Journal of Statistical and Econometric Methods, vol.8, no.1, 2019, 69-82
ISSN: 1792-6602 (print), 1792-6939 (online)
Scienpress Ltd, 2019

Estimation of Nested Error Non-parametric Unit Level Model

Patrick Munyangabo[1] Anthony Waititu[2] and Anthony Kibira Wanjoya[3]

Abstract

In this paper, we proposed a nested error nonparametric unit level model, when the linearity assumptions have been violated. The model formulation and parameter estimation of mean function were examined and proposed two theorems for asymptotic properties of mean function of proposed model were done. The simulation study was performed and it has shown that the mean square errors (MSE) and bias of our estimated were close to zero.

Keywords: Schwarzs inequality; Triangle inequality

[1] Department of Mathematics, Pan African University, Institute of Basic Sciences, Technology, and Innovation, Kenya. E-mail: munyangabopatrick@gmail.com
[2] Affiliation of the first author. E-mail: awaititu@jkuat.ac.ke
[3] Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, Kenya. E-mail: awanjoya@fsc.jkuat.ac.ke

Article Info: *Received* : September 23, 2018. *Revised* : October 21, 2018.
Published online : January 1, 2019

1 Introduction

In statistics, linear regression uses a linear approach to express the relationship between a dependent and one or more independent variables. In small area estimation, the approach for unit level model expresses the relationship between the variable of interest and the auxiliary variables as a linear regression model plus a random effect. It assumes the linearity among them. For instance, the well known model is found in [6], In case, where that assumption is violated, there are alternative approaches in the literature. [1] extended the model found in [6] for solving the issue of complex estimation, by incorporating nonparametric regression unit level model in small area estimation. The penalized spline regression was used. The work of [1] was extended by [3] to handle non continuous response variables. To deal with the problem of non-linearity in variable of interest, [7] proposed the M-quantile approach. A unit level log-normal model was proposed by [8] to deal with the problem of violation of normality assumptions. The proposed nested error in unit level nonparametric model is the purpose of this paper. We begin by the model formulation, then we perform parameter estimation of mean function and finally, we derive the asymptotic properties of our estimator.

2 Formulation of Proposed model

We consider a situation where the variable of interest and the covariates of auxiliary variable of proposed model has non linear relationship among them. Because the assumptions of the linear model with normal errors are violated, linear predictors are inefficient. That is why we look at the case where units in the population are assumed to have log-normal distributions. The locally weighted regression (loess or lowess) was used to approximate the mean function of local polynomial in the proposed model. The parameters of proposed estimator can be derived using least square method. The proposed model for the variable of interest, y_{jd}, is defined as

$$log(y_{jd}) := l_{jd} = m(x_{jd}) + u_j + e_{jd}. \tag{1}$$

where the function $m(\cdot)$ is unknown, but estimated by locally weighted regression (loess or lowess), with parameter polynomial model in one variable is given by:

$$m(x, \beta) = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots \beta_p x^p \tag{2}$$

The model (2) is called the p^{th} order model.

According to [10], polynomial models are also useful to approximating unknown functions and possibly very complex non-linear relationship. The $\log(y_{jd})$ is the logarithm of variable of interest, $\mathbf{X}_{jd} = (1, x_{1,jd}, \cdots, x_{p-1,jd})$ is a p-vector of auxiliary variables on unit j of area d, u_d is an area-level random effect, and e_{jd} is the error term. The u_d and e_{jd} are mutually independent with zero mean and variance σ_u^2 and σ_e^2 respectively.

3 Parameter Estimation for mean function of proposed Model

In order to estimate the parameters of mean function of proposed Model, the weighted function for local weight regression (loess or lowess) was defined, to get good approximation of local polynomial. The Taylor's series expansion were used to enable us the properties of the estimator, then the least square method was employed to estimate the local polynomial parameters.

3.1 Weighted function

The local weight regression (loess or lowess) commonly uses the kernel function known as the weight tricube [9]. The tricube weight function is defined as

$$t(d) = \begin{cases} (1 - |d|^3)^3 & ; \quad 0 \leq d \leq 1 \\ 0 & ; \quad \text{Otherwise} \end{cases}$$

The weight sequence is defined as

$$w_i(x_0) = t\left(\frac{|x_0 - x_i|}{b}\right) \tag{3}$$

where b is the bandwidth and depending on targets covariance x_0.

3.2 Parameter estimation of Mean function

The local polynomial of mean function is

$$m(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_p x^p = \sum_{j=1}^{p} \beta_p x^p \tag{4}$$

and the $k^{\text{th}}, k > 0$, derivative of $m(x)$ is given by

$$m^{(k)}(x) = \sum_{j=0}^{p} \prod_{i=0}^{k-1} (j-i)\beta_j x^{j-k} \tag{5}$$

The good approximation of local polynomial was obtained by application Taylor's Theorem as follows

$$\begin{aligned} m(x) &= m(x_0) + m'(x_0)(x-x_0) + \frac{1}{2!}m'(x_0)(x-x_0)^2 + \cdots + \frac{1}{p!}m^{(p)}(x_0)(x-x_0)^p \\ &= m(x_0) + \sum_{k=1}^{p} \frac{1}{k!} m^{(k)}(x_0)(x-x_0)^k \end{aligned} \tag{6}$$

Now, assuming that $x_0 = 0$, we have

$$\begin{aligned} m'(x_0) &= \sum_{j=0}^{p} \prod_{i=0}^{0} (j-i)\beta_j x_0^{j-1} = \sum_{j=0}^{p} j\beta_j x_0^{j-1} = 0 + \beta_1 x_0^0 + \underbrace{\sum_{j=2}^{p} j\beta_j x_0^{j-1}}_{=0} = \beta_1 \\ m''(x_0) &= \sum_{j=0}^{p} \prod_{i=0}^{1} (j-i)\beta_j x_0^{j-2} = \sum_{j=0}^{p} j(j-1)\beta_j x_0^{j-2} = 2\beta_2 + \underbrace{\sum_{j=3}^{p} j(j-1)\beta_j x_0^{j-2}}_{=0} \\ m'''(x_0) &= \sum_{j=0}^{p} \prod_{i=0}^{2} (j-i)\beta_j x_0^{j-2} = \sum_{j=0}^{p} j(j-1)(j-2)\beta_j x_0^{j-3} = 6\beta_3 + 0 \end{aligned} \tag{7}$$

In general, considering the formula in (5)

$$\begin{aligned} m^{(k)}(x_0) &= \sum_{j=0}^{k} \prod_{i=0}^{k-1} (j-i)\beta_j x^{j-k} + \underbrace{\sum_{j=k+1}^{p} \prod_{i=0}^{k-1} (j-i)\beta_j x^{j-k}}_{=0} \\ &= \prod_{i=0}^{k-1} (k-i)\beta_j x^{k-k} \\ &= k!\beta_k \end{aligned} \tag{8}$$

Replacing the result from (8) in (6), gives

$$\begin{aligned} m(x) = m(x_0) + \sum_{k=1}^{p} \frac{1}{k!} k! \beta_k (x - x_0)^k = \sum_{k=0}^{p} \beta_k (x - x_0)^k \\ = m(x_0) + \beta_1 (x - x_0) + \beta_2 (x - x_0)^2 + \cdots + \beta_p (x - x_0)^p \end{aligned} \tag{9}$$

We have kept low as possible the order of the polynomial due to the reason that the arbitrary fitting of higher order polynomial can increase of risk of over-fitting and to generate bad prediction of model (1)[13].

Putting (3) and (9), in model (2), we have

$$\boldsymbol{\zeta} = \left\{ l_i - [m(x_0) + \beta_1 (x - x_0) + \beta_2 (x - x_0)^2 + \cdots + \beta_p (x - x_0)^p] \right\}^2 w_i(x_0). \tag{10}$$

By minimize (10), we have

$$\boldsymbol{\zeta} = \sum [l_i - m(x_0) - \beta_1 (x - x_0) - \beta_2 (x - x_0)^2 - \cdots - \beta_p (x - x_0)^p]^2 w_i(x_0) \tag{11}$$

The (11) can be written as

$$\boldsymbol{\zeta} = (\mathbf{L} - \mathbf{X}\boldsymbol{\beta})^T \mathbf{W} (\mathbf{L} - \mathbf{X}\boldsymbol{\beta}) \tag{12}$$

where

$$\begin{aligned} L &= \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{bmatrix} \\ X &= \begin{bmatrix} 1 & (x_1 - x_0)^1 & (x_1 - x_0)^2 & \dots & (x_1 - x_0)^p \\ 1 & (x_2 - x_0)^1 & (x_2 - x_0)^2 & \dots & (x_2 - x_0)^p \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & (x_n - x_0)^1 & (x_n - x_0)^2 & \dots & (x_n - x_0)^p \end{bmatrix} \\ \beta^T &= \begin{bmatrix} \beta_0 & \beta_1 & \beta_2 & \dots & \beta_p \end{bmatrix} \end{aligned}$$

$$\mathbf{W} = \begin{bmatrix} t\left(\frac{|x_0 - x_1|}{b}\right) & 0 & 0 \dots & 0 \\ 0 & t\left(\frac{|x_0 - x_2|}{b}\right) & 0 \dots & 0 \\ \vdots & \vdots & \vdots \ \ddots & \vdots \\ 0 & 0 & 0 \dots & t\left(\frac{|x_0 - x_n|}{b}\right) \end{bmatrix}$$

Then from (12), we have

$$\boldsymbol{\zeta} = (\mathbf{L}^T\mathbf{W}\mathbf{L} - \mathbf{L}^T\mathbf{W}\mathbf{X}\boldsymbol{\beta} - \boldsymbol{\beta}^T\mathbf{X}^T\mathbf{W}\mathbf{L} + \boldsymbol{\beta}^T\mathbf{X}^T\mathbf{W}\mathbf{X}\boldsymbol{\beta}) \tag{13}$$

First derivative of (13) with respect of $\boldsymbol{\beta}$ gives

$$\frac{\partial\boldsymbol{\zeta}}{\partial\boldsymbol{\beta}} = -2\mathbf{X}^T\mathbf{W}\mathbf{L} + 2\boldsymbol{\beta}\mathbf{X}^T\mathbf{W}\mathbf{X} \tag{14}$$

Since $\mathbf{W}$ is a diagonal matrix and evaluating (14) at point 0, we have

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}\mathbf{L} \tag{15}$$

Given the condition that the $(\mathbf{X}^T\mathbf{W}\mathbf{X})$ is a nonsingular matrix. The quantity $m(x_0)$ is then estimated by the fitted intercept parameter (ie. by β_0) as this defines the position of the estimated local polynomial curve at the point x_0[13].

4 Asymptotic Properties

4.1 Deriving the Asymptotic bias

Theorem 4.1. ***Proposed theorem****: Let $(x_1; l_1), \cdots, (x_n; l_n)$ be a random sample of bivariate data and x be a close point to x_0 then $\widehat{m}(x)$ is the consistency estimator of $m(x_0)$ when the smoothing parameter satisfy the condition $b \to 0$, as $n \to \infty$, then*
$\widehat{m}(x) \xrightarrow{D} m(x_0)$.

Proof. Here, prove properties for $p = 0$ (Nadaya-Watson kernel estimation). According to [12] the mean can be approximated as follows:

$$\mathbb{E}[\widehat{m}(x_0)] = \frac{\mathbb{E}\left[\sum_{i=1}^{n} t\dfrac{(x_i - x_0)}{b} l_i\right]}{\mathbb{E}\left[\sum_{i=1}^{n} t\dfrac{(x_i - x_0)}{b}\right]} = \frac{A}{B} \tag{16}$$

1.Assuming that (x_i, l_i) are independent identically distributed, then

$$A = n\mathbb{E}\left[t\frac{(x - x_0)}{b} l\right] \tag{17}$$

$$A = n \iint t \frac{(x - x_0)}{b} l f(x, l)\, dx\, dl \tag{18}$$

Changing the variables $s = \dfrac{(x - x_0)}{b}$, then

$$A = nb \iint t(s) l f(sb + x_0, l)\, ds\, dl \tag{19}$$

where

$$f(sb + x_0, l) = f(l|sb + x_0) f(sb + x_0) \tag{20}$$

Using series expansion

$$\begin{aligned} f(sb + x_0) &= f(x_0) + f'(x_0)sb + \frac{1}{2} f''(x_0)s^2b^2 + \mathbf{O}(s^2b^2) \\ m(sb + x_0) &= (x_0) + m'(x_0)sb + \frac{1}{2} m'(x_0)s^2b^2 + \mathbf{O}(s^2b^2) \end{aligned}$$

2.Assuming that f and m are differentiable

$$\begin{aligned} A \;=\; & nb\Big[f(x_0)m(x_0) \int (ts)\, ds + \frac{1}{2} f(x_0)m''(x_0)b^2 \int s^2 t(s)\, ds + \\ & + f'(x_0)m'(x_0)b^2 \int s^2 t(s)\, ds + \frac{1}{2} f'(x_0)m(x_0)b^2 \int s^2 t(s)\, ds \Big] + \mathbf{O}(s^2b^2) \end{aligned} \tag{21}$$

1.Notation $\sigma_t^2 = \int s^2 t(s)\, ds$ that is Roughness of t

$$\begin{aligned} A \;=\; & nb \left[f(x_0)m(x_o) + \frac{1}{2} f(x_0)m''(x_0)b^2\sigma_t^2 + f'(x_0)m'(x_0)b^2\sigma_t^2 + \frac{1}{2} f''(x_0)m(x_0)b^2\sigma_t^2 \right] \\ & + \mathbf{O}(s^2b^2) \end{aligned} \tag{22}$$

$$\begin{aligned} A \;=\; & nb \left[f(x_0)m(x_0) + \frac{b^2}{2}\sigma_t^2 (f(x_0)m'(x_0) + 2f'(x_0)m'(x_0) + f''(x_0)m(x_0)) \right] \\ & + \mathbf{O}(s^2b^2) \end{aligned} \tag{23}$$

Now, we have to calculate

$$B = \mathbb{E}\left[\sum_{i=1}^{n} t \frac{(x_i - x_0)}{b} \right] \tag{24}$$

Using assumption 1 and changing the variable $s = \dfrac{(x - x_0)}{b}$, we have

$$B = nb \int t(s) f(sb + x_0)\, ds$$

Using the Taylor series expansion, we have

$$B = nb \int t(s)[f(x_0) + f'(x_0)sb + \frac{1}{2} f''(x_0) s^2 b^2]\, ds + \mathbf{O}(s^2 b^2) \tag{25}$$

$$B \approx nf(x_0) \int t(s)\, ds + b^2 f(x_0) \int st(s)\, ds + \frac{1}{2} b^2 f''(x_0) \int s^2 t(s)\, ds \tag{26}$$

Application of assumption 2, we have

$$B \approx nbf(x_0) \left[1 + \frac{1}{2} b^2 \frac{f''(x_0)}{f(x_0)} \sigma_t^2 \right] \tag{27}$$

$$\frac{A}{B} = \frac{m(x_0) + \dfrac{b^2}{2} \sigma_t^2 \left[\dfrac{f(x_0) m'(x_0)}{f(x_0)} \right] + \dfrac{2 f'(x_0) m'(x_0)}{f(x_0)} + \dfrac{f''(x_0) m(x_0)}{f(x_0)}}{1 + \dfrac{1}{2} b^2 \left[\dfrac{f''(x_0) m(x_0)}{f(x_0)} \right] \sigma_t^2} \tag{28}$$

4.Assuming that $b \to 0$, so

$$\begin{gathered} \frac{f''(x_0)}{f(x_0)\sigma_t^2} \approx \left(1 - \frac{1}{2} b^2 \frac{f''(x_0)}{f(x_0)} \sigma_t^2 \right)^{-1} \\ B = 1 + \frac{1}{2} b^2 \frac{f''(x_0)}{f(x_0)\sigma_t^2} \approx \left(1 - \frac{1}{2} b^2 \frac{f''(x_0)}{f(x_0)} \sigma_t^2 \right)^{-1} \end{gathered} \tag{29}$$

$$\begin{aligned} \mathbb{E}[\widehat{m}(x_0)] &\approx m(x_0) - m(x_0) \frac{b^2}{2} \sigma_t^2 \frac{f''(x_0)}{f(x_0)} + \frac{b^2}{2} \sigma_t^2 m''(m_0) + 2 \frac{f'(x_0) m'(x_0)}{f(x_0)} \\ &\quad + \frac{f''(x_0) m(x_0)}{f(x_0)} \end{aligned} \tag{30}$$

$$\mathbb{E}[\widehat{m}(x_0)] \approx m(x_0) + \frac{b^2}{2} \sigma_t^2 \left[m''(x_0) + \frac{f'(x_0) m'(x_0)}{f(x_0)} \right] \tag{31}$$

$\mathbb{E}[\widehat{m}(x_0) - m(x_0)] \to 0$ when $b \to 0$ and $n \to \infty$

$$\widehat{m}(x_0) \xrightarrow{D} m(x_0) \tag{32}$$

□

4.2 Deriving the asymptotic variance

Theorem 4.2. ***Proposed theorem***: *Let $(x_1; l_1), \cdots, (x_n; l_n)$ be a random sample of bivariate data and x be a close point to x_0 then $Var[\widehat{m}(x_0)] \approx \dfrac{1}{f(x_0)}$ provided that $f(x_0)$ is positive, as $b^2 \to 0$.*

Proof. We that the variance can be obtained as:

$$Var[\widehat{m}(x_0)] = \mathbb{E}[\widehat{m}(x_0)^2] - \{[\mathbb{E}[\widehat{m}(x_0)]\}^2 \tag{33}$$

$$\mathbb{E}[\widehat{m}(x_0)] = \frac{\sum_{i=1}^{n} t\dfrac{(x_i - x_0)}{b} l_i}{\sum_{i=1}^{n} t\dfrac{(x_i - x_0)}{b}} = \frac{A}{B} \tag{34}$$

$$\mathbb{E}[\widehat{m}(x_0)]^2 = \frac{\mathbb{E}[A^2]}{\mathbb{E}[B^2]} \tag{35}$$

$$\mathbb{E}(A^2) = \mathbb{E}\left[\left(\sum_{i=1}^{n} t\frac{(x_i - x_0)}{b} l_i\right)\left(\sum_{i=1}^{n} t\frac{(x_i - x_0)}{b} l_i\right)\right] \tag{36}$$

Using the assumption 1, we have

$$\mathbb{E}(A^2) = n \times \mathbb{E}\left[t^2\frac{(x - x_0)}{b} l^2\right] \tag{37}$$

$$\mathbb{E}(A^2) = n \times \iint t^2\frac{(x - x_0)}{b} l^2 f(x,y)\, dx\, dy \tag{38}$$

Changing the variables put $s = \dfrac{x - x_0}{b}$, we get

$$\mathbb{E}(A^2) = nb \int t^2(s) f(sb + x_0)[m^2(sb + x_0) + 1]\, ds \tag{39}$$

Using series expansion first order,

$$\begin{aligned} f(sb + x_0) &= f(x_0) + f'(x_0)sb + \mathbf{O}(s^2b^2) \\ m^2(sb + x_0) &= m^2(x_0) + (m^2)'(x_0)sb + \mathbf{O}(s^2b^2) \end{aligned}$$

Now, we have

$$\mathbb{E}(A^2) = nb \int \left\{t^2(s)\left[f(x_0) + f'(x_0)sb\right]\left[m^2(x_0) + (m^2)'(x_0)sb + 1\right] + O(s^2b^2)\right\}\, ds \tag{40}$$

4.Assuming that $\int st^2\, ds$ is negligible we have

$$\mathbb{E}(A^2) \approx nb\sigma_t^2\left[f(x_0)m^2(x_0) + 1\right] \tag{41}$$

On the other hand, we have

$$\mathbb{E}(B^2) = \mathbb{E}\left[\sum_{i=1}^{n} t\frac{(x_i - x_0)}{b}\right] \tag{42}$$

Using the assumption 1, we get

$$\mathbb{E}(B^2) = n\int t^2\left(\frac{x - x_0}{b}\right) f(x)\, dx \tag{43}$$

Changing the variables put $s = \dfrac{x - x_0}{b}$, we have

$$\mathbb{E}(B^2) = nb\int t^2(s)f(x_0) + f'(x_0)sb + \mathbf{O}(sb^2)\, ds \tag{44}$$

then,

$$\mathbb{E}(B^2) \approx nbf(x_0)\sigma_t^2 \tag{45}$$

Using(41) and (45), then (35) gives

$$\begin{aligned}\mathbb{E}[\widehat{m}(x_0)^2] &\approx \frac{nb\sigma_t^2\left[f(x_0)m^2(x_0) + 1\right]}{nbf(x_0)\sigma_t^2}\\ \mathbb{E}[\widehat{m}(x_0)^2] &\approx m^2(x_0) + \frac{1}{f(x_0)}\end{aligned} \tag{46}$$

Using (46) and (31) in (33), we have

$$Var[\widehat{m}(x_0)] \approx m^2(x_0) + \frac{1}{f(x_0)} - \left[m(x_0) + \frac{b^2}{2}\sigma_t^2 m''(x_0)\right]^2 \tag{47}$$

$$Var[\widehat{m}(x_0)] = \frac{1}{f(x_0)} - m(x_0)b^2\sigma_t^2 m''(x_0) - \mathbf{O}\left(\frac{b^4}{4}\right) \tag{48}$$

$$Var[\widehat{m}(x_0)] \approx \frac{1}{f(x_0)} \quad \text{as } b^2 \to 0 \text{ and given that } f(x_0) > 0 \tag{49}$$

$\square$

5 Simulation Study

We evaluate the performance of our estimator though out the simulation. The model (1) was used for simulation, with one-dimensional covariate X_{jd} was generated form normal distribution. The variable of interest y_{jd} were simulated from normal distribution. The mean and variance used for the simulation were approximately calculated based on the (2013-2014) Rwanda Integrated Household Living Conditions Survey(EICV4). In our simulation, $N = 10000$ observations were generated and the small area population sizes $N_i \quad i, = 1, \cdots, m = 25$, were generated randomly so that $\sum_i N_i = N$ and was kept fixed throughout the simulation. The random errors were independently generated from a normal distribution with parameter $\mathcal{N}(0, \sigma_e^2)$. The random area effects v_j were generated from $\mathcal{N}(0, \sigma_u^2)$. The data plot is represented by the Figure 1.

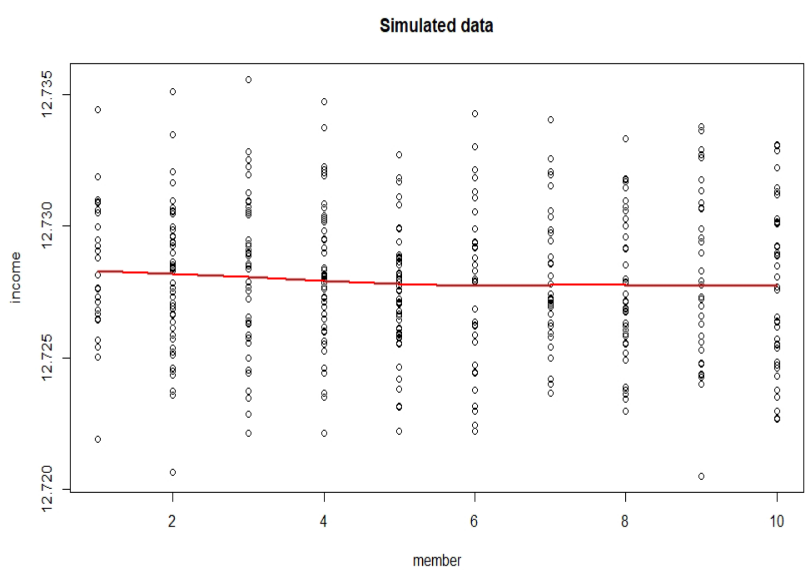

Figure 1: Plot of the simulated data and its fitting

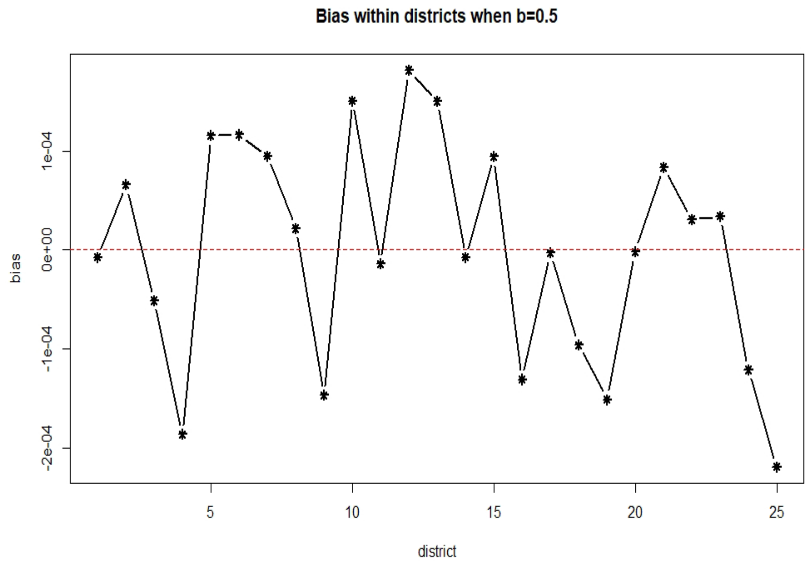

Figure 2: The distribution of bias within districts at $b = 0.5$.

6 Conclusion

Considering that the normality assumptions were violated the proposed model could work as an alternative in small area estimation. Since, our estimator was found to give smaller MSE and Bias almost close to zero when the bandwidth is equal to 0.5. Despite, the good performance of our estimator more research is needed to derive its optimal bandwidth.

Acknowledgments.
The authors acknowledge African Union and Pan African University, Institute for Basic Sciences Technology and Innovation, Kenya for funding and supporting this research project.

Table 1: Calculated MSE and Bias for simulated data at different bandwidth

Bandwidth	**MSE**	**Bias**
0.01	2.212724e-07	7.602842e-17
0.1	2.212724e-07	7.602842e-17
0.3	1.395272e-07	7.602842e-17
0.5	6.800175e-08	1.012576e-17
0.8	1.395272e-07	-2.540015e-17

The MSE and Bias were found to be close to zero when the bandwidth is 0.5. For the rest of our analysis a bandwidth($b = 0.5$) was used to tackle the issue of over and under estimation.

References

[1] Opsomer, J. D., Claeskens, G., Ranalli, M. G., Kauermann, G., and Breidt, F.J., Non-parametric small area estimation using penalized spline regression, *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, **70**(1), (2008), 265-286.

[2] Pratesi, M., Ranalli, M. G., and Salvati, N., Nonparametric m-quantile regression using penalised splines, *Journal of Nonparametric Statistics*, **21**(3), (2009), 287-304.

[3] Lombardia, M. J. and Sperlich, S., Semiparametric inference in generalized mixed effects models, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **70**(5), (2008), 913-930.

[4] Pratesi, M., Ranalli, M. G., and Salvati, N., Semiparametric m-quantile regression for estimating the proportion of acidic lakes in 8-digit hucs of the northeastern us. *Environmetrics*, **19**(7), (2008), 687-701.

[5] Salvati, N., Ranalli, M., and Pratesi, M., Small area estimation of the mean using non-parametric m-quantile regression: a comparison when a linear mixed model does not hold, *Journal of Statistical Computation and Simulation*, **81**(8), (2011), 945-964.

[6] Battese, G. E., Harter, R. M., and Fuller, W. A., An error-components model for prediction of county crop areas using survey and satellite data, *Journal of the American Statistical Association*, **83**(401), (1988), 28-36.

[7] Tzavidis, N., Salvati, N., Pratesi, M., and Chambers, R., M-quantile models with application to poverty mapping, *Statistical Methods and Applications*, **17**(3), (2008), 393-411.

[8] Berg, E., Chandra, H., and Chambers, R., *Small Area Estimation for Lognormal Data*, pages 279-298. John Wiley Sons, Ltd 2016.

[9] Cleveland, W. S. and Devlin, S.J., Locally weighted regression: an approach to regression analysis by local fitting, *Journal of the American statistical association*, **83**(403), (1988), 596-610.

[10] Fan, J. and Gijbels, I., *Local polynomial modelling and its applications: monographs on statistics and applied probability* 66, volume 66. CRC Press, 1996.

[11] National Institute of Statistics of Rwanda, N., Rwanda poverty profile report, 2013/14, (2015)

[12] Seltman, H. Approximations for mean and variance of a ratio, (2012), unpublished note.

[13] Irizarry, R. A. and Bravo, H.C., Lecture 7, Smoothing, (2010).

[14] L. L. Seknewna, P. N. Mwita and B. K. Muema, Smoothed conditional scale function estimation in AR(1)-ARCH(1) processes, *Journal of Probability and Statistics*, 2018, (2018).