Pak. J. Statist.
2023 Vol. 39(3), 299-312

NEURAL NETWORK CALIBRATION ESTIMATION WITH HIGH DIMENSIONAL DATA IN SURVEY SAMPLING: EVIDENCE USING MONTE CARLO SIMULATION SCHEME

Laraib Brirah[1], Haris Khurram[2], Muhammad Ahmed Shehzad[1§] and **Aamna Khan[1]**

[1] Department of Statistics, Bahauddin Zakariya University Multan, Pakistan. Email: laraibbrirah786@gmail.com aamnaa@bzu.edu.pk

[2] National University of Computer and Emerging Sciences Chiniot-Faisalabad Campus, Pakistan. Email: haris.khurram@nu.edu.pk

[§] Corresponding author Email: ahmad.shehzad@bzu.edu.pk

ABSTRACT

Calibration is commonly used in survey sampling to include auxiliary information at the estimation stage. Calibrating the observation weights on the population means (or totals) of the auxiliary variables implicitly assumes on a linear superpopulation regression model. When auxiliary information is available for all units the population, more complex modeling can be handled by means of model calibration .This article explores the estimation of finite population totals when auxiliary data is included, either univariate or multivariate. In this work we introduce a new type of model calibration nonparametric estimator for the finite population mean based on neural network learning. The proposed neural network model calibration estimators can handle any linear or nonlinear working models. More precisely, we adopt neural network learning to estimate the functional relationship between the survey variable and the auxiliary variables. Under suitable regularity conditions, the proposed estimators are proven to be design consistent. The performance of the proposed estimators for finite-size samples is investigated by means of simulation studies. The uniqueness of our work is that we observed the result by single auxiliary variable and also by increasing No. of parameters and show the efficient estimator.

KEYWORDS

Survey sampling; Auxiliary information; Generalized regression estimator; Calibration; Model-assisted approach; Neural networks; Nonparametric regression.

Highlights of this Paper:

- We combine artificial neural network learning approach and calibration technique to estimate the total of survey variable by using auxiliary information.
- The aim is to improve the estimate of totals when data have some linear and non-linear relationship.
- The presented paper is the part of this project in which, some theoretical and mathematical justification and aspect of the proposed approach is presented.

© 2023 Pakistan Journal of Statistics

1. INTRODUCTION

In research, sampling is very helpful. Survey sampling consists of selecting a representative sample of a population for use as part of a sample survey. The term "auxiliary variables" is most frequently used in relation to the use of such variables, which are accessible to all units in the population, in ratio estimation, regression estimation, and calibration estimation. Auxiliary variables are variables about which information is available prior to sampling. Auxiliary information helps survey statisticians improve estimates in many ways. A family of estimators called calibration estimators uses a similar collection of auxiliary data to generate estimates. It is important to rely on auxiliary information to obtain a better estimate of a population statistic. Calibration estimator is now widely used in worldwide statistical surveys. The estimator improves estimates by using data from auxiliary variables. Wu and Sitter (2001) introduced model calibration, where nonlinear parametric regression models and generalized linear regression models are used to obtain model-assisted estimators by generalizing the calibration method of Deville and Särndal (1992). In this article we extend model calibration by assuming more general models than those suggested by Wu and Sitter (2001) and employ neural networks to obtain the fitted values to calibrate on. This allows more flexible prediction and straightforward insertion of multivariate auxiliary information.

2. CALIBRATION ESTIMATION

Calibration approach first developed by Deville and Särndal (1992), Wu and Sitter (2001) extend the concept of model calibration to generalized linear regression model and non-parametric regression models. According to the Montanari and Ranalli (2005), the calibration procedure is based on linear relationship among the survey variables y and the auxiliary variables $\boldsymbol{x}$. A calibration approach is a weighting approach that uses auxiliary information to calculate what will match the known population value. The model calibration is extended to include more general super-population models, and we acquire the fitted values for calibration using neural network approach. Using neural network learning, we determine the functional relationship among the auxiliary variables x and the survey variables y. Although neural networks have been utilized for survey sampling and imputation, their use for model calibration is novel. Neural network approach has been shown to greatly enhance prediction of the value in a relevant variable in non sampling units. This feature, in particular, improves the efficiency of the generated estimators when the underlying functional relationship is somewhat complicated. For a finite population mean, we propose neural network model calibration estimators.

2.2 Notations and Equations of Calibration Technique

Let us considered a sets of N units from a finite population U labeled as $U = \{1, \dots, N\}$. The population (U) size is not always known. From which the probability sample $s(s \subseteq \boldsymbol{U})$, Q auxiliary variables are represented by a vector x is known for every unit in the population, for instance through census data, administrative registers, remote sensing, or prior surveys so the vector is represent as $\boldsymbol{x_k} = \left(x_{k1}, \dots, x_{kq} \dots, x_{kQ}\right)'$, the auxiliary vector is known as $\forall_i \in U$.

Let's say that a sample is taken from a population without replacement and the population $U = \{1, \dots, N\}$ with a size sampling design $P(.)$ this is fixed size design and a known inclusion probability. A random sample represented by s of fixed size n taken from U in accordance with probability sample plan with inclusion probabilities as $\pi_k = \Pr(k \in s)$ satisfy that $\pi_k > 0$, is known as 1st order inclusion probability and $\pi_{kl} = \Pr(k, l \in s)$ is known as a 2nd order inclusion probability all $k, l \in U$. The study variable y_k is the value of variable interest of y for k_{th} elements of the population which additionally has a vector value for an auxiliary variable, and y_k is known for all $k \in s$. Here n is number of observation and number of auxiliary variables is represented by Q.

According to Deville and Särndal (1992) the calibration of population total of t_y is,

$$\hat{t}_y = \sum_s w_k y_k \tag{1}$$

All chi-squared distance estimators result in asymptotically identical results when the distance is calculated using alternative distance measurements in Deville and Särndal (1992). The average distance $E_p\left\{\frac{\sum_s (w_k - d_k)^2}{d_k q_k}\right\}$. The distance in equation between the initial weight d_k and the new weight w_k was quite arbitrary chosen as

$$\Phi_s(w) = \frac{\sum_s (w_k - d_k)^2}{d_k q_k} \tag{2}$$

where $w = (w_k, k \in s)$ each unit in the sample is given a vector of weights, and the $q_k's$ are well-known positive constants that may be used to account for the variability of the observations and are independent of d_k similarly, $1/q_k =$ Positive weight unrelated to d_k and the uniform weights $1/q_k = 1$ or $q_k = 1$ for all units k.

Thus, the resulting estimator of t_y is

$$\hat{t}_{y,w} = \hat{t}_{y,\pi} + \left(\boldsymbol{t}_x - \hat{\boldsymbol{t}}_{x,\pi}\right)' \widehat{\boldsymbol{B}}_{\boldsymbol{s}} \tag{3}$$

where $\hat{t}_{x,\pi} = \sum_s d_k \boldsymbol{x}_k$ denotes the H.T (Horvitz-Thompson) estimator for the auxiliary information of vector $\boldsymbol{x}$ and $\widehat{\boldsymbol{B}}_{\boldsymbol{s}} = \boldsymbol{T}_s^{-1} \sum d_k q_k \boldsymbol{x}_{\boldsymbol{k}} y_k$. Here $\widehat{\boldsymbol{B}}_{\boldsymbol{s}}$ is represent the weighted estimator for the multiple regression coefficients. As a result, $\hat{t}_{y,w}$ implicitly assumes that the auxiliary variables and the survey variable have a linear connection. Wu and Sitter (2001) suggest taking into account more complex models and generalizing calibration process by model calibration by noting that "it is the relationship between y and x, hopefully captured by the working model that determines how the auxiliary information should best be used."

3. ARTIFICIAL NEURAL NETWORK (ANN) MODEL

The first mathematical model of a neuron was established by Warren McCulloch and Walter Pitts in 1943. They establish the fundamental mathematical model for neuron, which symbolizes a single cell of brain system that takes input, analyses those inputs and produces an output. A neural network learns from training data and increases its accuracy over time. By using concepts of biological neuron network, an artificial neuron network

(ANN) mimics the organization human brain. ANNs. The neurons are interconnected and exchange signals. The nodes may take data and do straightforward operations to do it. Other neurons get the outcome of these processes. The activation or node value is its output. A neural network has three or more interconnected layers. Input neurons make up the top layer. These neurons communicate with the deeper layer, which communicate with the top output layer. The hidden layers are the terms for the middle layers.

Mc-Culloch and Pitts (1943) identified the neuron as a binary threshold unit

$$y = f(x_1.w_1 + x_2.w_2 + \cdots + x_n.w_n \ + b) = f\left\{\sum_{i=1}^{n}(x_k w_k + b)\right\} \tag{4}$$

where f is activation function and b is denoted as bias. Here x denotes the input variables and y denotes the output variable and w denote the weights. This computation is expressed as a transfer function.

$$\sum_{i=1}^{n} w_i X_i + b. \tag{5}$$

4. THE NEURAL NETWORK CALIBRATION ESTIMATOR

In 1996, Nordbotten employed auxiliary data from administrative registries along with neural networks for imputation. The use of neural networks calibration is novel, allowing enabling more accurate prediction and easier addition of multivariate auxiliary data. Neural network, among other things, are a highly well-liked learning technique. Although multivariate data may be handled by the programmed quickly and rapidly, neural networks are frequently employed in practice. In such a model, there are three components: inputs, outputs, and hidden variables-neurons-which transform the information coming from inputs into outputs in a nonlinear fashion. Each connection is weighted. Each hidden unit receives linear connections of inputs as its input; In order to send signals to the output and an activation function called $\varphi(.)$ is employed. The final output is created by adding a further constant to a linear combination of these signals.

Assume that the following superpopulation model can adequately represent connection among survey variable *y* and auxiliary variables *x* using first and second moments,

$$E_\xi(y_k) \ = y_{nn}(x_k) \; for \; k = 1,2,3, \dots, N$$

$$V_\xi(y_k) = V(x_k) \; for \; k = 1,2,3, \dots, N$$

$$C_\xi(y_k \, , y_l) = 0 \; for \; k \neq l$$

Here V_ξ denotes variance and C_ξ denotes covariance with respect to super population models. We suppose that $(y_1, x_1), \dots, (y_k, x_k)$ are mutually independent.

Based on these definitions, we can estimate the calibration of the neural network model for t_y

$$\hat{t}_{y,nn}^{mc} = \sum\nolimits_{k=1}^{n} y_k w_k$$

Here w_k represent the calibrated weights are aimed at minimizing the distance function in $\varphi_s = \frac{\sum_s (w_k - d_k)^2}{d_k q_k}$ under the constraints, $\sum_{k=1}^{n} w_k = 1$ and $\sum_{k=1}^{n} w_k \, \hat{y}_{nn} = \sum_{k=1}^{n} \hat{y}_{nn}$. The proposed estimator was as a result of applying the method of Deville and Sarndal (1992) to generate the ideal weights as

$$\hat{t}_{y,nn}^{mc} = \hat{t}_{y,\pi} + \left\{ \sum_{k=1}^{N} \hat{y}_{nn} - \sum_{k=1}^{n} d_k \, \hat{y}_{nn} \right\} \hat{B}_{nn} \tag{6}$$

Here

$$\hat{B}_{nn} = \frac{\sum_{k=1}^{n} d_k \, q_k (\hat{y}_{nn} - \check{y}_{nn})(y_k - \check{y})}{\sum_{k=1}^{n} d_k q_k \, (\hat{y}_{nn} - \check{y}_{nn})^2},$$

$$\check{y}_{nn} = \frac{\sum_{k=1}^{n} d_k q_k \, \hat{y}_{nn}}{\sum_{k=1}^{n} d_k q_k},$$

$$\check{y} = \frac{\sum_{k=1}^{n} d_k q_k y_k}{\sum_{k=1}^{n} d_k q_k},$$

$$\hat{y}_{nn} = \frac{\sum_{k \in s} d_k q_k \hat{y}_{nn}}{\sum_{k \in s} d_k q_k}$$

Considered

$$X = \hat{y}_{nn} - \frac{\sum_{k \in s} d_k q_k \hat{y}_{nn}}{\sum_{k \in s} d_k q_k},$$

$$Y = y_k - \frac{\sum_{k=1}^{n} d_k \, q_k y_k}{\sum_{k=1}^{n} d_k q_k}$$

then

$$\hat{\beta}_{nn} = \frac{\sum_{i=1}^{n} d_k q_k XY}{\sum_{i=1}^{n} d_k q_k (X)^2}$$

Based on the working model, $\hat{t}_{y,nn}^{mc}$ may be thought of the working model can be viewed as a generalized regression estimator is represent as $E_{\xi}(y_k) = \alpha + \beta f(\boldsymbol{x}_k)$. Here $\hat{t}_{y,nn}^{mc}$ uses estimates of $y_{nn}(\boldsymbol{x}_k)$ as auxiliary variable in a generalized regression procedure.

4.1 Important Properties and Assumptions of $\hat{t}_{y,nn}^{mc}$

The important design properties of $\hat{t}_{y,nn}$, the Taylor series of the fitted values of $\hat{y}_{nn}$ will be used. In order to do this, a set of regularity requirements based on the parameters $\hat{\theta}$ & $\tilde{\theta}$ and the function $\hat{y}_{nn}(.)$ in an asymptotic framework is required. We take it for granted that there is a list of sampling designs and finite populations, all of which are indexed by the letter v. The population size and the sample size respectively N_v & n_v approach to infinity as $v \to \infty$.

We rely on the following assumptions to justify our theoretical findings.

a) Firstly, the errors of $\varepsilon_i$are iid and have mean zero and variance $V(\boldsymbol{x_k})$, uniformly for all N.
b) Considered $\boldsymbol{x_k}$, each N has a fixed $\boldsymbol{x_k}$ in relation to the supperpopulation model ξ **.**
c) For each v, $\boldsymbol{x_k}$ is identically and independently distributed from fixed and an unknown distribution for $F(x) = \int_{-\infty}^{xk} f(t_1, t_2, \dots, t_Q) d\boldsymbol{t}$, where $f(\cdot)$ is a density which is strictly positive.
d) The survey variables are bound with a probability of 1 at the fourth moment with ξ.
e) The design rates bounded that is $limsup\ nN^{-1}\ as\ v \to \infty =\ \pi$ here $\pi\ \in (0, 1)$.
f) The sampling design $p(s)$ ensures that the Horvitz-Thompson estimator of the population total t_y is design consistent and asymptotically normally distributed with variance $O(n^{-1})$. Using the Horvitz-Thompson variance estimator, we can measure the variance of any study variable y with a bounded fourth moment.
g) φ is an activation function in (3.12), in other words, it is a symmetric sigmoid function that can be differentiated into any of order, is also assumed to be linearly independent for the class of functions $\{\varphi(bt + b_0), b > 0\} \cup \{\varphi \equiv 1\}$. These constraints are met by the logistic activation function $\varphi(t) = [1 + \exp(-t)]^{-1}$; in 1997, Hwang and Ding (1997) provide more examples of sigmoidal functions fulfilling these requirements.

Theorems require the conditions listed below. The proof of Theorems contains descriptions of some of the notations used here.

i. $\hat{\theta} = \tilde{\theta} + O_p(n^{-1/2})$ And $\tilde{\theta} \to \theta$
ii. For each $\boldsymbol{x_i}$, $|\partial f(\boldsymbol{x_i}, \theta)/\partial\theta| \le h(\boldsymbol{x_i}, \theta)$ for θ in a neighborhood of θ and $\sum_{k=1}^{N} h(\boldsymbol{x_i}, \theta) = O(1)$
iii. The essential design weights, $d_k = \frac{1}{\pi_k}$, fulfill the asymptotically normal distribution of the Horvitz-Thompson estimator for certain population totals.

4.1.1 Asymptotically Design-Unbiased Estimators*:*

The proposed estimator is Asymptotically Design-Unbiased Estimators as shown by Theorem 1.

Theorem 1: Asymptotically Design-Unbiased Estimators

Suppose the working-model used to create the estimators are represent above as super-population models by using condition (i) to (iii) describe above, $\boldsymbol{\hat{t}_{y,nn}}$ is equal to $\boldsymbol{\hat{t}_{y\pi}} + \boldsymbol{O_p(n^{-1/2})}$ and are thus asymptotically design-unbiased estimators for $\boldsymbol{t_y}$. This is also approximately model-unbiased under condition (i) see Wu and sitter (2001).

4.1.2 Design Consistency*:*

The proposed estimator is also design consistence as shown in Theorem 2

Theorem 2 (Montanari and Ranalli 2005) $\boldsymbol{\hat{t}_{y,nn}^{mc}}$ is consistence for $\boldsymbol{t_y}$ in the logic that $\boldsymbol{lim_{v\to\infty} P(|\hat{t}_{y,nn}^{mc} - t_y| < \in) = 1}$ with $\boldsymbol{\xi - \text{probablity}}$ 1 for any fixed $\in > \boldsymbol{0}$

4.1.3 Asymptotic Normality:

The proposed estimator also asymptotically normal as shown in Theorem 3

Theorem 3:

The asymptotic distribution of $\hat{\boldsymbol{t}}_{\boldsymbol{y,nn}}$ is as under

$$\frac{\hat{t}_{y,nn}^{mc} - t_y}{\sqrt{V\left(\tilde{t}_{y,nn}^{mc}\right)}} \to N(0,1) \tag{7}$$

where $\tilde{t}_{y.nn}^{mc}$ is the population variance of artificial neural network calibration estimator

$$\tilde{t}_{y,nn}^{mc} = \hat{t}_{y,\pi} + \left\{\sum_{k=1}^{N} \tilde{y}_{nn} - \sum_{k=1}^{n} d_k\, \tilde{y}_{nn}\right\} \tilde{\beta}_{nn} \tag{8}$$

With

$$\tilde{\beta}_{nn} = \frac{\sum_{k=1}^{N} q_k\, (\tilde{y}_{nn} - \bar{y}_{nn})(y_k - \bar{y})}{\sum_{k=1}^{N} q_k\, (\tilde{y}_{nn} - \bar{y}_{nn})^2} \tag{9}$$

And $\bar{y}_{nn} = \sum_{k=1}^{N} \tilde{y}_{nn}$, whose design variance is given by

$$V\left(\hat{t}_{y,nn}^{mc}\right) = \sum_{k}^{N} \sum_{l}^{N} \left(\pi_{k,l} - \pi_k \pi_l\right) \frac{E_k}{\pi_k} \frac{E_l}{\pi_l} \tag{10}$$

where $\pi_{k,l}$ are second order inclusion probabilities, $E_k = y_k - \tilde{y}_{nn}\, \tilde{\beta}_{nn}$ further detail available in (Montanari & Ranalli (2005)).

5. SIMULATION STUDIES

We carry out a Monte Carlo simulation analysis to examine the effectiveness of the provided estimators for estimating the total population. The concept and organization of our inquiry be inspired by Breidt and Opsomer's (2000) simulation study, which focused on a single auxiliary and allowed for comparisons. However, certain elements have also been modified and added to offer fresh perspectives on the subject. The simulation research examines the behaviors of the following t_y estimators of population totals.

1. Horvitz-Thompson Estimator ($\hat{t}_{y\pi}$)
2. Calibration Estimator ($\hat{t}_{y,w}$)
3. Neural Network Model calibration ($\hat{t}_{y,nn}^{mc}$)

The third estimator, which enables more intricate modelling of the regression function, joins the preceding two parametric estimators. Four separate models were used to construct the survey variables, and he uni-variate regression function or signal of each of these models is unique. We considered the following regression models

Linear: $y_1(x) = 1 + 2(x - 0.5)$
Quadratic: $y_2(x) = 1 + 2(x - 0.5)^2$
Exponential: $y_3(x) = exp(-8x)$
Cycle 1: $y_4(x) = 2 + \sin(2\,\pi\,x)$

With $x \in [0,1]$. The selection of such signals where the population x is generated as independently identically distributed uniform with [0, 1] random variables is described by Breidt and Opsomer (2000). Regression function with mean zero and variance one was used to produce population values for all survey variables. In simulation study of 1000 we use the population of size 1000 and different sample sizes (25, 50, 100, 200) drawn by simple random sampling and x generated as identically and independently distributed from uniform distribution with [0, 1] random variables. The population values for y are generated from $y_k(x)$ by adding error term in all cases. The R function neuralnet() has been used to fit neural networks, which is based on back propagation ('rprop+'). There are several software programmes for neural networks, both free and paid like nnet() etc. The values of the inputs and outputs scaled to the range [0, 1] and the activation function chose the logistic and using single hidden layers by increasing no. of neurons / units. We use different no. of parameters (1, 2, 5, and 10) to check the efficiency of the estimators.

5.1 Performance Evaluation Measures

Suppose that proposed estimator is $\hat{t}_{y*}$ and the H.T estimator is $\hat{t}_{y\pi}$ then the MSE of the estimator s time of no. of simulation will be Mean square is define as average squared difference between expected values and actual values and it's represented as

$$MSE = Var(\hat{t}_{y*}) + bias(\hat{t}_{y*})^2$$

Here bias is define as $bias(\hat{t}_{y*}) = E(\hat{t}_{y*}) - t_y$

We also calculate the Scaled Mean Square Error to check the efficiency of the estimators the formula of SMSE is

$$SMSE(\hat{t}_{y*}) = \frac{MSE(\hat{t}_{y*})}{MSE(\hat{t}_{y\pi})VP}$$

where $MSE(\hat{t}_{y*})$ is the MSE of the estimator s time of no. of simulation of proposed estimator and $MSE(\hat{t}_{y\pi})$ is the Horvitz-Thompson estimator, with $SMSE(\hat{t}_{y*})$ the goal is to compare an estimator's MSE to its minimum possible value. In light of this, the estimator's effectiveness increases when SMSE's value is reduced.

Proportion of variance is use for comparative analysis. Under the formula below, we can calculate the proportion of variance due to noise

$$VP = (S_y^2 - S_{y_{nn}}^2)/S_y^2$$

In this case, a survey variable has a population variance S_y^2, whereas a signal has a population variance of $S_{y_{nn}}^2$.

Table 1
Performance of ANN Calibration Estimator for the Four Models with Single Auxiliary Variable based on MSE

MSE														
M	**n**	**p**	**Linear**			**Quad**			**Exp**			**Cycle 1**		
			$\hat{t}_{y\pi}$	$\hat{t}_{y,w}$	$\hat{t}^{mc}_{y,nn}$	$\hat{t}_{y\pi}$	$\hat{t}_{y,w}$	$\hat{t}^{mc}_{y,nn}$	$\hat{t}_{y\pi}$	$\hat{t}_{y,w}$	$\hat{t}^{mc}_{y,nn}$	$\hat{t}_{y\pi}$	$\hat{t}_{y,w}$	$\hat{t}^{mc}_{y,nn}$
2	25	1	14637.2	2245.85	**60.155**	11938.4	1681.27	**1445.82**	6473.47	3479.38	**70.324**	6308.04	127.709	**16.11**
	50	1	5508.55	3007.71	**10.608**	8366.98	5428.53	**28.506**	8075.17	7125.67	**26.98**	380.133	17.223	**0.103**
	100	1	1616.79	2247.41	**12.059**	7495.95	6402.72	**5.052**	5849.32	5541.99	**20.26**	3849.29	5001.25	**0.117**
	200	1	1861.16	3740.82	**0.006**	9758.35	6799.19	**1.194**	7281.89	6473.57	**7.514**	7342.84	4097.85	**5.051**
3	25	1	6492.97	270.379	**34.905**	3879.33	13.505	**5.637**	1602.86	410.233	**8.961**	2662.02	691.326	**70.026**
	50	1	7956.19	513.343	**134.199**	1938.67	289.631	**4.447**	1157.57	221.791	**2.12**	3596.35	203.001	**92.037**
	100	1	358.966	259.468	**45.58**	591.041	46.621	**0.226**	61.978	3.531	**2.417**	4572.2	765.657	**2.07**
	200	1	162.498	35.927	**14.303**	718.12	99.522	**1.562**	153.386	49.597	**0.151**	1044.48	497.701	**0.034**
4	25	1	2197.76	629.002	**500.775**	1244.79	787.369	**5.02**	158.157	50.51	**22.377**	12662.2	1016.28	**21.608**
	50	1	112.882	345.057	**87.163**	278.01	87.697	**2.595**	87.744	1.898	**0.674**	493.891	142.794	**32.941**
	100	1	1796.69	291.145	**27.625**	222.849	60.19	**1.946**	33.072	1.082	**0.982**	517.261	218.454	**0.253**
	200	1	1749.01	25.706	**0.203**	550.55	87.811	**0.054**	124.133	0.64	**0.016**	1256.63	45.481	**0.427**

Table 2
Performance of ANN Calibration Estimator for the Four Models with Single Auxiliary Variable based on SMSE

MSE														
M	**n**	**p**	**Linear**			**Quad**			**Exp**			**Cycle 1**		
			$\hat{t}_{y\pi}$	$\hat{t}_{y,w}$	$\hat{t}_{y,nn}^{mc}$	$\hat{t}_{y\pi}$	$\hat{t}_{y,w}$	$\hat{t}_{y,nn}^{mc}$	$\hat{t}_{y\pi}$	$\hat{t}_{y,w}$	$\hat{t}_{y,nn}^{mc}$	$\hat{t}_{y\pi}$	$\hat{t}_{y,w}$	$\hat{t}_{y,nn}^{mc}$
2	25	1	4.82	0.737	**0.02**	1.294	0.182	**0.015**	1.484	0.798	**0.016**	4.55	0.092	**0.012**
	50	1	4.82	2.632	**0.009**	1.294	0.839	**0.004**	1.483	1.309	**0.004**	4.55	0.206	**0.001**
	100	1	4.82	6.7	**0.035**	1.294	1.015	**0.001**	1.484	1.056	**0.005**	5.912	4.55	**0**
	200	1	9.689	4.82	**0**	1.294	1.085	**0**	1.484	1.066	**0.002**	8.153	4.55	**0.006**
3	25	1	4.929	0.205	**0.026**	371.51	1.293	**0.539**	5.799	1.484	**0.032**	6.633	1.723	**0.174**
	50	1	4.928	0.318	**0.083**	1.293	0.193	**0.003**	1.484	0.284	**0.002**	37.498	6.033	**0.169**
	100	1	6.818	4.928	**0.865**	16.396	1.293	**0.006**	26.051	1.484	**1.016**	39.611	5.89	**0.018**
	200	1	22.292	4.928	**1.962**	9.332	1.199	**0**	4.59	1.805	**0**	18.657	4.169	**0**
4	25	1	4.857	1.39	**1.107**	2.042	1.291	**0.008**	4.66	1.423	**0.659**	81.856	6.561	**0.14**
	50	1	4.857	14.847	**3.75**	11.292	4.096	**0.038**	68.79	2.488	**0.529**	6.569	1.81	**0.438**
	100	1	4.857	0.787	**0.075**	66.291	4.783	**0.042**	45.492	1.898	**1.721**	6.561	4.606	**0.008**
	200	1	4.857	0.071	**0.001**	1.292	0.02	**0**	0.008	1.488	**0.0001**	181.49	6.57	**0.061**

Table 3
Performance of ANN Calibration Estimator for the four Models with P= 2,5,10 based on MSE

MSE														
M	**n**	**p**	**Linear**			**Quad**			**Exp**			**Cycle 1**		
			$\hat{t}_{y\pi}$	$\hat{t}_{y,w}$	$\hat{t}^{mc}_{y,nn}$	$\hat{t}_{y\pi}$	$\hat{t}_{y,w}$	$\hat{t}^{mc}_{y,nn}$	$\hat{t}_{y\pi}$	$\hat{t}_{y,w}$	$\hat{t}^{mc}_{y,nn}$	$\hat{t}_{y\pi}$	$\hat{t}_{y,w}$	$\hat{t}^{mc}_{y,nn}$
2	25	2	6833.02	5178.06	**70.732**	5587.14	4401.28	**7.571**	454.181	387.174	**350.412**	24912	2370.29	**120.952**
		5	246549	221917	**2404.18**	7360.68	1667.04	**493.906**	3497.34	3354.66	**19.133**	9453.69	6851.31	**43.36**
		10	166271	143269	**348.917**	13154.6	8985.64	**2.043**	142.541	76.223	**69.331**	16033.1	8210.78	**76.207**
	50	2	6354.1	5484.11	**28.896**	4345.79	4118.07	**18.039**	197.234	183.329	**0.372**	427.926	286.295	**45.058**
		5	32630.4	31132.1	**465.043**	7064.74	6106.29	**90.695**	237.51	225.478	**0.851**	19640	4197.28	**33.321**
		10	47322.1	44224.3	**2.584**	13154.6	10597.1	**181.278**	3514.58	3017.28	**12.53**	1459.84	854.17	**63.956**
	100	2	9767.71	9469.94	**0.876**	4373.21	4186.62	**3205.05**	500.734	483.104	**0.004**	7615.47	7392.04	**845.505**
		5	23932.7	23299.2	**45.321**	5973.97	5557.85	**15.219**	667.629	629.661	**8.028**	19.431	13.661	**0.127**
		10	75611.2	72434.7	**14.407**	2482.25	1315.16	**0.016**	464.023	379.041	**10.493**	1773.83	1229.03	**0.194**
	200	2	9767.71	9469.94	**0.876**	280.057	267.917	**0.007**	1633.26	1624.58	**0.628**	2337.96	2225.42	**148.16**
		5	5009.57	4894.05	**19.034**	36.95	21.846	**3.255**	602.376	593.667	**0**	3774.04	3553.02	**0.638**
		10	7261.97	7176.78	**48.7**	4087.22	3706.48	**0.138**	1312.6	1279.72	**4.219**	132.021	128.002	**0.057**

Table 4
Performance of ANN Calibration Estimator for the four Models with P= 2,5,10 based on MSE

M	n	p	Linear $\hat{t}_{y\pi}$	Linear $\hat{t}_{y,w}$	Linear $\hat{t}^{mc}_{y,nn}$	Quad $\hat{t}_{y\pi}$	Quad $\hat{t}_{y,w}$	Quad $\hat{t}^{mc}_{y,nn}$	Exp $\hat{t}_{y\pi}$	Exp $\hat{t}_{y,w}$	Exp $\hat{t}^{mc}_{y,nn}$	Cycle 1 $\hat{t}_{y\pi}$	Cycle 1 $\hat{t}_{y,w}$	Cycle 1 $\hat{t}^{mc}_{y,nn}$
2	25	2	9.128	6.917	**0.094**	1.241	0.978	**0.002**	1.06	0.904	**0.818**	1.065	0.942	**0.005**
		5	15.711	14.141	**0.153**	1.178	0.267	**0.079**	1.039	0.997	**0.006**	1.387	1.005	**0.006**
		10	18.194	15.677	**0.038**	1.169	1.092	**0**	1.046	0.559	**0.05**	1.026	0.526	**0.005**
	50	2	9.128	8.417	**0.044**	1.001	0.949	**0.004**	1.055	0.981	**0.002**	1.049	0.702	**0.11**
		5	15.711	14.989	**0.224**	1.178	1.018	**0.015**	1.039	0.987	**0.004**	1.005	0.215	**0.002**
		10	18.194	17.003	**0.001**	1.169	0.941	**0.016**	1.046	0.898	**0.004**	1.026	0.6	**0.045**
	100	2	9.128	8.85	**0.001**	1.24	1.185	**0.907**	1.061	1.018	**0**	1.05	1.019	**0.117**
		5	15.711	15.295	**0.03**	1.178	1.096	**0.003**	1.035	0.976	**0.012**	1.008	0.709	**0.007**
		10	18.194	17.43	**0.003**	1.169	0.619	**0**	1.046	0.854	**0.024**	1.026	0.711	**0**
	200	2	9.128	8.85	**0.001**	1.149	1.099	**0**	1.055	1.049	**0**	1.05	0.999	**0.067**
		5	13.583	13.27	**0.051**	1.004	0.594	**0.088**	1.018	1.003	**0**	1.024	0.964	**0**
		10	18.194	17.98	**0.122**	1.169	1.06	**0**	1.046	1.02	**0.003**	1.059	1.027	**0.001**

5.2 Results

In our study we present the comparison of the performance of the estimators that we used in our research work that is H.T estimator, Calibration estimator, and neural network calibration estimator. For this purpose we compute bias, mean square error (MSE) and scaled mean square (SMSE) of these estimators to compare the efficiency, the smaller the SMSE the greater the efficiency of the estimator or in other words the smallest the value MSE is good than others. We use Monte Carlo simulation then we analyze the results of our estimators.

Table 1 shows the results for MSE of the estimators when sample size is 25 and No. of parameters are ($p = 1$) the results of MSE, and SMSE are shown in table 1 and Table 2. We are interested in to check the efficiency of the estimators. In our given Table 1 and 2 the results for neural network calibration we clearly see that the decreasing MSE and also decreasing the SMSE as compare to H.T estimator and Calibration estimator. We using neural network with signal hidden layer with two units. Neural network calibration performs better than others two. In others words calibration estimator is more efficient that H.T estimator and neural network calibration estimator is more efficient than calibration estimator and H.T estimator. We also check the performance of ANN calibration estimator in all four populations' models by increasing sample size (25, 50, 100, 200), increasing no. units in hidden layer (2, 3, 4) with single auxiliary variable then we can clearly see that the MSE given in Table 1 decreasing in all cases. In this scenario neural network calibration estimator perform much better than others.

We also observed that by increasing sample size, increasing no. units in hidden layer with single auxiliary variable we can see that the neural networks calibration is most efficient than others two on the basis of results of MSE and SMSE. Furthermore, Table 3 and Table 4 shows the results for all four populations by increasing No. of parameters also ($P =$ 2, 5, 10) with single hidden layers with two units and see ANN calibration estimator also perform much better for all populations models as compare to calibration estimator and H.T estimator.

6. CONCLUSIONS

The resulting estimators are more efficient when the functional relationship is complex. Additionally, we are studying and combining neural network model with model calibration estimation for a finite population total. Through neural network learning, we determine the functional relationship between the survey variable and the auxiliary variables. There has been considerable research on the use of neural networks for solving real-life problems, but their application for calibration improved model calibration makes for more flexible predictions as well as straightforward insertion of multivariate analytic information. Firstly we can see that neural network calibration results are all behave well or more efficient on the base of Bias, MSE and SMSE as compare to calibration and H.T estimator, comparable in terms of the number of units selected for the hidden layer. All populations benefit from neural networks' performance. Overall, compared to the parametric H.T. and Calibration estimators, our intended estimator contributed to a good boost in efficiency. The scaled mean square error (SMSE) for neural network calibration is for all populations, usually very near to zero. We conclude on the base of our results that our results for the neural network calibration is always more effective than other methods. We also observe the results by increasing sample size and by increasing no. of parameters and see that neural network perform better in all situations and for all populations. The simulation studies are conducted for the univariate single variable case and also for multivariate variables and show that good gain in efficiency of neural network calibration with respect to the other parametric.

REFERENCES

1. Chambers, R.L. (1997). Weighting and calibration in sample survey estimation. In *Conference on Statistical Science Honouring the Bicentennial of Stefano Franscini's Birth*: Ascona November 18-20, 1996 (pp. 125-147). Birkhäuser Basel.
2. Chambers, R.L., Dorfman, A.H. and Wehrly, T.E. (1993). Bias robust estimation in finite populations using nonparametric calibration. *Journal of the American Statistical Association*, 88, 268-277.
3. Deville, J.C. and Sarndal, C.E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, 87, 376-382.
4. Di Ciaccio, A. and Montanari, G.E. (2001). A nonparametric regression estimator of a finite population mean. *Book of Short Papers*, CLADAG 2001, Palermo, 173-176.
5. Dorfman, A.H. (1992). Nonparametric regression for estimating totals in finite population. *Proceedings of the Section on Survey Research Methods*, 622-625, American Statistical Association, Alexandria, VA.
6. Dorfman, A.H. and Hall, P. (1993). Estimators of the finite population distribution function using nonparametric regression. *The Annals of Statistics*, 21, 1452-1475.
7. Hwang, J.T.G. and Ding, A.A. (1997). Prediction intervals for artificial neural networks. *Journal of the American Statistical Association*, 92, 748-757.
8. Ripley, B.D. (1996). *Pattern Recognition and Neural Networks*, Cambridge University Press, Cambridge.
9. Sarndal, C.E., Swensson, B. and Wretman, J. (1992). *Model Assisted Survey Sampling*, Springer-Verlag, New York.
10. Wu, C. (2002). Optimal calibration estimators in survey sampling, *Working Paper* 2002-01, Department of Statistics and Actuarial Science, University of Waterloo, Canada.
11. Wu, C. and Sitter, R. (2001). A model-calibration to using complete auxiliary information from survey data. *Journal of the American Statistical Association*, 96, 185-193.