Research Article

Estimation of a Finite Population Mean under Random Nonresponse Using Kernel Weights

Nelson Kiprono Bii ¹, Christopher Ouma Onyango,² and John Odhiambo ¹

¹Institute of Mathematical Sciences, Strathmore University, P.O. Box 59857-00200, Nairobi, Kenya

²Department of Statistics, Kenyatta University, P.O. Box 43844-00200, Nairobi, Kenya

Correspondence should be addressed to Nelson Kiprono Bii; nkiprono@strathmore.edu

Received 27 January 2020; Accepted 27 March 2020; Published 21 April 2020

Academic Editor: Aera Thavaneswaran

Copyright © 2020 Nelson Kiprono Bii et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Nonresponse is a potential source of errors in sample surveys. It introduces bias and large variance in the estimation of finite population parameters. Regression models have been recognized as one of the techniques of reducing bias and variance due to random nonresponse using auxiliary data. In this study, it is assumed that random nonresponse occurs in the survey variable in the second stage of cluster sampling, assuming full auxiliary information is available throughout. Auxiliary information is used at the estimation stage via a regression model to address the problem of random nonresponse. In particular, auxiliary information is used via an improved Nadaraya–Watson kernel regression technique to compensate for random nonresponse. The asymptotic bias and mean squared error of the estimator proposed are derived. Besides, a simulation study conducted indicates that the proposed estimator has smaller values of the bias and smaller mean squared error values compared to existing estimators of a finite population mean. The proposed estimator is also shown to have tighter confidence interval lengths at 95% coverage rate. The results obtained in this study are useful for instance in choosing efficient estimators of a finite population mean in demographic sample surveys.

1. Introduction

Many authors such as [1–4] have looked at estimation of a finite population mean in the presence of nonresponse using various assumptions. However, the estimators developed in these studies need improvements on the efficiency and the bias. In the sequence of improving estimation of a finite population mean in the presence of random nonresponse, an improved Nadaraya–Watson kernel regression estimator is proposed in this study. The improved Nadaraya–Watson kernel regression technique was first fronted by [5]. To compensate for random nonresponse, auxiliary information is used in this study via an improved Nadaraya–Watson kernel regression technique due to [5].

An improvement of the Nadaraya–Watson estimator [6, 7] has been proposed by [5] using local bandwidth factor

λ_{ij} determined using [8] algorithm. The improved Nadaraya–Watson estimator is given by

$$m_{\text{IMW}}(\hat{x}) = \frac{\sum_{i \in s} \sum_{j \in s} (Y_{ij} / \lambda_{ij}) K((x - X_{ij}) / \lambda_{ij} b)}{\sum_{i \in s} \sum_{j \in s} (1 / \lambda_{ij}) K((x - X_{ij}) / \lambda_{ij} b)}, \quad (1)$$

where b is a smoothing parameter while the local bandwidth factor λ_{ij} is given by

$$\lambda_{ij} = \left\{ \frac{m(\hat{X}_{ij})}{a} \right\}^{-\alpha}, \quad (2)$$

where a is an arithmetic mean given by $a = \sum_{i=1}^n \sum_{j=1}^n m(\hat{X}_{ij}) / mn$ while α is a sensitivity parameter which satisfies $0 \leq \alpha \leq 1$. It has been suggested by [8] that taking $\alpha = (1/2)$ produces good results.

2. The Proposed Estimator of Finite Population Mean Using Improved Nadaraya–Watson Kernel Regression Technique

Consider a finite population of size N consisting of M clusters with N_j elements in the j^{th} cluster. A sample of m clusters is selected so that n_{1i} units respond and n_{2i} units fail to respond. Let y_{ij} denote the value of the survey variable y for unit j in cluster i , for $i = 1, 2, \dots, N$, $j = 1, 2, \dots, N_i$, and let the population mean be given by

$$\bar{Y} = \frac{1}{N} \sum_{i=1}^M \sum_{j=1}^{N_i} Y_{ij}. \quad (3)$$

The proposed estimator is given by

$$\hat{\bar{Y}}_{\text{INW}} = \frac{1}{M} \left\{ \frac{1}{n_{1i}} \sum_{i \in s} \sum_{j \in s} y_{ij} + \frac{1}{n_{2i}} \sum_{i \in s} \sum_{j \notin s} \hat{y}_{ij} \right\}, \quad (4)$$

where $\hat{y}_{ij}$ is an estimator of the nonresponse component of the sample. Assuming auxiliary information X_{ij} is known throughout, $\hat{y}_{ij}$ can be obtained using the improved Nadaraya–Watson regression technique by

$$\hat{y}_{ij} = m_{\text{INW}}(\hat{x}_{ij}) = \frac{\sum_{i \in s} \sum_{j \in s} (1/\lambda_{ij}) K((x - X_{ij})/\lambda_{ij}b) Y_{ij}}{\sum_{i \in s} \sum_{j \in s} (1/\lambda_{ij}) K((x - X_{ij})/\lambda_{ij}b)}, \quad (5)$$

so that the estimator of the finite population mean can be rewritten as

$$\hat{\bar{Y}}_{\text{INW}} = \frac{1}{M} \left\{ \frac{1}{n_{1i}} \sum_{i \in s} \sum_{j \in s} y_{ij} + \frac{1}{n_{2i}} \sum_{i \in s} \sum_{j \notin s} m_{\text{INW}}(\hat{x}_{ij}) \right\}. \quad (6)$$

A special case where $n_{1i} = n_{2i} = n$ is assumed in this study. This simplifies mathematical computations so that equation (7) can be rewritten as

$$\hat{\bar{Y}}_{\text{INW}} = \frac{1}{Mn} \left\{ \sum_{i \in s} \sum_{j \in s} y_{ij} + \sum_{i \in s} \sum_{j \notin s} m_{\text{INW}}(\hat{x}_{ij}) \right\}, \quad (7)$$

where $m_{\text{INW}}(\hat{x}_{ij})$ is the improved Nadaraya–Watson kernel regression estimator given in equation (1), which is a weighted sum of the values of the survey variable Y_{ij} . Data are generated using a regression model given by

$$\hat{Y}_{ij} = m(\hat{x}_{ij}) + \hat{e}_{ij}, \quad (8)$$

where $m(\cdot)$ is an unknown smooth function of auxiliary random variables X_{ij} . It is assumed that the error term $\hat{e}_{ij}$ satisfies the following conditions:

$$\begin{aligned} E(\hat{e}_{ij}) &= 0, \\ \text{Var}(\hat{e}_{ij}) &= \sigma_{ij}^2, \\ \text{Cov}(\hat{e}_i, \hat{e}_j) &= 0, \\ &\text{for } i \neq j. \end{aligned} \quad (9)$$

Hence, the unspecified function of the auxiliary random variables $m(\hat{x}_{ij})$ is replaced by the improved Nadaraya–Watson kernel estimator $m_{\text{INW}}(\hat{x}_{ij})$. The estimator can be rewritten as

$$m_{\text{INW}}(\hat{x}_{ij}) = \sum_{i \in s} \sum_{j \in s} w(x_{ij}) Y_{ij}, \quad (10)$$

where $w(x_{ij}) = ((1/\lambda_{ij})K((x - X_{ij})/\lambda_{ij}b))/\sum_{i \in s} \sum_{j \in s} (1/\lambda_{ij})K((x - X_{ij})/\lambda_{ij}b)$ are the improved Nadaraya–Watson kernel weights, where $K(\cdot)$ is a given kernel function assumed to be symmetrical. Since the choice of the kernel function is not critical for the performance of the kernel regression estimator, a simplified Gaussian kernel with mean 0 and variance 1 is used in this study. This is given by

$$K(w) = \frac{1}{\sqrt{2\pi}} e^{-(w^2/2)} = \frac{1}{\sqrt{2\pi}} e^{-((x - X_{ij})/\lambda_{ij}b)^2/2}. \quad (11)$$

In this case, the improved Nadaraya–Watson kernel estimation at any point x_{ij} is given by

$$\hat{y}_{ij} = m_{\text{INW}}(\hat{x}_{ij}) = \frac{\sum_i \sum_j (1/\sqrt{2\pi}) e^{-((x - X_{ij})/\lambda_{ij}b)^2/2} Y_{ij}}{\sum_i \sum_j (1/\sqrt{2\pi}) e^{-((x - X_{ij})/\lambda_{ij}b)^2/2}}, \quad (12)$$

where b is the bandwidth while λ_{ij} is given in equation (2) due to [5].

This provides a way of estimating the nonresponse values of the survey variable Y_{ij} , in the i^{th} cluster given the auxiliary values x_{ij} , for a specified kernel function.

2.1. The Asymptotic Bias of the Proposed Estimator. The expected value of the proposed estimator is given by

$$E(\hat{\bar{Y}}_{\text{INW}}) = \frac{1}{Mn} \left\{ \sum_{i=1}^n \sum_{j=1}^m Y_{ij} + \sum_{i=n+1}^N \sum_{j=m+1}^M m_{\text{INW}}(\hat{x}_{ij}) \right\}. \quad (13)$$

Rewriting equation (5) using the property of symmetry associated with Nadaraya–Watson estimator,

$$m_{\text{INW}}(\hat{x}_{ij}) = \frac{\sum_{i \in s} \sum_{j \in s} K((X_{ij} - x_{ij})/\lambda_{ij}b) Y_{ij}}{\sum_{i \in s} \sum_{j \in s} K((X_{ij} - x_{ij})/\lambda_{ij}b)}, \quad i = 1, 2, \dots,$$

$$n; j = 1, 2, \dots, m. \quad (14)$$

Following the procedure by [9], equation (14) can be rewritten as

$$m_{\text{INW}}(\hat{x}_{ij}) = \frac{1}{g(\hat{x}_{ij})} \left[\frac{1}{mn(\lambda_{ij}b)} \sum_i \sum_j K\left(\frac{X_{ij} - x_{ij}}{\lambda_{ij}b}\right) Y_{ij} \right], \quad (15)$$

where $g(\hat{x}_{ij})$ is the estimated marginal density of auxiliary variables X_{ij} . The bias of the estimator can be written as

$$\text{Bias}\left(\widehat{\bar{Y}}_{\text{INW}}\right) = E\left(\widehat{\bar{Y}}_{\text{INW}} - \bar{Y}\right), \quad (16)$$

$$\begin{aligned} \text{Bias}\left(\widehat{\bar{Y}}_{\text{INW}}\right) &= E\left\{\frac{1}{Mn}\left[\sum_{i=1}^n \sum_{j=1}^m Y_{ij} + \sum_{i=n+1}^N \sum_{j=m+1}^M m_{\text{INW}}(\hat{x}_{ij})\right]\right. \\ &\quad \left.- \frac{1}{Mn}\left[\sum_{i=1}^n \sum_{j=1}^m Y_{ij} + \sum_{i=n+1}^N \sum_{j=m+1}^M Y_{ij}\right]\right\}, \end{aligned} \quad (17)$$

which reduces to

$$\text{Bias}\left(\widehat{\bar{Y}}_{\text{INW}}\right) = \frac{1}{Mn} E\left\{\sum_{i=n+1}^N \sum_{j=m+1}^M m_{\text{INW}}(\hat{x}_{ij}) - \sum_{i=n+1}^N \sum_{j=m+1}^M Y_{ij}\right\}. \quad (18)$$

Rewriting the regression model given by $Y_{ij} = m(X_{ij}) + e_{ij}$ as

$$Y_{ij} = m(x_{ij}) + [m(X_{ij}) - m(x_{ij})] + e_{ij}, \quad (19)$$

and substituting it in equation (15) gives

$$\begin{aligned} m_{\text{INW}}(\hat{x}_{ij}) &= \frac{1}{g(\hat{x}_{ij})} \left[\frac{1}{mn(\lambda_{ij}b)} \sum_i \sum_j K\left(\frac{X_{ij} - x_{ij}}{\lambda_{ij}b}\right) (m(x_{ij}) \right. \\ &\quad \left. + [m(X_{ij}) - m(x_{ij})] + e_{ij}) \right]. \end{aligned} \quad (20)$$

Hence, the first term in equation (18) before taking expectation is given as

$$\begin{aligned} &\frac{1}{Mn} \left\{ \frac{1/mnb \sum_{i=n+1}^N \sum_{j=m+1}^M K((X_{ij} - x_{ij})/\lambda_{ij}b) Y_{ij}}{g(\hat{x}_{ij})} \right\} \\ &= \frac{1}{Mng(\hat{x}_{ij})} \left\{ \frac{1}{mn(\lambda_{ij}b)} \sum_{i=n+1}^N \sum_{j=m+1}^M K\left(\frac{X_{ij} - x_{ij}}{\lambda_{ij}b}\right) m(x_{ij}) \right. \\ &\quad + \frac{1}{mn(\lambda_{ij}b)} \sum_{i=n+1}^N \sum_{j=m+1}^M K\left(\frac{X_{ij} - x_{ij}}{\lambda_{ij}b}\right) [m(X_{ij}) - m(x_{ij})] \\ &\quad \left. + \frac{1}{mn(\lambda_{ij}b)} \sum_{i=n+1}^N \sum_{j=m+1}^M K\left(\frac{X_{ij} - x_{ij}}{\lambda_{ij}b}\right) e_{ij} \right\}. \end{aligned} \quad (21)$$

Simplifying equation (21), the following is obtained:

$$\begin{aligned} &\frac{1}{Mn} \left\{ \frac{1}{mn(\lambda_{ij}b)g(\hat{x}_{ij})} \sum_{i=n+1}^N \sum_{j=m+1}^M K\left(\frac{X_{ij} - x_{ij}}{\lambda_{ij}b}\right) Y_{ij} \right\} \\ &= \frac{1}{Mn} \left(\frac{1}{mn(\lambda_{ij}b)g(\hat{x}_{ij})} \right) \left\{ \sum_{i=n+1}^N \sum_{j=m+1}^M g(\hat{x}_{ij}) m(x_{ij}) \right. \\ &\quad \left. + m_1(\hat{x}_{ij}) + m_2(\hat{x}_{ij}) \right\}, \end{aligned} \quad (22)$$

where

$$m_1(\hat{x}_{ij}) = \frac{1}{mn(\lambda_{ij}b)} \sum_{i=n+1}^N \sum_{j=m+1}^M K\left(\frac{X_{ij} - x_{ij}}{\lambda_{ij}b}\right) [m(X_{ij}) - m(x_{ij})], \quad (23)$$

$$m_2(\hat{x}_{ij}) = \frac{1}{mn(\lambda_{ij}b)} \sum_{i=n+1}^N \sum_{j=m+1}^M K\left(\frac{X_{ij} - x_{ij}}{\lambda_{ij}b}\right) e_{ij}. \quad (24)$$

Taking conditional expectation of equation (22) leads to

$$\begin{aligned} E\left[\sum_{i=n+1}^N \sum_{j=m+1}^M m_{\text{INW}}(\hat{x}_{ij})\right] &= \frac{1}{Mn} E\left[\frac{1}{mn(\lambda_{ij}b)} \sum_{i=n+1}^N \sum_{j=m+1}^M \left[m(\hat{x}_{ij}) \right. \right. \\ &\quad \left. \left. + \frac{m_1(\hat{x}_{ij})}{g(\hat{x}_{ij})} + \frac{m_2(\hat{x}_{ij})}{g(\hat{x}_{ij})} \right] \right]. \end{aligned} \quad (25)$$

The following theorem due to [10] and applied by [11] was used in obtaining asymptotic bias and variance of the estimator using conditional expectations.

Theorem 1. Let $K(w)$ be a symmetric density function with $\int wk(w)dw = 0$ and $\int w^2k(w)dw = k_2$. Assume n and N increase together such that $(n/N) \rightarrow \pi$ with $0 < \pi < 1$. Besides, assume the sampled and nonsampled values of x are in the interval $[c, d]$ and are obtained by densities d_s and d_{p-s} , respectively, where both are bounded away from zero on $[c, d]$ with continuous second derivatives. If for any variable Z , $E(Z/U = u) = A(u) + O(B)$ and $\text{Var}(Z/U = u) = O(C)$, then $Z = A(u) + O_p(B + C^{1/2})$.

Using this theorem, the asymptotic bias and variance are derived in the following sections. From the conditions of the error term stated in equation (9), it follows that $E(e_{ij}/X_{ij}) = 0$. Therefore, $E[m_2(\hat{x}_{ij})] = 0$. Thus, $E[m_1(\hat{x}_{ij})]$ can be obtained as follows:

$$\begin{aligned} E\left[\sum_{i=n+1}^N \sum_{j=m+1}^M [m_1(\hat{x}_{ij})]\right] &= \frac{1}{Mn} \left(\frac{1}{mn(\lambda_{ij}b)} \right) E\left\{ \sum_{i=n+1}^N \sum_{j=m+1}^M \right. \\ &\quad \left. \cdot K\left(\frac{X_{ij} - x_{ij}}{\lambda_{ij}b}\right) \times [m(X_{ij}) - m(x_{ij})] \right\}. \end{aligned} \quad (26)$$

Using substitution and change of variable technique given by

$$\left. \begin{aligned} w &= \frac{V - x_{ij}}{\lambda_{ij}b} \\ V &= x_{ij} + (\lambda_{ij}b)w \\ dV &= (\lambda_{ij}b)dw \end{aligned} \right\}, \quad (27)$$

equation (26) can be simplified to

$$E \sum_{i=n+1}^N \sum_{j=m+1}^M [m_1(\hat{x}_{ij})] = \frac{1}{Mn} \left\{ \frac{Mn - mn}{mn} \int k(w) [m(x_{ij} + (\lambda_{ij}b)w) - m(x_{ij})] \int g(x_{ij} + (\lambda_{ij}b)w) dw \right\}. \quad (28)$$

Using Taylor's series expansion about the point x_{ij} , the k^{th} order kernel can be derived as follows:

$$\begin{aligned} g(x_{ij} + (\lambda_{ij}b)w) &= g(x_{ij}) + g'(x_{ij})(\lambda_{ij}b)w + \frac{1}{2}g''(x_{ij})(\lambda_{ij}b)^2w^2 \\ &+ \dots + \frac{1}{k!}g^{(k)}(x_{ij})(\lambda_{ij}b)^kw^k + o((\lambda_{ij}b)^2). \end{aligned} \quad (29)$$

$$\begin{aligned} m(x_{ij} + (\lambda_{ij}b)w) &= m(x_{ij}) + m'(x_{ij})(\lambda_{ij}b)w + \frac{1}{2}m''(x_{ij})(\lambda_{ij}b)^2w^2 \\ &+ \dots + \frac{1}{k!}m^{(k)}(x_{ij})(\lambda_{ij}b)^kw^k + o((\lambda_{ij}b)^2). \end{aligned} \quad (30)$$

Therefore, expanding equation (28) up to order $o((\lambda_{ij}b)^2)$ and simplifying gives

Similarly,

$$\begin{aligned} E \sum_{i=n+1}^N \sum_{j=m+1}^M [m_1(\hat{x}_{ij})] &= \frac{1}{Mn} \left\{ \left(\frac{Mn - mn}{mn} \right) g(\hat{x}_{ij}) m'(x_{ij}) (\lambda_{ij}b) \int w k(w) dw + \left(\frac{Mn - mn}{mn} \right) g'(\hat{x}_{ij}) m'(x_{ij}) (\lambda_{ij}b)^2 \int w^2 k(w) dw \right. \\ &+ \left. \left(\frac{Mn - mn}{mn} \right) \frac{1}{2} g(\hat{x}_{ij}) m''(x_{ij}) (\lambda_{ij}b)^2 \times \int w^2 k(w) dw + o((\lambda_{ij}b)^2) \right\}. \end{aligned} \quad (31)$$

Using the conditions due to [10] given by $\int_{-\infty}^{\infty} k(w) dw = 1$, $\int_{-\infty}^{\infty} w k(w) dw = 0$, and $\int_{-\infty}^{\infty} w^2 k(w) dw = d_k$, the derivation in equation (31) can further be simplified to obtain

$$E \sum_{i=n+1}^N \sum_{j=m+1}^M [m_1(\hat{x}_{ij})] = \frac{1}{Mn} \left(\frac{Mn - mn}{mn} \right) \left[g'(x_{ij}) m'(x_{ij}) + \frac{1}{2} g(x_{ij}) m''(x_{ij}) \right] (\lambda_{ij}b)^2 d_k + o((\lambda_{ij}b)^2). \quad (32)$$

Hence, the expected value of the second term in equation (25) then becomes

$$E \sum_{i=n+1}^N \sum_{j=m+1}^M [m_1(\hat{x}_{ij})] = \frac{1}{Mn} \left\{ \left(\frac{Mn - mn}{mn} \right) \left[\frac{1}{2g(\hat{x}_{ij})} m''(x_{ij}) g(x_{ij}) + \frac{g'(x_{ij}) m'(x_{ij})}{g(\hat{x}_{ij})} \right] (\lambda_{ij}b)^2 d_k + o((\lambda_{ij}b)^2) \right\}. \quad (33)$$

Simplifying equation (33) gives

$$E \sum_{i=n+1}^N \sum_{j=m+1}^M [m_1(\hat{x}_{ij})] = \frac{1}{Mn} \left\{ \left(\frac{Mn - mn}{mn} \right) (\lambda_{ij}b)^2 d_k C(x) + o((\lambda_{ij}b)^2) \right\}, \quad (34)$$

where $C(x) = [g(\hat{x}_{ij})]^{-1} [(1/2)m''(x_{ij})g(x_{ij}) + g'(x_{ij})m'(x_{ij})]$ and $d_k = \int w^2 k(w)dw$.

Using the equation of the bias given in (16) and the conditional expectation in equation (25), the following equation for the conditional bias of the estimator was obtained:

$$\text{Bias}\left(\frac{\hat{\bar{Y}}_{\text{INW}}}{x_{ij}}\right) = \frac{1}{Mn} \left\{ \left(\frac{Mn - mn}{mn} \right) (\lambda_{ij}b)^2 d_k C(x) + o\left((\lambda_{ij}b)^2\right) \right\}. \quad (35)$$

In the next subsection, the asymptotic variance of the estimator is also derived.

2.2. Asymptotic Variance of the Proposed Estimator. Using equation (7), the conditional variance of the estimator is given as

$$\begin{aligned} \text{Var}\left(\frac{\hat{\bar{Y}}_{\text{INW}}}{x_{ij}}\right) &= \text{Var}\left\{ \frac{1}{Mn} \left\{ \sum_{i=1}^n \sum_{j=1}^m Y_{ij} + \sum_{i=n+1}^N \sum_{j=m+1}^M m_{\text{INW}}(\hat{x}_{ij}) \right\} \right\} \\ &= \left(\frac{1}{Mn} \right)^2 \text{Var}\left\{ \sum_{i=1}^n \sum_{j=1}^m Y_{ij} + \sum_{i=n+1}^N \sum_{j=m+1}^M m_{\text{INW}}(\hat{x}_{ij}) \right\}, \end{aligned} \quad (36)$$

$$\text{Var}\left(\frac{\hat{\bar{Y}}_{\text{INW}}}{x_{ij}}\right) = \left(\frac{1}{Mn} \right)^2 \text{Var}\left\{ \sum_{i=1}^n \sum_{j=1}^m Y_{ij} + \left(\frac{1}{mn(\lambda_{ij}b)g(\hat{x}_{ij})} \right) \left\{ \sum_{i=n+1}^N \sum_{j=m+1}^M g(\hat{x}_{ij})m(x_{ij}) + m_1(\hat{x}_{ij}) + m_2(\hat{x}_{ij}) \right\} \right\}. \quad (38)$$

From equation (24),

$$m_2(\hat{x}_{ij}) = \frac{1}{mn b} \sum_{i=n+1}^N \sum_{j=m+1}^M K\left(\frac{X_{ij} - x_{ij}}{\lambda_{ij}b}\right) e_{ij}. \quad (39)$$

Hence,

$$\text{Var} \sum_{i=n+1}^N \sum_{j=m+1}^M [m_2(\hat{x}_{ij})] = \frac{1}{(Mn)^2} \left(\frac{Mn - mn}{mn(\lambda_{ij}b)} \right)^2 \sum_{i=1}^n \sum_{j=1}^m \text{Var}(D_x), \quad (40)$$

where $D_x = K((X_{ij} - x_{ij})/\lambda_{ij}b)e_{ij}$. Expressing equation (40) in terms of expectation, the following equation is obtained

$$\begin{aligned} \text{Var} \sum_{i=n+1}^N \sum_{j=m+1}^M [m_2(\hat{x}_{ij})] &= \frac{1}{(Mn)^2} \left[\frac{(Mn - mn)^2}{mn(\lambda_{ij}b)^2} \right] \\ &\quad \{E[D_x]^2 - [E(D_x)]^2\}. \end{aligned} \quad (41)$$

Using the fact that the conditional expectation $E(e_{ij}/X_{ij}) = 0$, the second term in equation (41) reduces to zero. Therefore,

where $m_{\text{INW}}(\hat{x}_{ij})$ is given by

$$m_{\text{INW}}(\hat{x}_{ij}) = \frac{1}{g(\hat{x}_{ij})} \left[\frac{1}{mn(\lambda_{ij}b)} \sum_i \sum_j K\left(\frac{X_{ij} - x_{ij}}{\lambda_{ij}b}\right) Y_{ij} \right], \quad (37)$$

where $g(\hat{x}_{ij}) = (1/mn(\lambda_{ij}b)) \sum_i \sum_j K((x - X_{ij})/\lambda_{ij}b)$ is the estimated marginal density of auxiliary variables X_{ij} ; for details see [6, 7]. Rewriting the regression model $Y_{ij} = m(X_{ij}) + e_{ij}$ as $Y_{ij} = m(x_{ij}) + [m(X_{ij}) - m(x_{ij})] + e_{ij}$ and substituting in equation (37) leads to

$$\text{Var} \sum_{i=n+1}^N \sum_{j=m+1}^M [m_2(\hat{x}_{ij})] = \frac{1}{(Mn)^2} \left[\frac{(Mn - mn)^2}{mn(\lambda_{ij}b)^2} \right] \sigma_{ij}^2, \quad (42)$$

where $E(e_{ij}/X_{ij})^2 = \sigma_{ij}^2$.

Let $X = X_{ij}$ and $x = x_{ij}$ and make the following substitutions:

$$\left. \begin{aligned} w &= \frac{X - x}{\lambda_{ij}b} \\ X - x &= (\lambda_{ij}b)w \\ dX &= (\lambda_{ij}b)dw \end{aligned} \right\}, \quad (43)$$

so that

$$\begin{aligned} \text{Var} \sum_{i=n+1}^N \sum_{j=m+1}^M [m_2(\hat{x}_{ij})] &= \frac{(Mn - mn)^2}{mn(\lambda_{ij}b)^2 (Mn)^2} \\ &\quad \int K\left(\frac{X - x}{\lambda_{ij}b}\right)^2 \sigma_x^2 g(X) dX. \end{aligned} \quad (44)$$

Using the change of variables technique and simplifying, equation (44) reduces to

$$\text{Var} \sum_{i=n+1}^N \sum_{j=m+1}^M [m_2(\hat{x}_{ij})] = \frac{(Mn - mn)^2}{mn(\lambda_{ij}b)(Mn)^2} \int K(w)^2 \sigma_x^2 g(x) dw + o\left(\frac{1}{mn(\lambda_{ij}b)}\right). \quad (45)$$

Following the same procedure for getting the variance of $m_2(\hat{x}_{ij})$, $\text{Var} \sum_{i=n+1}^N \sum_{j=m+1}^M [m_1(\hat{x}_{ij})]$ can similarly be obtained as follows:

$$\text{Var} \sum_{i=n+1}^N \sum_{j=m+1}^M [m_1(\hat{x}_{ij})] = \frac{1}{(Mn)^2} \text{Var} \sum_{i=n+1}^N \sum_{j=m+1}^M \left[\frac{1}{mn(\lambda_{ij}b)} \sum_{i=1}^n \sum_{j=1}^m K\left(\frac{X_{ij} - x_{ij}}{\lambda_{ij}b}\right) \right] \times [m(X_{ij}) - m(x_{ij})]. \quad (46)$$

Equation (46) can be rewritten as

$$\text{Var} \sum_{i=n+1}^N \sum_{j=m+1}^M [m_1(\hat{x}_{ij})] = \frac{(Mn - mn)^2}{mn b^2 (Mn)^2} \text{Var} K\left(\frac{X_{ij} - x_{ij}}{\lambda_{ij}b}\right)^2 [m(X_{ij}) - m(x_{ij})]^2 g(X) dX, \quad (47)$$

where $X = (\lambda_{ij}b)w + x$ so that $dX = (\lambda_{ij}b)dw$. Changing variables and applying Taylor's series expansion about the point x_{ij} leads to

$$\text{Var} \sum_{i=n+1}^N \sum_{j=m+1}^M [m_1(\hat{x}_{ij})] = \frac{(Mn - mn)^2}{mn(\lambda_{ij}b)^2 (Mn)^2} \int K(w^2) [m(x + (\lambda_{ij}b)w) - m(x)]^2 \times g\left(x + (\lambda_{ij}b)w\right) dw, \quad (48)$$

which gives

$$\text{Var} \sum_{i=n+1}^N \sum_{j=m+1}^M [m_1(\hat{x}_{ij})] = \frac{(Mn - mn)^2}{mn(\lambda_{ij}b)^2 (Mn)^2} \int K(w^2) [m(x) + m'(x)(\lambda_{ij}b)w + \dots - m(x)]^2 (g(x) + g'(x)(\lambda_{ij}b)w) dw. \quad (49)$$

Following the procedure by [12] and simplifying, equation (49) reduces to

$$\text{Var} \sum_{i=n+1}^N \sum_{j=m+1}^M [m_1(\hat{x}_{ij})] = o\left[\frac{(Mn - mn)^2 b^2}{mn(\lambda_{ij}b)}\right]. \quad (50)$$

For large samples, as $n \rightarrow N, m \rightarrow M$, and $b \rightarrow 0$, then $mn(\lambda_{ij}b) \rightarrow \infty$. Hence, the variance in equation (49) asymptotically tends to zero, i.e., $\text{Var} \sum_{i=1}^N \sum_{j=1}^M [m_1(\hat{x}_{ij})] \rightarrow 0$, so that the variance of the estimator of the population mean reduces to

$$\text{Var}\left(\frac{\hat{\bar{Y}}_{\text{INW}}}{x_{ij}}\right) = \frac{(Mn - mn)^2}{mn(\lambda_{ij}b)(Mn)^2} \sum_{i=n+1}^N \sum_{j=m+1}^M \text{Var}\left[m(x_{ij}) + \frac{m_1(\hat{x}_{ij}) + m_2(\hat{x}_{ij})}{g(\hat{x}_{ij})}\right]. \quad (51)$$

Simplifying equation (51) leads to

$$\text{Var}\left(\frac{\widehat{\widehat{Y}}_{\text{INW}}}{x_{ij}}\right) = \frac{(Mn - mn)^2}{mn(\lambda_{ij}b)(Mn)^2[g(\widehat{x}_{ij})]^2} \text{Var}\left\{\sum_{i=n+1}^N \sum_{j=m+1}^M [m_2(\widehat{x}_{ij})]\right\}. \quad (52)$$

Substituting equation (45) in equation (52) yields the following:

$$\text{Var}\left(\frac{\widehat{\widehat{Y}}_{\text{INW}}}{x_{ij}}\right) = \frac{1}{(Mn)^2} \left\{ \frac{(Mn - mn)^2 \int K(w)^2 \sigma_{x_{ij}}^2 dw}{mn(\lambda_{ij}b)g(\widehat{x}_{ij})} + o\left[\frac{(Mn - mn)^2}{mn(\lambda_{ij}b)} + \frac{1}{mn(\lambda_{ij}b)}\right] \right\}. \quad (53)$$

2.3. Mean Squared Error of the Proposed Estimator. The conditional MSE of the estimator of the finite population mean combines the conditional squared bias and the conditional variance of the estimator, that is,

$$\text{MSE}\left(\frac{\widehat{\widehat{Y}}_{\text{INW}}}{x_{ij}}\right) = \text{Var}\left(\frac{\widehat{\widehat{Y}}_{\text{INW}}}{x_{ij}}\right) + \text{Bias}^2\left(\frac{\widehat{\widehat{Y}}_{\text{INW}}}{x_{ij}}\right), \quad (54)$$

which on simplification leads to

$$\begin{aligned} \text{MSE}\left(\frac{\widehat{\widehat{Y}}}{X_{ij} = x_{ij}}\right) &= \frac{1}{(Mn)^2} \left\{ \frac{(Mn - mn)^2 H(w) \sigma_{x_{ij}}^2}{mn(\lambda_{ij}b)g(\widehat{x}_{ij})} \right. \\ &\quad + \left[\frac{(Mn - mn)^2}{4(mn)^2 (Mn)^2} (\lambda_{ij}b)^2 d_k^2 \left[m''(x_{ij})g(x_{ij}) + \frac{2g'(x_{ij})m'(x_{ij})}{g(\widehat{x}_{ij})} \right]^2 \right. \\ &\quad \left. \left. + o\left(\frac{1}{Mn} \left\{ \frac{(Mn - mn)^2}{mn(\lambda_{ij}b)} + \frac{1}{mn(\lambda_{ij}b)} \right\}\right) \right] \right\}, \end{aligned} \quad (55)$$

where $H(w) = \int K(w)^2 dw$ and $d_k = \int w^2 K(w) dw$.

From equation (55), it is noted that if the sample size is large, that is, as $n \rightarrow N$ and $m \rightarrow M$, the MSE of $\widehat{\widehat{Y}}_{\text{INW}}$ due to the kernel tends to zero for a sufficiently small bandwidth. The estimator $\widehat{\widehat{Y}}$ is therefore asymptotically consistent since its MSE converges to zero in probability.

3. Simulation Study

A simulation experiment was conducted using R code in order to compare the performance of the proposed estimator in two-stage cluster sampling with the transformed estimator due to [13] and the nonparametric regression estimator due to [14]. An asymptotic framework is used where both the population number of clusters and the sample number of clusters are large. The number of clusters within each cluster N_i is held constant so that no cluster dominates the population.

Both linear and nonlinear mean functions of auxiliary random variables due to [14] were considered in generating data, where $x \in (0, 1)$. The equations of the mean functions used in simulating the data are given in Table 1.

The population auxiliary values x_{ij} of size $M = 2000$ are generated as identical and independently distributed uniform $(0, 1)$ random variables. The survey values are only known for the respondents in the selected sample. Using the auxiliary values, the nonresponse values are generated, that is, for every generated value $x_{ij}, i = 1, 2, \dots, M; j = 1, 2, \dots, N_i$, the mean survey nonresponse values are generated as

$$\widehat{y}_{ij} = \frac{1}{M} \left\{ \frac{m(\widehat{x}_{ij})}{N_i} + \frac{\widehat{e}_{ij}}{N_i} \right\}, \quad (56)$$

where $\widehat{e}_{ij}$ are identically and independently distributed normal random variables with mean zero and variance one. Besides, a Gaussian kernel with mean zero and variance one was used. A Gaussian kernel was used since it has smooth and continuous derivatives at every data point. Besides, an optimal bandwidth generated using cross-validation technique due to [15] was used. It has been noted by [15] that this bandwidth would lead to more informative estimates compared to other choices. The local bandwidth factor λ_{ij}

TABLE 1: Equations of mean functions simulated.

Mean function	Equation
Linear	$m_1(\hat{x}) = 1 + 2(x - 0.5)$
Quadratic	$m_2(\hat{x}) = 1 + 2(x - 0.5)^2$
Sine	$m_3(\hat{x}) = 2 + \sin(2\pi x)$
Exponential	$m_4(\hat{x}) = \exp(-8x)$
Bump	$m_5(\hat{x}) = 1 + 2(x - 0.5) + \exp\{-200(x - 0.5)^2\}$
Jump	$m_6(\hat{x}) = 1 + 2(x - 0.5)I_{x \leq 0.65} + 0.65I_{x \geq 0.65}$

given in equation (2) was generated using the algorithm due to [8].

At stage one, a sample of clusters is generated first by simple random sampling using a sample of size $m = 200$. At stage two, subsamples of elements within every selected cluster are generated by simple random sampling with replacement using a random sample of size n_i . The nonresponse mean survey values were then generated using equation (56). The estimates of the finite population mean were then computed using the estimator in equation (7). The values of bias and mean squared error values were also computed. The 95% confidence intervals were then constructed for the estimators of the finite population mean for comparative purposes.

4. Simulation Results

The values of the bias, mean squared error, and confidence interval lengths are given in the following tables. Note that $\hat{\bar{Y}}_{INW}$ is the estimator of the finite population mean proposed in this study and $\hat{\bar{Y}}_{TDM}$ is the transformation of data method estimator of the finite population mean due to [13] whereas $\hat{\bar{Y}}_{REG}$ is the nonparametric regression estimator due to [14]. Both $\hat{\bar{Y}}_{TDM}$ and $\hat{\bar{Y}}_{REG}$ were used for comparative purposes with the proposed estimator.

The biases of the estimators considered are presented in Table 2. Negative values of the bias imply underestimation while positive values of the bias indicate overestimation of the finite population mean by the different estimators. The proposed estimator has relatively smaller values of the bias followed by transformation of data method estimator due to [13]. The nonparametric-based estimator due to [14] has larger values compared to the other two estimators. It is also observed that the three estimators have relatively closer values of the bias in the quadratic mean function though the transformation of data method has positive bias at this mean function. Generally, among the three estimators of the finite population mean, the proposed estimator using the improved Nadaraya–Watson kernel regression technique performs better than the other two estimators in terms of bias.

Mean squared error combines both the variance and the squared bias terms of an estimator. The mean squared error values presented in Table 3 were simulated using the different mean functions indicated. The quadratic mean function gives the smallest value of the mean squared error of the proposed estimator followed by the linear function. The estimator due to [14] has the largest value of the mean

TABLE 2: Summary results of bias.

Estimator	$\hat{\bar{Y}}_{INW}$	$\hat{\bar{Y}}_{TDM}$	$\hat{\bar{Y}}_{REG}$
Linear	-0.00213	-0.1667	-0.3312
Quadratic	-0.0132	0.04171	-0.0966
Sine	-0.0521	-0.6416	-1.2311
Exponential	-0.0041	0.3592	0.7225
Bump	-0.0032	-0.2358	-0.4685
Jump	-0.0188	-0.2466	-0.4743

TABLE 3: Summary results of MSE values.

Estimator	$\hat{\bar{Y}}_{INW}$	$\hat{\bar{Y}}_{TDM}$	$\hat{\bar{Y}}_{REG}$
Linear	0.0334	0.1097	0.1321
Quadratic	0.0093	0.1455	0.5835
Sine	0.4215	1.5157	1.555
Exponential	0.3430	0.5220	1.3780
Bump	0.0634	0.2195	0.2508
Jump	0.2250	0.2951	1.1611

TABLE 4: Summary results of 95% CI lengths.

Estimator	$\hat{\bar{Y}}_{INW}$	$\hat{\bar{Y}}_{TDM}$	$\hat{\bar{Y}}_{REG}$
Linear	0.7164	1.2984	1.4249
Quadratic	0.8270	1.4951	2.994
Sine	1.5269	2.5451	4.888
Exponential	2.2958	2.8322	4.6016
Bump	0.9872	1.8365	1.9630
Jump	1.8594	2.1297	4.2239

squared error in the jump function. Generally, it is noted from Table 3 that the mean squared error values for the proposed estimator are relatively smaller than the rest of the estimators considered. The transformation of data method estimator due to [13] follows closely in the second place with smaller mean squared error values compared to nonparametric regression-based estimator due to [14]. From this comparison of the mean squared error values, it can be concluded that the proposed estimator is more efficient than the other two estimators considered. It has got smaller MSE values in all the mean functions and thus outperforms the others in terms of efficiency.

The 95% upper and lower confidence intervals were constructed for the estimators of the finite population mean. Confidence interval lengths were then obtained. The results are given in Table 4. From the values obtained, it is noted that the confidence interval lengths for the proposed estimator are much tighter than those of the estimators due to [13, 14]. Hence, at 95% level of confidence, the estimator proposed in this study performs better than its rival estimators.

5. Conclusion

This study has developed an estimator of the finite population mean in two-stage cluster sampling, assuming random nonresponse occurs in the survey variable in the second stage of cluster sampling. Complete auxiliary

information is assumed to be available in both stage one and stage two of cluster sampling. Kernel weights developed using the improved Nadaraya–Watson regression technique were used in the estimation process. The theoretical properties of the proposed estimator such as asymptotic bias, variance, and mean squared error were derived. Simulation results show that the proposed estimator has smaller values of the bias, smaller mean squared error values, and tighter confidence interval lengths compared to the other estimators. Therefore, the estimator of the finite population mean proposed in this study dominates the estimators due to [13, 14], respectively.

Data Availability

The data used to support the theoretical findings were generated via simulation using R statistical package.

Disclosure

The abstract was to be presented in a conference organized by the World Academy of Science, Engineering and Technology, but due to financial constraints, participation and presentation in the conference was withdrawn and the organizer was informed accordingly.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] N. K. Bii, C. O. Onyango, and J. Odhiambo, “Estimating a finite population mean under random non response in two stage cluster sampling with replacement,” *Journal of Physics: Conference Series*, vol. 1132, Article ID 012071, 2018.
- [2] H. P. Singh and S. Kumar, “Estimation of mean in presence of non-response using two phase sampling scheme,” *Statistical Papers*, vol. 51, no. 3, pp. 559–582, 2010.
- [3] H. P. Singh and S. Kumar, “Combination of regression and ratio estimate in presence of nonresponse,” *Brazilian Journal of Probability and Statistics*, vol. 25, no. 2, pp. 205–217, 2011.
- [4] H. P. Singh, S. Kumar, and M. Kozak, “Improved estimation of finite-population mean using sub-sampling to deal with non response in two-phase sampling scheme,” *Communications in Statistics-Theory and Methods*, vol. 39, no. 5, pp. 791–802, 2010.
- [5] S. Demir and Ö. Toktamiş, “On the adaptive Nadaraya-Watson kernel regression estimators,” *Hacettepe Journal of Mathematics and Statistics*, vol. 39, no. 3, 2010.
- [6] E. A. Nadaraya, “On estimating regression,” *Theory of Probability & Its Applications*, vol. 9, no. 1, pp. 141–142, 1964.
- [7] G. S. Watson, “Smooth regression analysis,” *Sankhyā: The Indian Journal of Statistics*, vol. Series A, pp. 359–372, 1964.
- [8] H. Läuter, “Silverman, B. W.: density estimation for statistics and data analysis. chapman & hall, London–New York 1986, 175 pp., £ 12.-,” *Biometrical Journal*, vol. 30, no. 7, pp. 876–877, 1988.
- [9] K. Takezawa, *Introduction to Nonparametric Regression*, vol. 606, John Wiley & Sons, Hoboken, NJ, USA, 2005.
- [10] A. H. Dorfman, “Nonparametric regression for estimating totals in finite populations,” in *Proceedings of the Section on Survey Research Methods*, pp. 622–625, American Statistical Association, Alexandria, VA, USA, 1992.
- [11] D. Ruppert, S. J. Sheather, and M. P. Wand, “An effective bandwidth selector for local least squares regression,” *Journal of the American Statistical Association*, vol. 90, no. 432, pp. 1257–1270, 1995.
- [12] E. Bach, J. O. Shallit, J. Shallit, and S. Jeffrey, *Algorithmic Number Theory: Efficient Algorithms*, vol. 1, MIT Press, MIT Press, Cambridge, MA, USA, 1996.
- [13] N. K. Bii, C. O. Onyango, and J. Odhiambo, “Boundary bias correction using weighting method in presence of nonresponse in two-stage cluster sampling,” *Journal of Probability and Statistics*, vol. 2019, Article ID 6812795, 8 pages, 2019.
- [14] F. J. Breidt and J. D. Opsomer, “Local polynomial regression estimators in survey sampling,” *The Annals of Statistics*, vol. 28, no. 4, pp. 1026–1053, 2000.
- [15] C. R. Loader, “Bandwidth selection: classical or plug-in?” *The Annals of Statistics*, vol. 27, no. 2, pp. 415–438, 1999.