	−12.5	0.2	65	102	58	65	95	99	89	95
					0.6	1.6	21.6	−4.9	1.5	48	53	46	48	95	97	95
					0.5	0.2	−1.4	237.7	−41.1	0.7	62	113	36	61	95	99
		30	0.8	0.2	0.6	−0.8	18.1	−40.5	−0.7	55	61	42	56	95	96	85
					0.6	1.1	156.8	−12.6	1.0	63	100	57	62	95	99	92
					0.6	1.3	21.6	−5.3	1.3	47	52	46	47	95	97	95
		10	0.5	0.2	0.6	−2.2	261.4	−54.3	0.6	61	108	33	57	96	100	63
					0.6	−1.3	−42.2	17.8	−1.3	54	59	41	54	95	97	86
					0.6	−1.3	−42.2	17.8	−1.3	54	59	41	54	95	97	86
1000	500	10	0.8	0.2	−1.1	141.4	13.1	−1.1	46	72	42	46	95	99	89	95
					0.6	−4.0	53.4	−9.7	−4.0	34	37	33	34	95	97	94
					0.5	0.2	−3.7	255.4	−44.3	−3.2	43	80	25	44	95	99
		30	0.8	0.2	0.6	−0.3	19.5	−39.2	−0.3	39	44	30	40	95	97	85
					0.6	−1.5	148.6	−14.6	−1.6	44	71	41	44	95	99	92
					0.6	−3.7	15.7	−9.8	−3.6	33	37	32	34	95	96	94
		10	0.5	0.2	0.6	−3.9	251.2	−55.3	−3.3	41	77	23	40	94	99	63
					0.6	−0.1	20.0	−40.8	−0.1	38	42	29	39	95	97	85
					0.6	−0.1	20.0	−40.8	−0.1	38	42	29	39	95	97	85
	200	10	0.8	0.2	−1.0	141.4	−13.1	−1.0	46	72	42	46	95	99	89	95
					0.6	−3.9	15.3	−9.7	−4.0	34	37	33	34	94	97	94
					0.5	0.2	−3.8	225.4	−44.3	−3.3	44	80	25	44	96	99
		30	0.8	0.2	0.6	−0.3	19.5	−39.2	−0.6	39	44	30	40	95	97	85
					0.6	−1.6	148.6	14.6	−1.5	45	71	41	44	95	99	92
					0.6	−3.7	15.7	−9.8	−3.6	33	37	32	34	95	96	94
		10	0.5	0.2	0.6	−4.0	251.2	−55.3	−3.3	42	77	23	40	95	99	63
					0.6	−0.1	20.0	−40.8	−0.1	38	42	29	39	95	97	85
					0.6	−0.1	20.0	−40.8	−0.1	38	42	29	39	95	97	85

Therefore,  $\alpha_i = 1/n$ . It is a special case of the estimator with form (1), that is,  $\alpha_i$ ,  $i = 1, \dots, n$ , were all same. The  $m$ th imputed-sample estimator is  $\hat{\theta}_n^m = \sum_{i=1}^n y_i^{*m}/n$ .

If  $y_i$  was missing, the conditional expectation of the imputed value was  $\eta_i \equiv E(y_i^{*m} | y_1, \dots, y_r) = x_i(X_r^T X_r)^{-1} X_r^T Y_{obs}$  for  $m = 1, \dots, M$ , where  $X_r$  was a  $r \times 2$  matrix with rows  $X_i = [1 \ x_i]$ ,  $i = 1, \dots, r$ . To estimate the variance of the multiple imputation estimator, we used the bootstrap procedure described in Section 3. Denote the bootstrap sample as  $\{y_1^b, \dots, y_n^b\}$ , the response items as  $Y_{obs}^b = \{y_1^b, \dots, y_{r'}^b\}$  and the non-response items as  $Y_{mis}^b = \{y_{r'+1}^b, \dots, y_n^b\}$  in the bootstrap sample. The corresponding independent variable for  $y_i^b$  is  $x_i^b$ . Then, we calculated

$$\begin{aligned} \hat{V}_1 &= \frac{r}{n^2(n-1)M} \sum_{m=1}^M \sum_{i=1}^n (y_i^{*m} - \bar{y}^m)^2, \text{ where } \bar{y}^m = n^{-1} \sum_{i=1}^n y_i^{*m}; \\ \hat{V}_2 &= \frac{1}{M(M-1)} \sum_{m=1}^M (\sum_{j=r+1}^n \frac{1}{n} y_j^{*m} - H_1)^2, \text{ where } H_1 = \frac{1}{M} \sum_{m=1}^M \sum_{j=r+1}^n \frac{1}{n} y_j^{*m}; \\ \hat{V}_3 &= \frac{1}{B-1} \sum_{b=1}^B (\sum_{j=r'+1}^n \frac{1}{n} \eta_j^b - H_2)^2, \text{ where } H_2 = \frac{1}{B} \sum_{b=1}^B \sum_{j=r'+1}^n \frac{1}{n} \eta_j^b \text{ and} \\ \eta_j^b &= x_j^b (X_r^{bT} X_r^b)^{-1} X_r^{bT} Y_{obs}^b; \\ \hat{V}_4 &= 2 \sum_{j=r'+1}^n \frac{1}{n} \sum_{i=1}^r \frac{1}{n} (B-1)^{-1} \sum_{b=1}^B (\bar{\eta}_j^b - H_3)(\bar{Y}_{r'}^b - H_4), \text{ where } \bar{\eta}_j^b = (n-r')^{-1} \\ &\sum_{i=r'+1}^n \eta_j^b, \bar{Y}_{r'}^b = (r')^{-1} \sum_{i=1}^{r'} y_i^b, H_3 = B^{-1} \sum_{b=1}^B \bar{\eta}_j^b, \text{ and } H_4 = B^{-1} \sum_{b=1}^B \bar{Y}_{r'}^b. \end{aligned}$$

We derived the variance estimator based on Kim's method [16]:  $\hat{V}_k = W_M + (1 + M^{-1})B_M + \frac{2}{n^2}\mathbf{1}^T \mathbf{X}_{n-r}^T (\mathbf{X}_r^T \mathbf{X}_r)^{-1} \mathbf{S}_r \hat{\sigma}^2$ , where  $\mathbf{1}$  is a vector with elements 1 of length  $(n - r)$ ;  $\mathbf{X}_{n-r}$  is an  $(n - r) \times 2$  matrix with rows  $\mathbf{X}_j = [1 \ x_j]$ ,  $j = r + 1, \dots, n$ ;  $\mathbf{S}_r = [r \ \sum_{i=1}^r x_i]$  is a  $2 \times 1$  matrix; and  $\hat{\sigma}^2 = (n - 1)^{-1} M^{-1} \sum_{m=1}^M \sum_{i=1}^n (y_i^{*m} - \bar{y}^m)^2$ .

The simulation results are shown in Table 2. We noticed that the biases of both Rubin's method and Kim's method were close to zero and the coverage rates were close to the nominal level. This indicated that the covariance assumption was satisfied and the bias term (4) was close to zero. The newly proposed method and the traditional bootstrap method had a similar performance as Rubin's method and Kim's method. This showed that they performed as well as the other two methods when their assumptions were satisfied. The newly proposed method was much faster than the traditional bootstrapping method. For a case with  $n = 1000$ ,  $B = 500$ ,  $M = 30$ , and 5000 Monte Carlo runs, the new method needed about half an hour, while the traditional bootstrap method needed about eight hours. In addition, same as in example 1, the  $B$  considered in example 2 had no significant effects, while larger  $n$  and  $M$  values led to smaller widths of the confidence intervals.

**Table 2.** Simulation results for the linear regression mean estimation.

Scen	n	B	M	Rbias (%)				Mwidth ( $\times 10^2$ )				95cov ( $\times 10^2$ )			
				$\hat{V}$	$\hat{V}_{MI}$	$\hat{V}_K$	$\hat{V}_B$	$\hat{V}$	$\hat{V}_{MI}$	$\hat{V}_K$	$\hat{V}_B$	$\hat{V}$	$\hat{V}_{MI}$	$\hat{V}_K$	$\hat{V}_B$
1	500	500	10	0.7	0.4	0.4	0.1	21	22	22	22	95	95	95	96
			30	0.8	0.2	0.3	0.1	20	20	20	20	95	95	95	95
		200	10	0.6	0.4	0.4	−0.1	21	22	22	22	95	95	95	96
			30	0.7	0.2	0.3	0.0	20	20	20	20	95	95	95	95
	1000	500	10	−1.8	−2.4	−2.3	−2.3	14	15	15	16	94	95	95	96
			30	−2.4	−2.9	−2.9	−2.9	14	14	14	14	94	95	95	95
		200	10	−2.0	−2.4	−2.4	−2.4	14	15	15	16	94	95	95	96
			30	−2.4	−2.9	−2.9	−2.8	14	14	14	14	94	95	95	95
2	500	500	10	−6.0	−5.0	−5.0	−5.0	13	14	14	14	94	94	94	95
			30	−5.0	−4.0	−4.0	−4.1	13	13	13	13	94	95	95	95
		200	10	−6.0	−5.0	−5.0	−5.0	13	14	14	14	94	94	94	95
			30	−5.0	−4.0	−4.0	−4.0	13	13	13	13	94	95	95	95
	1000	500	10	−1.0	0.4	0.4	0.0	9	10	10	10	95	95	95	96
			30	−0.6	0.8	0.8	0.6	9	9	9	9	95	95	95	95
		200	10	−1.1	0.4	0.4	−0.1	9	10	10	10	95	95	95	96
			30	−0.8	0.8	0.8	0.7	9	9	9	9	95	95	95	95

### 3.3. A Real Data Analysis

In order to illustrate the proposed method, we applied the newly proposed method to the Hox pupil popularity data [34]. The dataset collects information about pupil popularity, sex, school, and teacher experience. Pupil popularity is measured using a self-rating scale that ranges from 0 (very unpopular) to 10 (very popular). The sex denotes the gender of the students, which is boy or girl. The school variable indicates which schools the students are from. It includes 100 schools. The teacher experience records the teacher's teaching experience in years. There are 2000 observations in the dataset.

Before doing any analysis on the dataset, we usually show descriptive statistics first. We were interested in the average pupil popularity for the domain third school. We selected the first 50 schools. Therefore, the data set had 998 observations. The values for the variable school are always observed. For the variable pupil popularity, there were 405 missing values. For the third school, there were 18 observations for pupil popularity and six were missing among them. Therefore, the domain proportion was 1.8% and the total response rate was approximately 60%. Then, the method used in example 1 in Section 4 was used to calculate the average pupil popularity of the third school. We used  $M = 30$  and  $B = 500$ . Rubin's point estimator was 6.41. The four methods used in the simulation study were



used to estimate the variance of Rubin's point estimator. The bootstrap variance estimator  $\hat{V}$  was 0.059, with the corresponding 95% confidence interval of (5.935, 6.887). Rubin's variance estimator  $\hat{V}_{MI}$  was 0.109, with the corresponding 95% confidence interval of (5.762, 7.060). Kim's variance estimator  $\hat{V}_k$  was 0.053, with the corresponding 95% confidence interval of (5.961, 6.862). The traditional bootstrap variance estimator  $\hat{V}_B$  was 0.066, with the corresponding 95% confidence interval of (5.932, 6.946). The results showed that the new proposed bootstrap variance estimator results were close to the traditional bootstrap variance estimator results. We observed that Rubin's variance estimator  $\hat{V}_{MI}$  was larger than the bootstrap variance estimator  $\hat{V}$  and Kim's variance estimator  $\hat{V}_k$  was smaller than the bootstrap variance estimator  $\hat{V}$ . The results from the simulation studies indicated that the covariance assumption may not have been satisfied. In addition, Rubin's variance estimator may have resulted in a larger coverage probability and Kim's variance estimator may have resulted in a lower coverage probability of the confidence interval than the nominal level.

#### 4. Discussion and Conclusions

In this study, we proposed a new bootstrap variance estimator for a multiple imputation estimator with the complete-sample estimator in the form of (1). It does not need the covariance assumption used in Kim's method, and thus, it has a wider range of applications. The simulations showed that it performed similarly to the traditional bootstrap method but greatly saved computational time. Moreover, the new method performed well for both situations with the covariance assumption satisfied or not. In addition, Kim's method can only be applied to a situation where the imputed values can be written as a linear function of the observed values. However, the new method has no such restriction, and it can be applied to the multiple imputation estimator with any form of imputed values. Further research can be conducted to investigate its performance in such situations.

This newly proposed method can be extended to a more general class of estimators with the form  $\sum_{i \in A} g(y_i)$ , such as a maximum likelihood estimator, where the proportion of  $y$  is less than any constant and the quantile of  $y$ . Future research can be conducted to investigate the performance when using these estimators.

**Author Contributions:** Conceptualization, L.Y. and Y.Z.; methodology, L.Y. and Y.Z.; software, L.Y.; validation, L.Y. and Y.Z.; writing—original draft preparation, L.Y.; writing—review and editing, L.Y. and Y.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data used in this article can be found in reference [34].

**Conflicts of Interest:** The authors declare no conflict of interest.

#### Appendix A

Proof of the approximate number of degrees of freedom:

As in Rubin and Schenker [35], we derived the degrees of freedom of the  $t$ -distribution using (i) the conditional on the observed observations, imputed observations, and bootstrap observations; (ii) the assumption that  $\hat{V}_1$  is fixed; and (iii) the assumption that  $\hat{V}/\hat{V}_{\infty, \infty}$ , where  $\hat{V}_{\infty, \infty}$  is the probability limit of  $\hat{V}$  as both  $M$  and  $B$  go to infinity, is distributed as  $\chi_v^2/v$ .

Let  $H_5 = \sum_{i \in A_R} \alpha_i (\bar{Y}_{r'}^b - H_4)$ ,  $H_6 = \sum_{j \in A_k} \alpha_j (\bar{\eta}_j^b - H_3)$ ,  $\hat{V}_5 = (B-1)^{-1} \sum_{b=1}^B (H_5 + H_6)^2$ ,  $\hat{V}_6 = (B-1)^{-1} \sum_{b=1}^B H_5^2$ , and  $\hat{V}_7 = (B-1)^{-1} \sum_{b=1}^B H_6^2$ . Let  $\alpha = \hat{V}_{2, \infty} / \hat{V}_2$ , where  $\hat{V}_{2, \infty} = \text{plim}_{M \rightarrow \infty} \hat{V}_2$ . Let  $\beta = \hat{V}_{3, \infty} / \hat{V}_3$ ,  $\gamma = \hat{V}_{5, \infty} / \hat{V}_5$ ,  $\kappa = \hat{V}_{6, \infty} / \hat{V}_6$ , and  $\phi = \hat{V}_{7, \infty} / \hat{V}_7$ , where  $\hat{V}_{\infty} = \text{plim}_{B \rightarrow \infty} \hat{V}$ . Let  $D^*$  be the dataset, which consists of the observations, multiple imputations, and bootstrap observations. Then,  $\alpha^{-1} | D^*$  follows an  $\chi^2$  distribution with degrees of freedom  $M-1$ . Similarly,  $\beta^{-1} | D^*$ ,  $\gamma^{-1} | D^*$ ,  $\kappa^{-1} | D^*$ , and  $\phi^{-1} | D^*$  follow

an  $\chi^2$  distribution with degrees of freedom  $B - 1$ . Therefore,  $\text{var}(\alpha^{-1}|D^*) = 2/(M - 1)$ ,  $\text{var}(\beta^{-1}|D^*) = \text{var}(\gamma^{-1}|D^*) = \text{var}(\kappa^{-1}|D^*) = \text{var}(\phi^{-1}|D^*) = 2/(B - 1)$ . Furthermore  $\text{cov}(\beta^{-1}, \gamma^{-1}|D^*) = 2(1 - r^2)/(B - 1)$ ,  $\text{cov}(\beta^{-1}, \kappa^{-1}|D^*) = 2/(B - 1)$ ,  $\text{cov}(\kappa^{-1}, \gamma^{-1}|D^*) = 2(1 - r^2)/(B - 1)$ ,  $\text{cov}(\phi^{-1}, \gamma^{-1}|D^*) = 2r^2/(B - 1)$ , and the covariances between any other two terms are zero.

Then, we can write  $\hat{V} = \hat{V}_1 + \hat{V}_2 + \hat{V}_3 + \hat{V}_5 - \hat{V}_6 - \hat{V}_7$ . Let  $f = \hat{V}_2/\hat{V}_1$ ,  $g = \hat{V}_3/\hat{V}_1$ ,  $h = \hat{V}_5/\hat{V}_1$ ,  $a = \hat{V}_6/\hat{V}_1$ , and  $b = \hat{V}_7/\hat{V}_1$ . Using a Taylor expansion, we derived

$$\frac{\hat{V}}{\hat{V}_{\infty, \infty}} = \frac{1+f+g+h-a-b}{1+\alpha f+\beta g+\gamma h-\kappa a-\phi b} \\ \approx 1 + \frac{f}{1+f+g+h-a-b}(\alpha^{-1} - 1) + \frac{g}{1+f+g+h-a-b}(\beta^{-1} - 1) \\ + 1 + \frac{h}{1+f+g+h-a-b}(\gamma^{-1} - 1) + \frac{a}{1+f+g+h-a-b}(\kappa^{-1} - 1) - \frac{b}{1+f+g+h-a-b}(\phi^{-1} - 1).$$

Let  $\pi$  be the response rate for the interested domain. We obtained

$$\text{var}\left(\frac{\hat{V}}{\hat{V}_{\infty, \infty}} \middle| D^*\right) = \frac{2}{M-1} \frac{f^2}{(1+f+g+h-a-b)^2} + \frac{2}{B-1} \frac{g^2+h^2+a^2+b^2+2gh(1-\pi^2)-2ga-2ha(1-\pi^2)-2hb\pi^2}{(1+f+g+h-a-b)^2} \\ = 2 \frac{(M-1)^{-1}f^2 + (B-1)^{-1}(g^2+h^2+a^2+b^2+2gh(1-\pi^2)-2ga-2ha(1-\pi^2)-2hb\pi^2)}{(1+f+g+h-a-b)^2}.$$

Thus, the degrees of freedom are

$$v = \frac{(1+f+g+h-a-b)^2}{(M-1)^{-1}f^2 + (B-1)^{-1}(g^2+h^2+a^2+b^2+2gh(1-\pi^2)-2ga-2ha(1-\pi^2)-2hb\pi^2)} \\ = \frac{(1+f+g+h-a-b)^2}{(M-1)^{-1}\hat{V}_2^2 + (B-1)^{-1}[\hat{V}_3^2 + \hat{V}_5^2 + \hat{V}_6^2 + \hat{V}_7^2 + 2(1-\pi^2)\hat{V}_5(\hat{V}_3 - \hat{V}_7) - 2\hat{V}_3\hat{V}_7 - 2\hat{V}_5\hat{V}_6\pi^2]}.$$

## References

1. Clogg, C.C.; Rubin, D.B.; Schenker, N.; Schultz, B.; Weidman, L. Multiple imputation of industry and occupation codes in census public-use samples using bayesian logistic-regression. *J. Am. Stat. Assoc.* **1991**, *86*, 68–78. [\[CrossRef\]](#)
2. Schafer, J.L.; Ezzati-Rice, T.M.; Johnson, W.; Khare, M.; Little, R.J.A.; Rubin, D.B. The NHANES III multiple imputation project. *Race Ethn.* **1996**, *60*, 28–37.
3. Gelman, A.; King, G.; Liu, C. Not asked and not answered: Multiple imputation for multiple surveys. *J. Am. Stat. Assoc.* **1998**, *93*, 846–857. [\[CrossRef\]](#)
4. Davey, A.; Shanahan, M.J.; Schafer, J.L. Correcting for selective nonresponse in the National Longitudinal Survey of Youth using multiple imputation. *J. Hum. Resour.* **2001**, *36*, 500–519. [\[CrossRef\]](#)
5. Taylor, J.M.G.; Cooper, K.L.; Wei, J.T.; Sarma, A.V.; Raghunathan, T.E.; Heeringa, S.G. Use of multiple imputation to correct for nonresponse bias in a survey of urologic symptoms among African-American men. *Am. J. Epidemiol.* **2002**, *156*, 774–782. [\[CrossRef\]](#)
6. Rubin, D.B. *Multiple Imputation for Nonresponse in Surveys*; John Wiley: Hoboken, NJ, USA, 1987.
7. Rubin, D.B. Multiple imputation after 18+ years. *J. Am. Stat. Assoc.* **1996**, *91*, 473–489. [\[CrossRef\]](#)
8. Meng, X.L. Multiple-imputation inferences with uncongenial sources of input. *Stat. Sci.* **1994**, *9*, 538–558.
9. Fay, R.E. *When Are Inferences from Multiple Imputation Valid?* Proceedings of the Section on Survey Research Methods; U.S. Bureau of the Census: Washington, DC, USA, 1992; pp. 227–232.
10. Fay, R.E. Valid inferences from imputed survey data. *Surv. Res. Methods* **1993**, 41–48.
11. Binder, D.A.; Sun, W.M.; Amer Stat, A. Frequency valid multiple imputation for surveys with a complex design. *Surv. Res. Methods* **1996**, 281–286.
12. Wang, N.; Robins, J.M. Large-sample theory for parametric multiple imputation procedures. *Biometrika* **1998**, *85*, 935–948. [\[CrossRef\]](#)
13. Nielsen, S.F. Proper and improper multiple imputation. *Int. Stat. Rev.* **2003**, *71*, 593–607. [\[CrossRef\]](#)
14. Robins, J.M.; Wang, N.S. Inference for imputation estimators. *Biometrika* **2000**, *87*, 113–124. [\[CrossRef\]](#)
15. Yang, S.; Kim, J.K. Fractional imputation in survey sampling: A Comparative Review. *Stat. Sci.* **2016**, *31*, 415–432. [\[CrossRef\]](#)
16. Kim, J.K.; Michael Brick, J.; Fuller, W.A.; Kalton, G. On the bias of the multiple-imputation variance estimator in survey sampling. *J. R. Stat. Soc. Ser. B* **2006**, *68*, 509–521. [\[CrossRef\]](#)
17. Efron, B.; Tibshirani, R. *An Introduction to the Bootstrap*; Chapman and Hall: New York, NY, USA, 1993.
18. Sarndal, C.E. Methods for estimating the precision of survey estimates when imputation has been used. *Surv. Methodol.* **1992**, *18*, 241–252.
19. Rao, J.N.K.; Shao, J. Jackknife variance-estimation with survey data under hot deck imputation. *Biometrika* **1992**, *79*, 811–822. [\[CrossRef\]](#)
20. Rao, J.N.K. On variance estimation with imputed survey data. *J. Am. Stat. Assoc.* **1996**, *91*, 499–506. [\[CrossRef\]](#)
21. Shao, J.; Sitter, R.R. Bootstrap for imputed survey data. *J. Am. Stat. Assoc.* **1996**, *91*, 1278–1288. [\[CrossRef\]](#)
22. Shao, J.; Steel, P. Variance estimation for survey data with composite imputation and nonnegligible sampling fractions. *J. Am. Stat. Assoc.* **1999**, *94*, 254–265. [\[CrossRef\]](#)

23. Haziza, D. Imputation and inference in the presence of missing data. In *Handbook of Statistics: Sample Surveys: Theory Methods and Inference*; Rao, C.R., Pfeffermann, D., Eds.; Elsevier: Amsterdam, The Netherlands, 2009; Volume 29A, pp. 215–246.
24. Kim, J.K.; Rao, J.N.K. A unified approach to linearization variance estimation from survey data after imputation for item nonresponse. *Biometrika* **2009**, *96*, 917–932. [[CrossRef](#)]
25. Chen, S.; Haziza, D.; Léger, C.; Mashreghi, Z. Pseudo-population bootstrap methods for imputed survey data. *Biometrika* **2019**, *106*, 369–384. [[CrossRef](#)] [[PubMed](#)]
26. Lu, K.F.; Li, D.Y.; Koch, G.G. Comparison between two controlled multiple imputation methods for sensitivity analyses of time-to-event data with possibly informative censoring. *Stat. Biopharm. Res.* **2015**, *7*, 199–213. [[CrossRef](#)]
27. Gao, F.; Liu, G.F.; Zeng, D.; Xu, L.; Lin, B.; Diao, G.; Golm, G.; Heyse, J.F.; Ibrahim, J.G. Control-based imputation for sensitivity analyses in informative censoring for recurrent event data. *Pharm. Stat.* **2017**, *16*, 424–432. [[CrossRef](#)] [[PubMed](#)]
28. Schomaker, M.; Heumann, H. Bootstrap inference when using multiple imputation. *Stat. Med.* **2018**, *37*, 2252–2266. [[CrossRef](#)] [[PubMed](#)]
29. Darken, P.; Nyberg, J.; Ballal, S.; Wright, D. The attributable estimand: A new approach to account for intercurrent events. *Pharm. Stat.* **2020**, *19*, 626–635. [[CrossRef](#)]
30. Nguyen, T.L.; Collins, G.S.; Pellegrini, F.; Moons, K.G.; Debray, T.P. On the aggregation of published prognostic scores for causal inference in observational studies. *Stat. Med.* **2020**, *39*, 1440–1457. [[CrossRef](#)]
31. Bartlett, J.W.; Hughes, R.A. Bootstrap inference for multiple imputation under uncongeniality and misspecification. *Stat. Methods Med. Res.* **2020**, *29*, 3533–3546. [[CrossRef](#)]
32. Satterthwaite, F.E. An approximate distribution of estimates of variance components. *Biom. Bull.* **1946**, *2*, 110–114. [[CrossRef](#)]
33. Schenker, N.; Welsh, A.H. Asymptotic results for multiple imputation. *Ann. Stat.* **1988**, *16*, 1550–1566. [[CrossRef](#)]
34. Hox, J.J. *Multilevel Analysis: Techniques and Applications*; Lawrence Erlbaum: Mahwah, NJ, USA, 2002.
35. Rubin, D.B.; Schenker, N. Multiple imputation for interval estimation from simple random samples with ignorable nonresponse. *J. Am. Stat. Assoc.* **1986**, *81*, 366–374. [[CrossRef](#)]