Estimating Error Distribution Using Single Index Model

Munaf Yousif Hmood and Rabab Ali Al Suhail
Department of Statistics, College of Administration and Economics, University of Baghdad, 10071 Baghdad, Iraq
munaf.yousif@coadec.uobaghdad.edu.iq, rababsally91@gmail.com

Keywords: Semiparametric Least Squares, Refined Minimum Average Variance Estimation, Refined Outer Product of Gradients, The Empirical Distribution, Kernel Distribution Function.

Abstract: The single index model is regard as a semiparametric model; this kind of models is more flexible and less restrictive than parametric models of conditional mean functions. In this paper we discuss the estimation of error distribution depending on the single index model, our application focuses on bricks production for several factories across Iraq. We compare three estimation methods (Semiparametric Least Squares (SLS), Refined Outer Product of Gradients (rOPG) and Refined Minimum Average Variance Estimation (rMAVE). Then the error distribution is estimated by using two approaches, the empirical distribution and Kernel distribution functions. The results indicate that the rOPG performs best for the single index model, while the empirical distribution function provides the more accurate estimation for error distribution. We use the normal distribution test to certify that the residuals follow a normal distribution for the used single index model estimation methods. Additionally, we compare our results with the multiple linear regression models to give an insight to about the correct specification of the appropriate model.

1 INTRODUCTION

Regression model can be used to study the effect of the explanatory variables X's on the dependent variable Y and then build the suitable model to get optimal forecasts. There are three cases for regression model, the first one is dealt with a certain known population with unknown parameters, so this kind of model is a parametric model and is estimated by parametric methods.

This kind of models may not necessarily accurate due to an appropriate representative for the population under study although it's worked under the strict conditions, therefore we resort to the nonparametric models that do not set strict conditions and assumptions about the structure of the model and do not require any knowledge about the error distribution as in the first case with parametric model.

The third case can be a combination between the other two cases above or a smooth function for the parametric index, so that this kind of method is an extension of linear regression with greater flexibility with using nonlinear smoothing functions, and this kind of models named as single - index- models.

This model can receive most of the information about the relation between X's and Y, and then try to avoid the curse of dimensionality problem accompanying by the nonparametric methods.

A considerable literature has addressed the estimation of the single Index model (SIM) and error distribution, Ichimura (1993) [1] proposed two estimators for estimating SIM, the semiparametric least squares estimator (SLS) and the weighted semiparametric least squares estimator (WSLS). In 2001 Hristache, juditsky and Spokoiny (HJS) [2] proposed the HJS method for estimating SIM, that improves the ADE procedure through iterative refinement, typically requiring more than two iterations. Later, Xia (2006) [3] developed the refined outer product of gradients (rOPG) estimator as an alternative to the (HJS) method, addressing its limitations when the link function is symmetric. Xia also applied the refined Minimum Average Variance Estimation (rMAVE) method, and derived approximate distributions for both estimators Müller et al. in 2012 [4] investigated the use of the empirical distribution function based on the residuals within the partial linear regression (PLR) framework. In (2018), Müller et al. [5] proposed an efficient estimator for the error distribution using a Kernel distribution function. More recently, Rabab and Hmood (2022) [6] estimated the error distribution estimate for the SIM using both the empirical distribution function and the Kernel distribution function, while the SIM itself was estimated using the refined Minimum Average Variance Estimation (rMAVE) method, their simulation experiments showed that the Kernel distribution function provided best efficient performance.

2 SINGLE INDEX MODEL

Estimating the error distribution becomes essential necessary after estimating the regression function. This is because the single index model does not impose a predetermined form for the error distribution. Instead, the error distribution is estimated after the regression function has been estimated allowing assess how closely the estimated regression curve approximates the true underlying relationship in the data, this helps to determine whether the errors are approximately normally distributed.

The single index model is a semiparametric model, and does not assume that the link function is known, so it is more flexible and less restrictive than parametric models of conditional mean functions, such as linear models and Probit probability models, flexibility is important in applications because there is usually no justification for assuming that the link function is known in advance, and seriously misleading results can be obtained if incorrect specifications of the link function are assumed [7].

The single-index model is based on the fact that the appropriate model for the data is a parametric model containing unknown parameters that can be estimated initially by one of the parametric regression methods such as Maximum Likelihood or least squares methods, and then it can be considered that the estimate resulting from the substitute of the estimated parameters in the parametric model follows a parametric model, i.e., a model that does not have specific parameters, but contains an unknown function without specific parameters, and there is a nonlinear relationship between the explanatory variables and the response variable, and the data resulting from the estimation of the first stage is estimated using one of the nonparametric estimation methods [6], [8].

The single index model can be written as [7]:

$$Y = G(X^T\beta) + e. \tag{1}$$

Where:

- $Y \in R$: denote the vector of a dependent variable with degree n×1.
- $X^T \in R$: denote the row vector of independent variables of degree 1×p.
- β: The unknown vector of parameters of a degree p×1 and must satisfy the following conditions:
  $$||\beta|| = 1 \text{ or } \beta^T\beta = 1 \text{ and } \beta_1 > 0.$$
- G: denote the unknown link function of degree n×1.
- e: denote the vector of degree n×1 of the random error with zero mean and variance σ_e^2.

3 SEMIPARAMETRIC LEAST SQUARES (SLS)

This method was introduced by Ichimura (1993) and is similar to nonlinear parametric regression, in this way the regression function $E[Y|X^T\beta]$ can first be estimated locally using any smoothing method with a certain index β that satisfy $\beta^T\beta = 1$ and then the vector of parameters β can be estimated by minimizing the difference between the observed response and the estimated regression function with respect to β. [1], [7].

To estimate the parameter vector β, the following idea is followed: create an objective function to estimate the parameter vector β with a convergence rate $n^{-0.5}$within the local function of the link function G. Because it is unknown function, it must be replaced by nonparametric estimates, and then the objective function is maximized or minimized with respect to the parameter vector β.

The objective function can be stimulated by minimizing the variance in the data that cannot be explained by the traditional regression, this residual variance can be written as [9]:

$$\mathrm{Var}\{Y|X^T\beta\} = E[\{Y - G(X^T\beta)\}^2|X^T\beta].$$

With $X^T\beta$ being the linear index, where the previous equation leads us to the well-known nonlinear least squares estimator (NLS):

$$\min_{\beta} \frac{1}{n}\sum_{i=1}^{n}\left(Y - G(X_i^T\beta)\right)^2. \tag{2}$$

The above equation (2) can be used to estimate the vector of parameters β if the link function G is known. Otherwise, if the link function G is unknown we first need to estimate it, however, the Kernel method does not estimate $G(X_i^T\beta)$ directly, because G is the conditional mean of $(Y|X^T\beta)$ that depends on β, so the two step estimator is likely to be inefficient, because of that Ichimura in 1993 proposed replacing G with a leave -one-out NW estimator:

$$\hat{G}_{-i}(X_i^T\beta) = \frac{\sum_{i\neq j}^{n} K\left(\frac{(X_i - X_j)^T\beta}{h}\right) Y_j \ I(X_j \in A_n)}{\sum_{i\neq j}^{n} K\left(\frac{(X_i - X_j)^T\beta}{h}\right) \ I(X_j \in A_n)}. \quad (3)$$

Where:

- K: is a Kernel function, we assumed Gaussian Kernel.
- h: is the (Bandwidth) and was estimated by cross validation method.
- $I(X_j \in A_n)$: Weight function or trimming term.

Hence the semiparametric least squares estimator is:

$$\hat{\beta} = \min_{\beta} \frac{1}{n}\sum_{i=1}^{n}\left(Y - \hat{G}_{-i}(X_i^T\beta)\right)^2 1(X_i \in Av). \quad (4)$$

4 REFINED OUTER PRODUCT OF GRADIENTS (ROPG)

The method of the average derivative estimation (ADE) is the first approach that uses the structure of the model directly because the parameter vector β is proportional to the derivatives $\partial E(Y|X = x)/\partial x = G'(x^T\beta)\beta$, where the parameter vector β can be estimated by using the Average Derivative method (ADE) which computes the average of the derivatives or gradients of the regression function. the ADE method uses a high-dimensional kernel to estimate derivatives, this method is easy to implement with an easy-to-understand algorithm but suffers from the curse of dimensionality [10], [11].

Hristache et al. in 2001 [2] adopted the same idea of the ADE method and proposed a dynamic procedure to adapt the model structure by reducing the kernel dimensions and instead of considering the average derivatives as in HJS, Xia used the outer product of derivatives, the refining scheme of Kernel weights is used in the estimation, and the resulting estimator (refined external product estimator) is easy to implement with the same simplicity of estimation in HJS.[3]

Suppose that: $g(x) = E(Y|X = x)$, then

$$\nabla g(x) = \frac{\partial}{\partial x} g(x) = \frac{\partial}{\partial x} G(x^T\beta) = G'(x^T\beta)\beta. \quad (5)$$

Therefore, the derivative or gradient of the regression function at any point has the same direction of β and based on these observations, Powell et al. 1989 [11] and Hardle and Stoker [10] in 1989 proposed estimating β by $E\langle\nabla g(X)\rangle$ and they called this method the Average derivative estimation (ADE).

Assuming that f (x) is the probability density function of X, and using the integration by parts:

$$E\langle\nabla g(X)\rangle = -E\langle Y\nabla \log f(X)\rangle.$$

They proposed estimating $\nabla \log f(x)$ instead of estimating $\nabla g(x))$ by smoothing a higher-order Kernel due to the use of a high-dimensional density function, where the ADE method still suffers from The curse of dimensionality is another drawback in this method, which is $E(G'(X^T\beta)) = 0$, where $\nabla g(X) = E(G'(X^T\beta))\beta = 0$, so that this method failed to estimate the vector of parameters β to overcome this drawback, Samarov1993 proposed the outer product of gradients:

$\langle\nabla g(X)\nabla^T g(X)\rangle$

Noting that:

$$E\langle\nabla g(X)\nabla^T g(X)\rangle = E\left\langle\left[G'(X^T\beta)\right]^2\right\rangle \beta\beta^T.$$

The vector of parameters β has one nonzero eigenvalue, so the index β is the eigenvector corresponding to the largest eigenvalue of $E\langle\nabla g(X)\nabla^T g(X)\rangle$ and to implement the estimation Xia 2006 estimated the gradients by local linear smoothing, from the following minimization problem [3]:

$$\min_{a_j b_j} \sum_{i=1}^{n}\left(Y_i - a_j - b_j X_{ij}^T\right)^2 w_{ij}. \quad (6)$$

Where:

$$X_{ij}^T = X_i^T - X_j^T.$$

w_{ij}: Indicates the weight and depends on the distance between X_i^T, X_j^T.

Then calculate: $\hat{\sum} = \frac{1}{n}\sum_{j=1}^{n} \hat{b}_j \hat{b}_j^T$.

Where $\hat{b}_j$ is the minimization of equation (6), since the first eigenvector β of $\sum$ is the estimator of β. To improve efficiency and to allow the estimation to adapt to the structure of the dependence of Y on X, Xia, Tong, Li and Zhu 2002 taking the kernel weight $w_{ij} = K_h(X_{ij}^T\hat{\beta})$, this method is called the refined Outer Product of Gradients (rOPG). This method can be summarized with the following algorithm [6], [12]:

An initial estimate of the parameter vector B is obtained by using the ordinary least squares method or it may initially imposed.

1) Calculate the vector $\binom{a_j}{b_j}$ as follows:

$$\binom{a_j}{b_j} = \left[\sum_{i=1}^{n} K_h(X_{ij}^T\beta)\binom{1}{X_{ij}}\binom{1}{X_{ij}}^T\right]^{-1} \sum_{i=1}^{n} K_h(X_{ij}^T\beta)\binom{1}{X_{ij}} Y_i. \quad (7)$$

2) Calculate the first eigenvector corresponding to the largest eigenvalue of:

$$\hat{\Sigma} = n^{-1} \sum_{j=1}^{n} \hat{\rho}_j \, b_j \, b_j^T. \quad (8)$$

Where:

$$\hat{f}(X_j^T\beta) = n^{-1} \sum_{i=1}^{n} K_h(X_{ij}^T\beta), \hat{\rho}_j = \rho_n\left(\hat{f}(X_j^T\beta)\right).$$

- $\rho_n(.)$: denoted to the trimming function that can be used to adjust the boundary points:

$$\rho_n(\varsigma) = \begin{cases} 1 & if\ \varsigma \geq 2c_O n^{-\epsilon} \\ \frac{exp\{(2c_O n^{-\epsilon}-\varsigma)^{-1}\}}{exp\{(2c_O n^{-\epsilon}-\varsigma)^{-1}\}+exp\{(\varsigma-c_O n^{-\epsilon})^{-1}\}} & if\ 2c_O n^{-\epsilon} > \varsigma > c_O n^{-\epsilon} \\ 0 & if\ \varsigma \leq c_O n^{-\epsilon} \end{cases}. \quad (9)$$

- c_O,ϵ: constants denote to the trimming parameters and their values are $(\epsilon, c_O < 1/_{20})$.

3) Repeat (b) and (c) with updating β until convergence.

The final eigenvector in the algorithm is an estimator (rOPG) of β and is referred to as $\hat{\beta}_{rOPG}$.

When compared with the HJS method, the rOPG method directly utilizes the single-index weight $w_{ij} = K_h(X_{ij}^T\beta)$.

In this method, we replace the average derivative used in the HJS method with the outer product to avoid the issue of $E(G^{'}(X^T\beta)) = 0$.

5 REFINED MINIMUM AVERAGE VARIANCE ESTIMATION METHOD (RMAVE)

The Minimum Average Variance Estimation (MAVE) method was proposed by Xia, Tong, Li and Zhu in 2002. This method is highly effective for estimating the single index model and is also suitable for combining with other techniques to incorporate additional statistical requirements [12].

The main idea of (MAVE) method is to locally approximate the smoothed link function and then estimate the parameter vector β by minimizing the total approximation errors [1], [3].

$$\hat{\beta} = \min_{\beta:=\beta/|\beta|} \sum_{j=1}^{n} \sum_{i=1}^{n} \left(Y_i - a_j - d_j X_{ij}^T\beta\right)^2 w_{ij}. \quad (10)$$

Xia, Tong, Li and Zhu in 2002 improved MAVE method using refined Nadaraya-Watson Weight function:

$$\ddot{w}_{ij} = \frac{K_h\left(X_{ij}^T\hat{\beta}\right)}{\sum_{i=1}^{n} K_h\left(X_{ij}^T\hat{\beta}\right)}. \quad (11)$$

This method is called (Refined MAVE) and works by using (10), with replacing $\ddot{w}_{ij}$ instead of w_{ij} and then replacing the resulting estimator $\hat{\beta}$ from equation (10) in (11) and newly re-estimate β by substituting (11), and repeat this procedure until obtaining convergence [12].

The estimation of error distribution relies on empirical functions that use the residuals of the single-index model. These residuals are obtained after the regression function $\hat{G}(.)$ has been estimated:

$$\hat{e}_i = Y_i - \hat{G}\left(X_i^T\hat{\beta}\right). \quad (12)$$

We present the two methods as follows:

5.1 Estimating the Empirical Distribution Function of the Error

The empirical distribution function is the simplest way to estimate the cumulative distribution function, given a finite random sample $\hat{e}_1, \dots, \hat{e}_n, n \in N$. If e is the random error variable of the population, its distribution function is defined as:

$$F(t) = P[e \leq t], \quad t \in R \quad (13)$$

The statistical estimate of F(t) based on the random samples $\hat{e}_1, \dots, \hat{e}_n, n \in N$, i.e. based on the residual is the so-called empirical distribution function defined as follows: [13], [14].

$$\hat{F}_n(t) = \frac{1}{n}\{\#\ of\ \hat{e}_i\text{'s} \leq t\} = \frac{1}{n}\sum_{i=1}^{n} I(\hat{e}_i \leq t)\ , t \in R\ , (14)$$

referring to $I(.)$ to the index function for all t.

The estimator of the empirical distribution function is a step function and jumps each step by a height of (1 / n) at each point of the observed sample $\hat{e}_i$.

5.2 Estimating the Error Distribution Function Using the Kernel Method

The Kernel estimator for the error distribution is essentially a smoothed version of the empirical distribution function $\hat{F}_n(t)$ constructed from the residuals. This smoothing is necessary because the empirical distribution function is not smoothed due to it jumps by (1 / n) at each sample point, so a smoothed estimate of the cumulative distribution function F(t) can be obtained by integrating the Kernel density function $\hat{f}(t)$ based on the residuals, so that we get the smoothed estimator $\hat{F}_*(t)$:[5]

$$\hat{F}_*(t) = \int_{-\infty}^{t} \hat{f}(v)dv = \frac{1}{nh}\sum_{i=1}^{n}\int_{-\infty}^{t} K\left(\frac{v-\hat{e}_i}{h}\right)dv$$

Let W be a cumulative Kernel function defined as:

$$\mathbb{W}(t) = \int_{-\infty}^{t} K(v)dv$$

So we get the estimator in its final form:

$$\hat{F}_*(t) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{W}\left(\frac{t-\hat{e}_i}{h}\right) \quad (15)$$

h is the bandwidth and was estimated by the cross-validation method.

6 RESULTS

In this section, the methods mentioned in the previous sections were applied on one of the important industrial sectors, which is the production of bricks, which can be considered as one of the vital and important products of the national economy and this is due to its direct contribution to the reconstruction, construction, and there is an urgent need to produce bricks due to the increase in population and the need for houses and residential complexes. [15]

Our sample consisted of real data for 337 brick production plants in the private sector for 15 governorates in Iraq for the year 2021 and the data was obtained from the Ministry of planning. The dependent variable is the sum of the production of bricks, and the explanatory variables are represented by the factors affecting production, as X_1 is represented by wages, X_2 is represented by the sum of workers, and X_3 is represented by the sum of the value of supplies. The explanatory variables are within a smoothing function $G(.)$.

For the purpose of estimating the single index model (1), we will used three methods:

- Semiparametric least squares SLS.
- Refined Outer Product of Gradients (rOPG).
- Refined Minimum Average Variance Estimation method (rMAVE).

After extracting the estimated equation, the residuals for each model are found using (15), and then the normality of the residuals is examined. The error distribution function for these residuals is then estimated using two methods:

- The empirical distribution function $\hat{F}_n(t)$.
- Kernel distribution function $\hat{F}_*(t)$.

The following Tables 1 and 2 shows the values of the mean squares error of the error estimation methods:

Table 1: Mean Squared errors (MSE) of the empirical distribution function $\hat{F}_n(t)$.

t	$\hat{F}_n(t)$			
	SLS	rOPG	rMAVE	Best
-2	0.0005198	0.0003510	0.00035103	rMAVE
-1.5	0.004462	0.0064231	0.00642318	SLS
-1	0.0025185	0.0013739	0.00137399	rOPG
0	0.0253457	0.0210233	0.02102332	rOPG
1	0.0707785	0.0639697	0.06396978	rOPG
1.5	0.0870862	0.0795148	0.07951485	rOPG
2	0.0954919	0.0875554	0.08755550	rOPG

Table 2: Mean Squared Errors (MSE) of kernel distribution function $\hat{F}_*(t)$.

t	$\hat{F}_*(t)$			
	SLS	rOPG	rMAVE	Best
-2	0.0005198	0.00050772	0.00050771	rMAVE
-1.5	0.0044622	0.00443458	0.00443457	rMAVE
-1	0.0025185	0.00251388	0.00251387	rMAVE
0	0.0253457	0.02499716	0.02499715	rMAVE
1	0.0707785	0.07076324	0.07076321	rMAVE
1.5	0.0870862	0.08705806	0.08705802	rMAVE
2	0.0954919	0.09545076	0.09545075	rMAVE

To test the normal distribution of the semiparametric model the residuals is tested after estimating the single index model with the three estimation methods. One-Sample Kolmogorov-Smirnov test statistic was used and the test results were presented in Table 3:

Table 3: Normal distribution test for the residuals of SIM estimation methods.

	SLS residuals	rMAVE residuals	rOPG residuals
n	337	337	337
Mean	13.99	14.01	13.98
Std.Deviation	0.443	0.450	0.442
Kolmogorov-Smirnov	1.192	1.157	1.228
Distribution type	Normal	Normal	Normal
Asymp sig. (2-T)	0.117	0.137	0.098

Figures 1, 2, and 3 illustrate the behavior of the residuals resulting from the application of three different methods for estimating a single-index model.

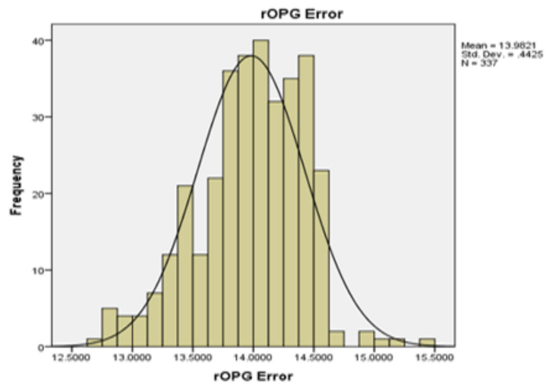

Figure 1: The behavior of the residuals using the rOPG method.

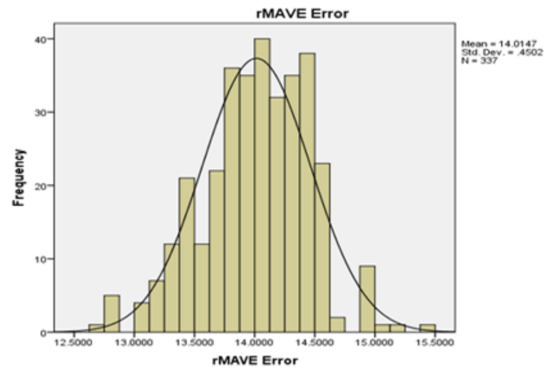

Figure 2: The behavior of the residuals using the rMAVE method.

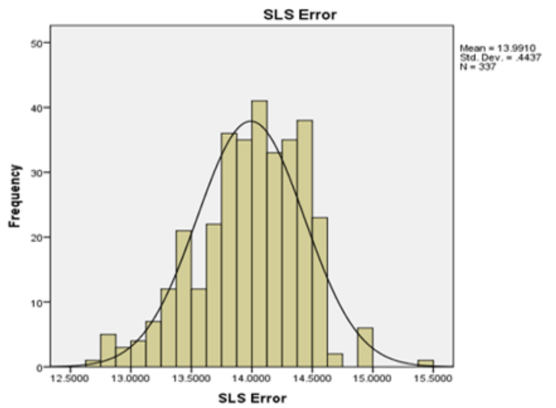

Figure 3: The behavior of the residuals using the SLS method.

To evaluate the performance of the Single Index model (SIM) relative to the Multiple Linear Regression (MLR) Model, the normality of (MLR) was assessed distribution and subjected to One-Sample Kolmogorov-Smirnov test for normality, the results are summarized in Table 4 below:

Table 4: Normal distribution test for the residuals of multiple regression estimation method.

	Error
n	337
Mean	222937.30
Std.Deviation	131983.77
Kolmogorov-Smirnov	1.160
Distribution type	Not Normal
Asymp sig(2-T)	0.0135

The following Figure 4 indicates the behavior of the residuals that resulting from estimating multiple regression model by OLS method.

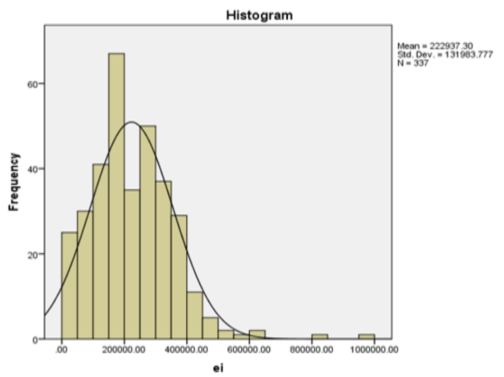

Figure 4: The behavior of the residuals using the OLS method.

7 DISCUSSION

From the results presented in Tables 1 and 2, it can be observed that the empirical distribution function $\hat{F}\_n(t)$ generally provides better performance across most values of t, except at $t = -1.5$, where the kernel-based distribution function estimator $\hat{F}\_*(t)$ shows superior performance.

Regarding the estimation of the single-index regression function, the best-performing method in terms of error distribution is the rOPG approach. In this case, the empirical distribution function $\hat{F}\_n(t)$ yields the most accurate estimation results when combined with the rOPG method. In contrast, the kernel-based estimator $\hat{F}\_*(t)$ performs better when the model is estimated using the rMAVE method.

The results presented in Table 3 and Figures 1–3 indicate that the residuals follow an approximately normal distribution. However, as shown in Table 4 and Figure 4, a noticeable left-skewness is observed, suggesting a deviation from normality. This indicates that the residuals do not fully satisfy the assumptions of a normal distribution, and therefore a parametric modeling approach may not be appropriate for this dataset. These findings are further supported by formal normality test results.

8 COCLUSIONS

Based on the analysis results presented in the tables, comparing the methods (rOPG, rMAVE, and SLS) leads to the following conclusions:

1) The model estimated using the reparametrized outer product gradient (rOPG) method exhibits the most favorable error distribution function, making it the best model for representing the brick production function.
2) The method employing the empirical distribution function $\hat{F}_n(t)$is the most effective for estimating the error distribution function. This indicates its capability in providing the most accurate representation of the brick production function in Iraq.
3) The residuals from all single-index model estimation methods follow a normal distribution, demonstrating the accuracy of all methods.
4) The residuals obtained from the rOPG method exhibit the closest adherence to a normal distribution, which is also evident from the corresponding plots.
5) Comparing the normality test results of the residuals from the fully parametric and semiparametric models underscores the significance of correct model specification on the study's outcomes.

REFERENCES

[1] H. Ichimura, “Semiparametric least squares (SLS) and weighted SLS estimation of single-index models,” J. Econom., vol. 58, pp. 71-120, 1993.

[2] M. Hristache, A. Juditsky, and V. Spokoiny, “Direct estimation of the index coefficient in a single-index model,” Ann. Stat., vol. 29, no. 3, pp. 595-623, 2001, doi: 10.1214/aos/1009210681.

[3] Y. Xia, “Asymptotic distributions for two estimators of the single-index model,” Econometric Theory, vol. 22, pp. 1112-1137, 2006.

[4] U. U. Müller, A. Schick, and W. Wefelmeyer, “Estimating the error distribution function in semiparametric additive regression models,” J. Stat. Plan. Inference, vol. 142, pp. 552-566, 2012.

[5] U. U. Müller, A. Schick, and W. Wefelmeyer, “Estimating the error distribution function in nonparametric regression with multivariate covariates,” Stat. Probab. Lett., vol. 79, pp. 957-964, 2009.

[6] R. A. Abd-alrahman and M. Y. Hmood, “On error distribution with single index model,” Int. J. Agric. Stat. Sci., vol. 18, pp. 2445-2450, 2022.

[7] J. L. Horowitz, Semiparametric Methods in Econometrics, New York, NY, USA: Springer, 2012.

[8] J. S. Simonoff and C. L. Tsai, “Score tests for the single index model,” Technometrics, vol. 44, pp. 142-151, 2002.

[9] Y. Xia, W. K. Härdle, and O. B. Linton, “Optimal smoothing for a computationally and statistically efficient single index estimator,” in Exploring Research Frontiers in Contemporary Statistics and Econometrics, I. Van Keilegom and P. Wilson, Eds., Heidelberg, Germany: Physica, 2011, pp. 229-261.

[10] W. Härdle and T. M. Stoker, “Investigating smooth multiple regression by the method of average derivatives,” J. Am. Stat. Assoc., vol. 84, pp. 986-995, 1989.

[11] J. L. Powell, J. H. Stock, and T. M. Stoker, “Semiparametric estimation of index coefficients,” Econometrica, vol. 57, pp. 1403-1430, 1989.

[12] Y. Xia, H. Tong, W. K. Li, and L. Zhu, “An adaptive estimation of dimension reduction space (with discussion),” J. R. Stat. Soc. Ser. B, vol. 64, pp. 363-410, 2002.

[13] A. Hanebeck, Nonparametric Distribution Function Estimation, Master’s thesis, Karlsruhe Institute of Technology, Germany, 2020.

[14] H. L. Koul, U. U. Müller, and A. Schick, “Estimating the error distribution in a single-index model,” in From Statistics to Mathematical Finance, D. Ferger, W. González Manteiga, T. Schmidt, and J. L. Wang, Eds., Cham, Switzerland: Springer, 2017, pp. 209-233.

[15] A. J. Al-Janabi, “The brick industry in Iraq for the period (2000-2012) (analytical study),” Al-Adab J., vol. 110, pp. 433-456, 2014.