Chilean Journal of Statistics
Vol. 4, No. 2, September 2013, 61–74

Estimation Theory
Research Paper

A New Version of Local Linear Estimators

Munaf Y. Hmood[1,*] and Ulrich Stadtmüller [2]
[1]Department of Statistics,
Baghdad University, Baghdad, Iraq
[2]Department of Number and Probability Theory,
Ulm University, Ulm, Germany
(Received: 12 May 2011 · Accepted in final form: 10 July 2013)

Abstract

In the current article we proposed a new version of local linear estimators.The main idea here is to combine two estimators to produce a new estimator having the best features from the original estimators. The proposed estimator is a convex combination between fixed and variable bandwidth local linear estimators and the aim is to choose a suitable parameter steering this combination. We use least squares minimization to select this parameter. Simulations are conducted to study a finite sample behavior of the proposed estimator which is compared with both fixed and variable bandwidth local linear estimators. Also real data on Buying and Selling Auction of the Central Bank of Iraq for US$ and on the GDP of Iraq are applied to get an impression of the behaviour of the estimators. The final results indicate that the proposed estimator performs well being flexible and stable and can be considered as a good alternative to the individual estimators.

Keywords: Local linear estimator. Fixed bandwidth. Variable bandwidth. Combining two estimators. Rule of the thumb.

Mathematics Subject Classification: Primary 62G05 · Secondary 62G08

1. Introduction

A regression curve describes a general relationship between an explanatory variable X and a response variable Y. We resort to Kernel method to estimate the regression function when no assumption is made about the form of regression function. The kernel method has been studied extensively in the literature since the Nadaraya (1964) and Watson (1964) estimator. There are different kinds of this estimator (PC, GM, Local linear and polynomial estimators) see Mller (1987), Fan (1992) (1993), Fan and Gijbels (1996). The kernel method is studied in both fixed and random designs contexts, so if we have n independent observations $(X_i, Y_i)_{i=1}^n$, then the regression function in the case of random design is equal to:

*Corresponding author: Munaf Y. Hmood. Email:munaf_yousif@yahoo.com

ISSN: 0718-7912 (print)/ISSN: 0718-7920 (online)

http://www.soche.cl/chjs

$$m(x) = E(Y \mid X = x)$$

while $\sigma^2(X_i) = Var(Y_i \mid X_i = xi)$ is the variance function.
In the other case with fixed design $(x_i)_{i=1}^n$ are ordered non-random numbers, therefore the response $(Y_i)_{i=1}^n$ follow:

$$Y_i = m(x_i) + \sigma(x_i)\varepsilon_i, \qquad 1 \leq i \leq n \tag{1}$$

where the (ε_i) are iid random variables with

$$E(\varepsilon_i) = 0, Var(\varepsilon_i) = 1 \text{ hence } E(Y_i) = m(x_i), Var(Y_i \mid X_i = x_i) = \sigma^2(x_i).$$

Often it is assumed that $\sigma^2(x_i) \equiv \sigma^2$ for all i , in this case the model is said to be homoscedastic, otherwise the model is heteroscedastic.
The aim of current paper is to compare between so called local linear estimator for both fixed and variable bandwidth with the proposed estimator that combine these estimators by using a combinig (tuning) parameter that depends on the observations. There are many reasons for choosing local linear estimator, it has nice properties and for more details see Fan (1992, 1993), Fan and Gijbels (1992, 1996). The paper is organized as follows: In section 2 the proposed method is described. In section 3 local linear estimator with variable bandwidth is introduced. In section 4 the combining parameter selection procedure is explained. Bandwidth selection method algorithm is presented in section 5.The simulation experiment results are presented in section 6 and finally Real data applications on (Buying and Selling Auction of the Central Bank of Iraq for US$ and on the GDP of Iraq) and the recommendations are presented in sections 7 and 8 respectively.

2. THE PROPOSED ESTIMATOR

"Whether any method is always best and if not whether there are distinct situations where one is preferable and if there exist such estimator, Are there any estimators better than it?
"Jennen-Steinmetz and Gasser (1988)"

This question leads us to use local linear estimator which has good properties and also compares favorably with other estimators to build a new version of it. The main idea here is to use a suitable combination between fixed and variable bandwidths local linear estimator. The reason for studying this combination is due to fact that there exist experiments that require a fixed bandwidth and other a variable bandwidth choice. The fixed bandwidth is used when the unknown regression function shows approximately the same smoothing overall the estimation range and the data are homoscedastic; see Fan and Gijbels (2000).This means the bandwidth is used for the entire range of X in the regression function. On the other hand the variable bandwidth is used when the regression function shows different behavior in different regions or the data are heteroscedastic and this happens when the regression function has a complicated structure (low smoothness) (very rough), we therefore propose this new estimator with the aim of reducing the uncertainty. The key idea in this paper is to fit the regression function by:

$$\hat{m}(x, \gamma) = \gamma \hat{m}_V(x) + (1 - \gamma)\hat{m}_F(x) \tag{2}$$

Where $0 \leq \gamma \leq 1$ is a tuning parameter and is unknown and can be estimated from the

data. Furthermore,
$\hat{m}_V(x)$ denotes the variable bandwidth local linear estimator and
$\hat{m}_F(x)$ is the fixed bandwidth local linear estimator.
Other combinations of estimators such as spline, wavelets, etc are also possible. The rationale for the proposed estimator is that when the data conform closely to the variable bandwidth local linear estimator then γ should give greater weight to $\hat{m}_V(x)$, so that the estimated $\hat{\gamma}$ is close to one.
On the other hand, when the data conform to fixed bandwidth local linear estimator then more weight will be given to $\hat{m}_F(x)$ and the estimate $\hat{\gamma}$ is close to zero.

3. Variable Bandwidth Local Linear Estimator

In this paper we will just exhibit the variable bandwidth local linear estimator and its properties because this estimator is constructed similarly to the local linear with fixed bandwidth but the difference to the fixed bandwidth case that is the smoothing parameter is allowed to vary from one data point to another and it will proportional to some function of the observation .The idea is to use a larger kernel in regions of low design density points to avoid the roughness. To put the basic idea of this estimator assumes that the observed data $(X_i, Y_i)_{i=1}^n$ are a random sample from a certain population (Y, X). We minimize

$$\sum_{1}^{n}(Y_i - \beta_0 - \beta_1(X_i - x))^2 K_{h(X_i)}(X_i - x) \text{ With respect to } \beta_0 \text{ and } \beta_1 \tag{3}$$

Here $K_{h(X)} = \frac{1}{h}K(\frac{x}{h})$ is a kernel with bandwidth h and $h(x) = h.\lambda(x)$ with some small but positive h; see Silverman (1986).

$$\lambda(x) = (\frac{f(x)}{g})^{\frac{-1}{2}}$$

Where g denotes the geometric mean of probability density function $f(X_i)$ where $f(X_i)$ can be estimated by using kernel density estimator with reference bandwidth and kernel of degree 2. For other choices of $h(x)$; see Silverman (1986), Jennen-Steinmetz and Gasser (1988).
The solution of the aforementioned equation can be put in matrix notation. Assuming invertibility of $X^T W X$, this leads to the solution:

$$\underline{\hat{\beta}} = (X^T W X)^{-1} X^T W \underline{Y} \tag{4}$$

Where

$$X^T W X = \begin{pmatrix} S_0(x, h(.)) & S_1(x, h(.)) \\ S_1(x, h(.)) & S_2(x, h(.)) \end{pmatrix}$$

$$X^T W Y = \begin{pmatrix} T_0(x, h(.)) \\ T_1(x, h(.)) \end{pmatrix}$$

And $\underline{Y} = (Y_1, ..., Y_n)^T$ is the $(n \times 1)$ vector of responses. Furthermore

$$X = \begin{pmatrix} 1 & (X_1 - x) \\ . & . \\ . & . \\ . & . \\ 1 & (X_n - x) \end{pmatrix}$$

Is an $(n \times 2)$ design matrix.

$W = diag\{K_{h(X_1)}(X_1 - x), K_{h(X_2)}(X_2 - x), ..., K_{h(X_n)}(X_n - x)\}$ be an $(n \times n)$ diagonal matrix of kernels with variable bandwidth and

$$S_\lambda(x, h(.)) = \sum_{i=1}^{n} K_{h(X_i)}(X_i - x)(X_i - x)^\lambda, \lambda = 0, 1, 2 \tag{5}$$

$$T_k(x, h(.)) = \sum_{i=1}^{n} K_{h(X_i)}(X_i - x)(X_i - x)^k Y_i, k = 0, 1$$

Since the estimator of regression function is the intercept coefficient, i.e. $\hat{m}_V(x) \equiv \hat{\beta}_0$, we find

$$\hat{m}_V(x) = \underline{e_1}^T \underline{\hat{\beta}} = \frac{S_2(x, h(.))T_0(x, h(.)) - S_1(x, h(.))T_1(x, h(.))}{S_2(x, h(.))S_0(x, h(.)) - S_1^2(x, h(.))} \tag{6}$$

Where $\underline{e_1} = (1, 0)^T$.
To avoid zero in the denominator, we can modify (6) by adding $(nh)^{-1}$ in the denominator.

3.1 ASYMPTOTIC PROPERTIES

For the asymptotic properties of local linear estimator the following conditions must be satisfied; see Fan (1992).

(1) The regression function $m(x)$ has a bounded and continuous second derivative.
(2) The variance function $\sigma^2(x)$ is bounded and continuous.
(3) The kernel is symmetric about zero and has a bounded support and first derivative.
(4) The bandwidth h is a sequence satisfying $h \to 0$ and $nh \to \infty$ as $n \to \infty$.
(5) The marginal density f of the covariate X is continuous and bounded away from zero in an interval (a, b).

In addition to the above conditions assume that f' is continuous and suppose that $\lambda(.)$ is bounded function with $inf\{\lambda(z)\} > 0$. It follows from (6) that:

$$E\{\hat{m}_V(x) - m(x) \mid X_1, X_2, ..., X_n\} = \frac{1}{2} m''(x) \underline{e_1}^T (X^T W X)^{-1} X^T W \begin{bmatrix} (X_1 - x)^2 \\ . \\ . \\ . \\ (X_n - x)^2 \end{bmatrix} + ... \tag{7}$$

Where $\cdots$ denote terms of smaller order. In order to compute a simpler approximation of this term observes that:

$$S_\lambda(x,h(.)) = \begin{cases} n(h(x))^\lambda d_\lambda f(x) + o_p(nh^\lambda) & \lambda \text{ even} \\ n(h(x))^{\lambda+1} d_{\lambda+1} f'(x) + o_p(nh^{\lambda+1}) & \lambda \text{ odd} \end{cases}$$

Where $d_\lambda = \int u^\lambda K(u)du$. This will lead to:

$$X^T W X = \begin{pmatrix} nf(x) + o_p(n) & n(h(x))^2 f'(x) d_2 + o_p(nh^2) \\ n(h(x))^2 f'(x) d_2 + o_p(nh^2) & n(h(x))^2 f(x) d_2 + o_p(nh^2) \end{pmatrix}$$

$$X^T W \begin{bmatrix} (X_1 - x)^2 \\ . \\ . \\ . \\ (X_n - x)^2 \end{bmatrix} = \begin{pmatrix} S_2(x,h(.)) \\ S_3(x,h(.)) \end{pmatrix} = \begin{pmatrix} n(h(x))^2 f(x) d_2 + o_p(nh^2) \\ n(h(x))^4 f'(x) d_4 + o_p(nh^4) \end{pmatrix} \tag{8}$$

And then to

$$(X^T W X)^{-1} = \frac{\begin{pmatrix} n(h(x))^2 f(x) d_2 + o_p(n^{-1}) & -n(h(x))^2 f'(x) d_2 + o_p(n^{-1}) \\ -n(h(x))^2 f'(x) d_2 + o_p(n^{-1}) & nf(x) + o_p(n^{-1}h^{-2}) \end{pmatrix}}{(nh(x))^2 d_2 (f^2(x) - (h(x))^2 (f')^2(x) d_2)} \tag{9}$$

Combining these equations we get the conditional bias of $\hat{m}_V(x)$.

$$E\{\hat{m}_V(x) - m(x) \mid X_1, X_2, ..., X_n\} = \frac{1}{2} m''(x) \frac{d_2 (h(x))^2 (f^2(x) - (h(x))^2 (f'(x))^2 (\frac{d_4}{d_2}))}{(f^2(x) - (h(x))^2 (f'(x))^2 d_2)} + o_p(h^2) \tag{10}$$

Under the assumption on $f(x)$ has a bounded support on $[0,1]$ we find

$$E\{\hat{m}_V(x) - m(x) \mid X_1, X_2, ..., X_n\} = \frac{1}{2} m''(x) (h(x))^2 d_2 + o_p(h^2) \tag{11}$$

To approximate the conditional variance we need first to define:

$$r_\lambda(x,h(.)) = \begin{cases} n(h(x))^{\lambda-1} r_\lambda f(x) \sigma^2(x) + o_p(nh^\lambda) & \lambda \text{ even} \\ n(h(x))^{\lambda-1} r_\lambda f'(x) \sigma^2(x) + o_p(nh^\lambda) & \lambda \text{ odd} \end{cases}$$

Where $r_\lambda = \int u^\lambda K^2(u)du$
Then the conditional variance approximation of $\hat{m}_V(x)$ is:

$$Var\{\hat{m}_V(x) \mid X_1, ..., X_n\} = \underline{e}_1^T (X^T W X)^{-1} X^T \Sigma X (X^T W X)^{-1} \underline{e}_1 \tag{12}$$

Where

$\Sigma = diag\{K^2_{h(X_1)}(X_1 - x)\sigma^2(X_1), ..., K^2_{h(X_n)}(X_n - x)\sigma^2(X_n)\}$ is an $(n \times n)$ diagonal matrix.

$$X^T\Sigma X = \sum_{i=1}^{n} K^2_{h(X_i)}(X_i - x)\sigma^2(X_i)\begin{pmatrix} 1 & (X_i - x) \\ (X_i - x) & (X_i - x)^2 \end{pmatrix} = \begin{pmatrix} r_0(x, h(.)) & r_1(x, h(.)) \\ r_1(x, h(.)) & r_2(x, h(.)) \end{pmatrix}$$
$$= \begin{pmatrix} n(h(x))^{-1}r_0 f(x)\sigma^2(x) + o_p(n) & nr_1 f'(x)\sigma^2(x) + o_p(nh) \\ nr_1 f'(x)\sigma^2(x) + o_p(nh) & nh(x)r_2 f(x)\sigma^2(x) + o_p(nh^2) \end{pmatrix}$$

Hence the conditional variance becomes:

$$Var\{\hat{m}_V(x) \mid X_1, ..., X_n\} = \frac{\sigma^2(x)f(x)(f^2(x)r_0 - 2h(x)r_1 f'(x) + (h(x))^2 r_2 (f')^2(x))}{nh(x)(f^2(x) - (h(x))^2(f')^2(x)d_2)^2} + o_p\{(nh)^{-1}\}$$

If $f(x)$ has bounded support $[0, 1]$, then the conditional variance will be equal to:

$$Var\{\hat{m}_V(x) \mid X_1, ..., X - n\} = \frac{\sigma^2(x)r_0}{nh(x)f(x)} + o_p\{(nh)^{-1}\} \tag{13}$$

Theorem 3.1 (Fan and Gijbels 1992) Assume the conditions $(1-5)$ are satisfied and assume the weight function $w(.)$ is nonnegative, bounded and continuous function on (a, b) and that $f(x)$ is uniformly Lipschitz continuous of order $\nu > 0$. Suppose that $\lambda(.)$ is bounded and continuous function on (a, b) and $min\lambda(z) > 0$, then the conditional MISE of $\hat{m}_V(x)$ is:

$$E\{\int_{-\infty}^{\infty}(\hat{m}_V(x) - m(x))^2 w(x)dx \mid X_1, ..., X_n\} = \frac{d_2^2}{4}\int_{-\infty}^{\infty}(m''(x))^2(h(x))^4 w(x)dx + \tag{14}$$

$$\frac{r_0}{4}\int_{-\infty}^{\infty}\frac{\sigma^2(x)}{h(x)f(x)}w(x)dx + o_p\{(nh)^{-1} + h^4\}$$

Simple algebra yields the optimal global bandwidth:

$$h_{opt} = (\frac{r_0\int_{-\infty}^{\infty}\frac{\sigma^2(x)}{\lambda(x)f(x)}w(x)dx}{d_2^2\int_{-\infty}^{\infty}(m''(x))^2(\lambda(x))^4 w(x)dx})^{1/5}n^{-1/5} \tag{15}$$

Substituting this optimal choice in to (14) leads to:

$$AMISE(\hat{m}_V(x)) = \frac{5d_2^{\frac{2}{5}}r_0^{\frac{4}{5}}}{4n^{\frac{4}{5}}}(\int_{-\infty}^{\infty}(m''(x))^2(\lambda(x))^4 w(x)dx)^{1/5}(\int_{-\infty}^{\infty}\frac{\sigma^2(x)}{\lambda(x)f(x)}w(x)dx)^{4/5} \tag{16}$$

4. CHOICE OF COMBINING PARAMETER

One can choose a combining parameter by minimizing the least squares problem with respect to γ. That is we minimize:

$$\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2=\sum_{i=1}^{n}(Y_i-\gamma\hat{m}_V(X_i)-(1-\gamma)\hat{m}_F(X_i))^2$$

with respect to γ.
This will result to estimator of γ which is:

$$\hat{\gamma}=\frac{-\sum_{i=1}^{n}(Y_i-\hat{m}_F(X_i))(\hat{m}_F(X_i)-\hat{m}_V(X_i))}{\sum_{i=1}^{n}(\hat{m}_F(X_i)-\hat{m}_V(X_i))^2} \tag{17}$$

We must note that the above formula depends on the data hence we can get this estimator easily. Also we can put some restrictions on this estimator; if the value of this estimator is negative then we can put it equal to zero, and if its value is greater than one then we can set it equal to one.

5. BANDWIDTH CHOICE

There are different automatic bandwidth selectors that produce asymptotically optimal Kernel smoothers {for more details see Fan and Gijbels (1996), Härdle (1990), Hart (1997), Simonoff (1996)}. In this section we will discuss one of these methods which named rule of thumb.

5.1 RULE OF THUMB (ROT)

A rule of thumb method is a simple bandwidth selector obtained by using the global fixed bandwidth (15) and considering a reference parametric model for the unknown regression function $m(x)$. The following algorithm illustrates this selector.

5.2 ALGORITHM

- Take $m(.)$ to be a polynomial of degree four, i.e. $m(x)=a_0+a_1x+a_2x^2+a_3x^3+a_4x^4$. This choice would imply that $m''(x)$ is a quadratic function, leaving certain flexibility for fitting functions with a curvature.
- Fit the parametric model to the data.
- Fit the standardized residual sum of squares from this parametric fit $(\tilde{\sigma})^2$
- Substituting the estimated quantities $\tilde{\sigma}^2$ and $(\tilde{m}'')^2(x)$ in (15) will lead to the estimated expression

$$\hat{h}_{ROT}=\left(\frac{r_0\tilde{\sigma}^2\int_{-\infty}^{\infty}w_0(x)dx}{nd_2^2\int_{-\infty}^{\infty}(\tilde{m}''(x))^2w_0(x)f(x)dx}\right)^{1/5} \tag{18}$$

The denominator in (18) can be written as $nE\lfloor(\tilde{m}''(x))^2w_0(x)\rfloor$ which can be estimated by $\sum_{i=1}^{n}(\tilde{m}''(X_i))^2w_0(X_i))$.
- Then the Rule of thumb estimator is:

$$\hat{h}_{ROT}=\left(\frac{r_0\tilde{\sigma}^2\int_{-\infty}^{\infty}w_0(x)dx}{d_2^2\sum_{i=1}^{n}(\tilde{m}''(X_i))^2w_0(X_i)}\right)^{1/5} \tag{19}$$

For uniform density over $[a, b]$ and constant variance, the resultant ROT selector for h becomes:

$$\hat{h}_{ROT} = \left(\frac{r_0\tilde{\sigma}^2(b-a)}{nd_2^2\int_{-\infty}^{\infty}(\tilde{m}''(x))^2 w_0(x) f(x)dx}\right)^{1/5} \approx \left(\frac{r_0\tilde{\sigma}^2(b-a)}{d_2^2\sum_{i=1}^{n}(\tilde{m}''(X_i))^2 w_0(X_i)}\right)^{1/5} \quad (20)$$

6. SIMULATION RESULTS

In this section we investigate the performance of the proposed estimator and we present the results of comparison between this estimator with both fixed bandwidth and variable bandwidth local linear estimators.
In each of these experiments, we use Gaussian kernel. The following four regression functions were studied with the case of random design:

- $m_1(x) = x^2$
- $m_2(x) = \sin(\pi x)$
- $m_3(x) = \frac{1}{(x-0.3)^2+0.01} + \frac{1}{(x-0.9)^2+0.04} - 6$
- $m_4(x) = \sin(2x) + 2exp(-16x)$ See Fan and Gijbels (1992).

For the regression functions $m_1(x), m_2(x)$, and $m_3(x)$ we generate data according to $X \sim U(0,1)$ and $X \sim 2 - 2(1 - \frac{7}{8}U)^{\frac{1}{3}}, U \sim U(0,1)$. While for $m_4(x)$ we generate data according to $X \sim U(-2,2), X \sim N(0,1)$ and $X \sim 2 - 2(1 - \frac{7}{8}U)^{\frac{1}{3}}, U \sim U(0,1)$. The number of simulations samples (replications) was (500). The error terms were taken to be $N(0, \sigma^2)$ the same values of $\sigma^2(0.30, 0.09)$ were used with each function. Sample sizes $n = 50, 100$ were used for each experiment.
For a comparison we present the results of the three estimators with their empirical MASE (Mean Averaged Square error) and their respective standard errors.
Tables 1-2 summarize the results of the simulations. Each table has three lines (one line for each of the above estimators) and four columns.
The second column gives the value of the variance of error terms; the last three columns give the MASE (and the standard error) for the first three models with different sample sizes.
Table 3 also summarizes the simulation results but now by using the fourth model with different distributions of the random variable X as the columns.

Table 1: MASE and its standard error (in parenthesis) for the models 1,2 and 3 with $X \sim U(0,1)$

Regression function estimators	σ^2	$m_1(x)$		$m_2(x)$		$m_3(x)$	
N		50	100	50	100	50	100
LLSF (Fixed bandwidth)	0.30	0.0120(0.036)	0.0066(0.013)	0.022(0.11)	0.810(15.90)	843.3(16497)	61.3(431.8)
	0.09	0.0034(0.004)	0.0170(0.310)	0.023(0.25)	0.003(0.007)	3175(68605)	54.6(255.1)
LLSV (Variable bandwidth)	0.30	0.011(0.016)	0.007(0.004)	0.018(0.056)	0.014(0.040)	42.8(98.52)	33.0(9.7)
	0.09	0.006(0.026)	0.003(0.0012)	0.020(0.240)	0.003(0.0012)	42.4(56.95)	33.1(9.4)
Proposed estim.	0.30	0.0098(0.007)	0.0064(0.0038)	0.014(0.010)	0.011(0.020)	38.03(12.54)	32.98(6.36)
	0.09	0.0032(0.002)	0.0027(0.001)	0.005(0.003)	0.003(0.0011)	39.82(16.73)	33.40(7.07)

Table 2: MASE and its standard error (in parenthesis) for the models 1, 2 and 3 with $X = 2 - 2(1 - 7/8U)^{1/3}, t \sim U(0, 1)$

Regression function estimators	σ^2	$m_1(x)$		$m_2(x)$		$m_3(x)$	
N		50	100	50	100	50	100
LLSF (Fixed bandwidth)	0.30	0.019(0.22)	0.007(0.040)	0.017(0.098)	0.0124(0.08)	89.6(390.7)	64.2(249.8)
	0.09	0.016(0.32)	0.002(0.001)	0.30(6.530)	0.020(0.28)	64.0(19.90)	53.4(24.70)
LLSV (Variable bandwidth)	0.30	0.014(0.030)	0.009(0.005)	0.0156(0.01)	0.012(0.0050)	83.8(520.5)	47.0(55.1)
	0.09	0.004(0.002)	0.004(0.002)	0.020(0.25)	0.005(0.0013)	61.5(118.7)	45.7(22.6)
Proposed estim.	0.30	0.010(0.0220)	0.006(0.0040)	0.013(0.008)	0.010(0.0048)	53.9(13.9)	45.8(8.30)
	0.09	0.002(0.0015)	0.002(0.0012)	0.005(0.007)	0.004(0.0013)	53.3(14.3)	45.5(7.87)

Table 3: MASE and its standard error (in parenthesis) for the model 4 with $X \sim U(-2, 2), X \sim N(0, 1)$ and $X = 2 - 2(1 - 7/8U)^{1/3}, t \sim U(0, 1)$

Regression function estimatiors	σ^2	$X \sim U(-2, 2)$		$X \sim N(0, 1)$		$X \sim 2 - 2(1 - 7/8t)^{1/3}$, $t \sim U(0, 1)$	
N		50	100	50	100	50	100
LLSF (Fixed bandwidth)	0.30	0.81(13.3)	0.15(1.84)	1.33(22.3)	0.37(2.190)	0.0290(0.23)	0.012(0.0054)
	0.09	0.14(0.65)	0.09(0.73)	0.77(7.73)	3.26(42.46)	0.0087(0.04)	0.010(0.110)
LLSV (Variable bandwidth)	0.30	0.15(0.94)	0.060(0.019)	0.36(3.36)	0.14(0.33)	0.020(0.07)	0.010(0.0047)
	0.09	0.36(2.65)	0.047(0.0111)	0.37(4.57)	0.13(0.54)	0.009(0.08)	0.004(0.011)
Proposed estim.	0.30	0.075(0.04)	0.058(0.016)	0.12(0.045)	0.11(0.038)	0.016(0.013)	0.011(0.0049)
	0.09	0.073(0.21)	0.046(0.011)	0.11(0.040)	0.10(0.030)	0.004(0.002)	0.003(0.0012)

From tables 1-3 we see that the adequate performance of the proposed estimator is confirmed in simulated data. It has the smallest MASE and a fairly stable standard error for all the models simulated, sample sizes used and error variances, except in few cases which indicates a close behavior for all estimators.

Also as is well known we see that the values of MASE decrease with increasing samples size. The same behavior is true for error variances used although for some cases a positive relationship between increasing MASE and error variance was observed.

Finally we provided for each simulated model figures to give an indication of what our results will mean in terms of the actual curves, note that in each figure the bold solid curve is $m(x)$, the dashed solid is the regression function estimator, and from these figures we observe that the proposed estimator has better performance than the other estimators near the left hand side although a close behavior inside the interior interval .The skip near the left hand side due to the design points and we refer to Fan and Gijbels (1992) to this case.

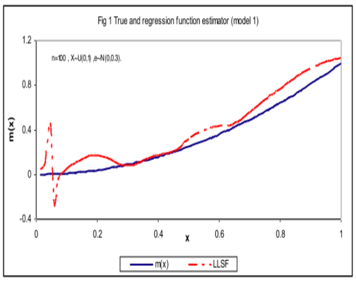


Fig 1 True and regression function estimator (model 1)
m(x)
x
m(x) LLSF

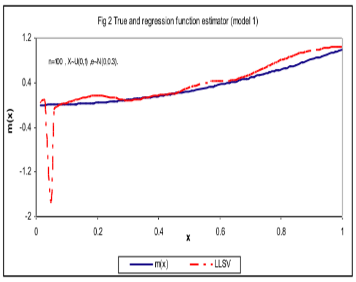


Fig 2 True and regression function estimator (model 1)
m(x)
x
m(x) LLSV

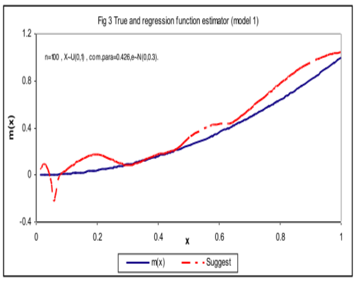


Fig 3 True and regression function estimator (model 1)
m(x)
x
m(x) Suggest

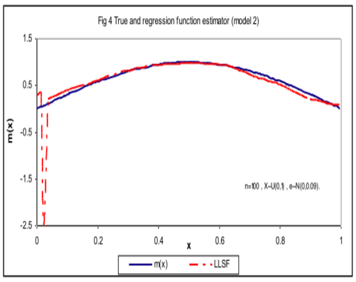


Fig 4 True and regression function estimator (model 2)
m(x)
x
m(x) LLSF

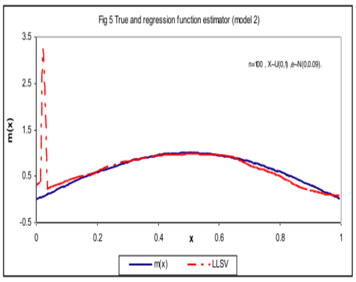


Fig 5 True and regression function estimator (model 2)
m(x)
x
m(x) LLSV

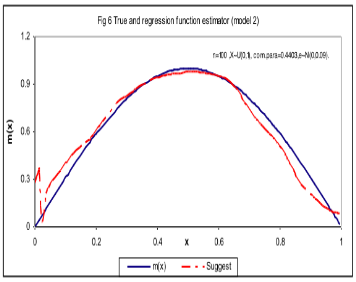


Fig 6 True and regression function estimator (model 2)
m(x)
x
m(x) Suggest

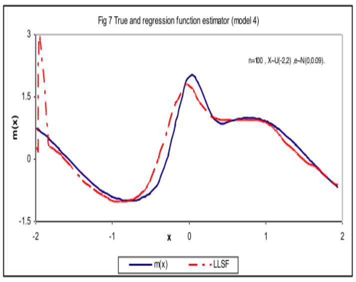


Fig 7 True and regression function estimator (model 4)
m(x)
x
m(x) LLSF

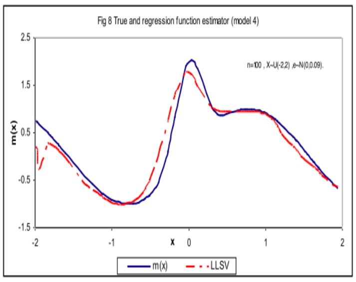


Fig 8 True and regression function estimator (model 4)
m(x)
x
m(x) LLSV

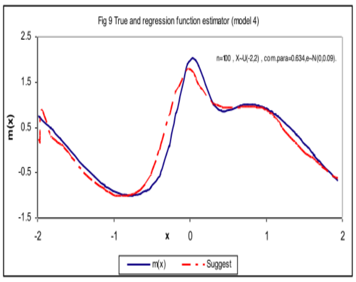


Fig 9 True and regression function estimator (model 4)
m(x)
x
m(x) Suggest

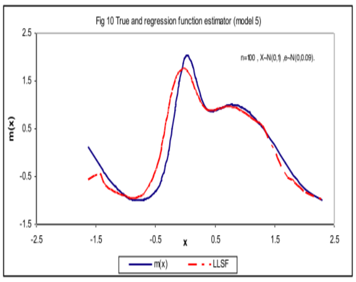


Fig 10 True and regression function estimator (model 5)
m(x)
x
m(x) LLSF

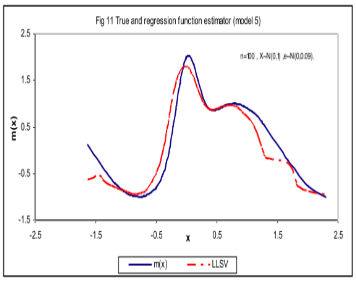


Fig 11 True and regression function estimator (model 5)
m(x)
x
m(x) LLSV

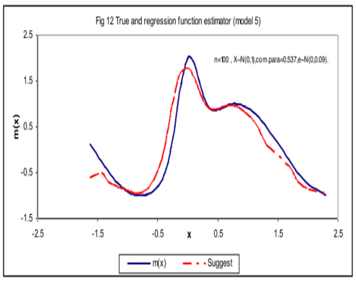


Fig 12 True and regression function estimator (model 5)
m(x)
x
m(x) Suggest

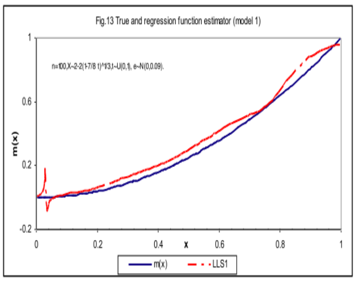


Fig.13 True and regression function estimator (model 1)
m(x)
x
m(x) LLS1

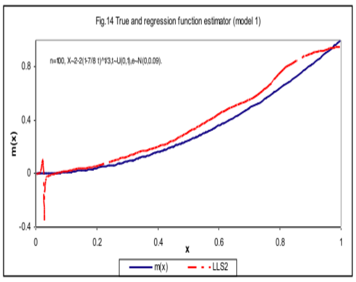


Fig.14 True and regression function estimator (model 1)
m(x)
x
m(x) LLS2

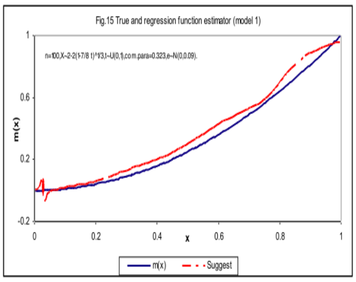


Fig.15 True and regression function estimator (model 1)
m(x)
x
m(x) Suggest

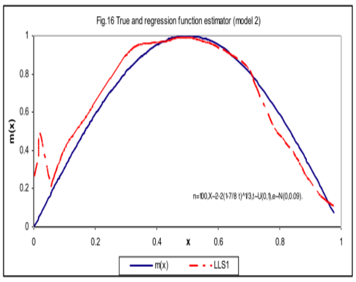


Fig.16 True and regression function estimator (model 2)
m(x)
x
m(x) LLS1

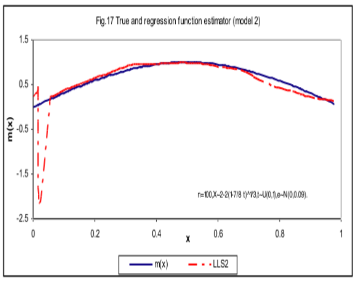


Fig.17 True and regression function estimator (model 2)
m(x)
x
m(x) LLS2

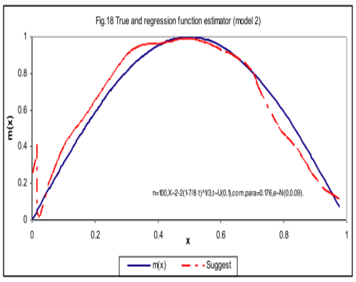


Fig.18 True and regression function estimator (model 2)
m(x)
x
m(x) Suggest

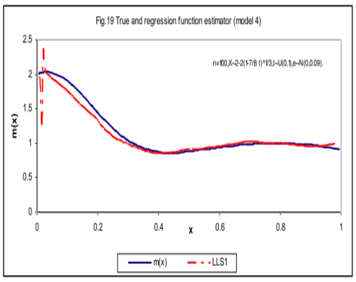


Fig.19 True and regression function estimator (model 4)

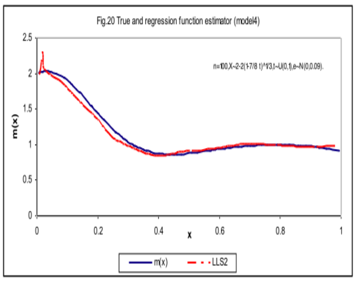


Fig.20 True and regression function estimator (model4)

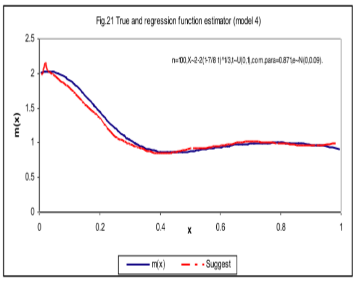


Fig.21 True and regression function estimator (model 4)

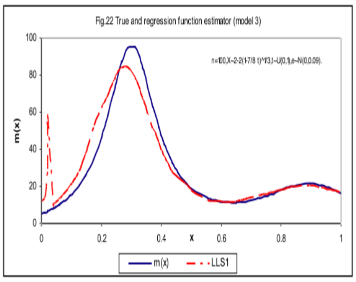


Fig.22 True and regression function estimator (model 3)

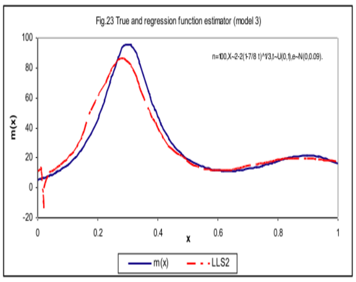


Fig.23 True and regression function estimator (model 3)

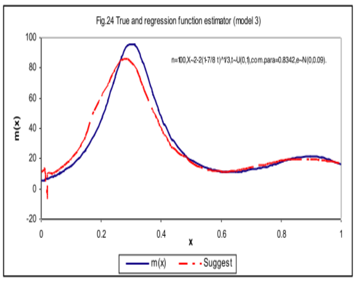


Fig.24 True and regression function estimator (model 3)

7. EMPIRICAL EXAMPLE

We now apply the estimators constructed previously to two different financial data. First real data set was (Buying and Selling Auction of the Central Bank of Iraq for US$ from 4/10/2003 to 31/12/2008) as a function in terms of the exchange rate of the Iraqi diners (ID) compared with the U.S. dollar ($) for the same period and the second data set was on the GDP of Iraq.

Figures (25), (26) and (27) present the plot of the actual data of (Buying and Selling Auction of the Central Bank of Iraq) and the aforementioned compared estimators respectively. We can observe preference the proposed estimated on the rest of the estimators, Notice that the estimated combining parameter was .

On the other hand the figures (28), (29) and (30) present the plot of the GDP of Iraq over the periods (1968-2010) with the aforementioned estimators' .From these figures equal performance of the proposed estimator with a variable bandwidth LLS is appear, i.e. . We note also converge the estimators (Proposed and LLSV) in figures (29) and (30) respectively with the real curve, this will refer to the acceptable representation of these estimators for the real curve.

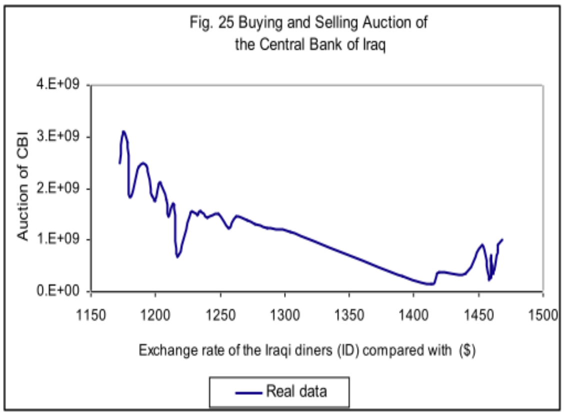


Fig. 25 Buying and Selling Auction of the Central Bank of Iraq

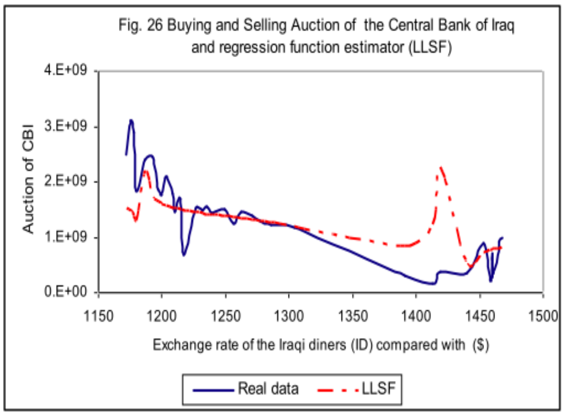
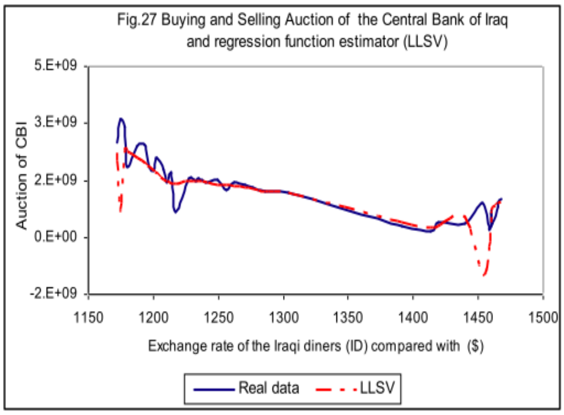
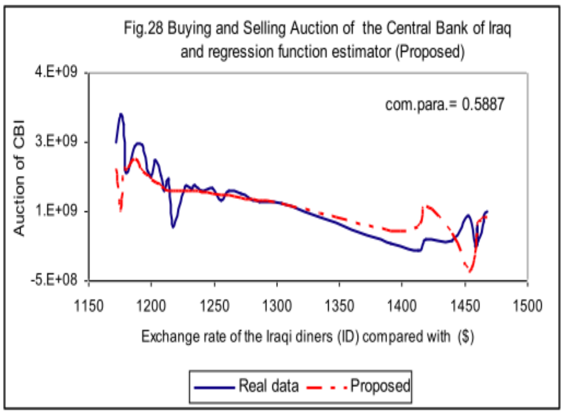
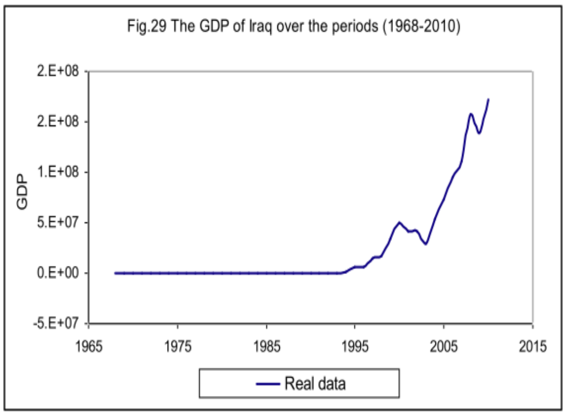
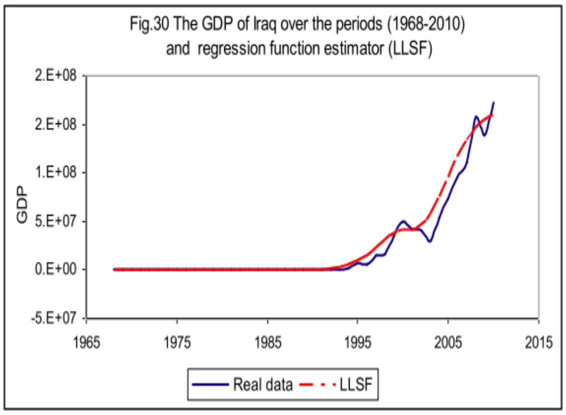


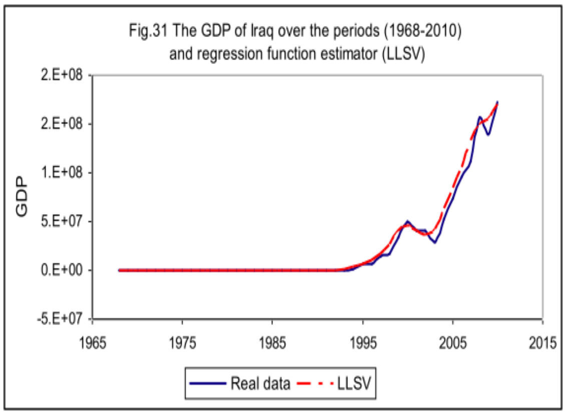
Fig. 26 Buying and Selling Auction of the Central Bank of Iraq and regression function estimator (LLSF)

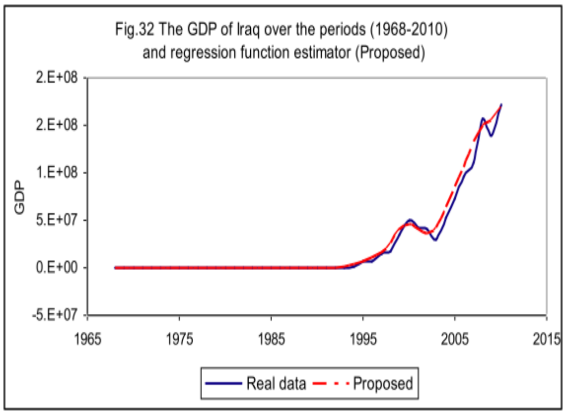












8. RECOMMENDATIONS

Finally, since the proposed estimator performs considerably well on data exhibiting varied patterns; we recommend its use in data analysis. Further studies can also be carried out to develop a general class of estimators of the type studies in this paper.

ACKNOWLEDGEMENTS

The authors wish to thank the Executive Editor, Prof. Víctor Leiva, and the anonymous reviewers for their constructive comments on the preliminary version of this paper

References

Breiman, L., Meisel, W., Purcell, E., 1977. Variable kernel estimators of multivariate densities. Technometrics, 19, 2,134-144.

Cheng, P.E., 1990. Applications of kernel regression estimators, a survey. Communications in Statistics-Theory and Methods, 19, 11, 4103-4134.

Cleveland, W.S., 1979. Robust locally weighted regression and smoothing scatter plots. Journal of the American Statistical Association, 74, 829-836.

Fan, J., 1992. Design adaptive nonparametric regression. Journal of the American Statistical Association, 87, 998-1004.

Fan, J., 1993. Local linear regression smoothers and their minimax efficiency. Annals of Statistics, 21, 196-216.

Fan, J., Gijbels, I., 1992. Variable bandwidth and local linear regression smoothers. Annals of Statistics, 20, 2008-2036.

Fan, J., Gijbels, I., 1995. Data driven bandwidth selection in local polynomial fitting: variable bandwidth and spatial adaptation. Journal of The Royal Statistical Society Series B-Statistical Methodology, 57,371-394.

Fan, J., Gijbels, I., 1996. Local Polynomial Modeling and Its Applications. Chapman and Hall, London.

Fan, J., Gijbels, I., 2000. Local Polynomial Fitting. In Smoothing and Regression, Approaches, Computation and Applications (M.G.Schimek, ed.).John Wiley & Sons, New York, 229-276.

Fan, J., Yao, Q., 2005. Nonlinear Time series, Nonparametric and Parametric Methods. Springer New York.

Härdle, W., 1990. Applied Nonparametric Regression. Cambridge University press, Boston.

Härdle, W., 1991. Smoothing Techniques with Implementation in S. Springer, New York.

Härdle, W., Marron, J.S., 1985. Optimal bandwidth selection in nonparametric regression function estimation. Annals of Statistics, 13, 1465-1481.

Härdle, W., Mller, M., Werwatz, A., Sperlich, S., 2004. Nonparametric and Semiparametric Models. Springer-Verlag Berlin Heidelberg.

Hart, J.D., 1997. Nonparametric Smoothing and Lack of Fit Tests. Springer, New York.

Jennen-Steinmetz, C., Gasser, T., 1988. A unifying approach to nonparametric regression estimation. JASA, 83, 1084-1089.

Loader, C., 1999. Bandwidth selection: classical or plug-in ? The Annals of Statistics, 27,415-438.

Mller, H.G., 1987. Weighted local regression and kernel methods for nonparametric curve fitting. JASA, 82, 231-238.

Mller, H.G., Stadtmüller, U., 1987. Variable bandwidth kernel estimators of regression curves. The Annals of Statistics, 15, 182-201.

Olkin, I., Spiegelman, C., 1987. A semiparametric approach to density estimation. JASA, 82, 858-865.

Rahman, M., Gokhale, D.V., Ullah, A., 1997. On testimation of a probability density: the normal case. Communications in Statistics - Theory and methods, 26, 263-278.

Sarda, P., Vieu, P., 2000. Kernel Regression. In Smoothing and Regression, Approaches, Computation and Applications (M.G.Schimek, ed.). John Wiley & Sons, New York, 43-70.

Silverman, B.W., 1984. Spline smoothing: the equivalent variable kernel method. Annals of Statistics, 12, 898-916.

Silverman, B.W., 1986. Density Estimation for Statistics and Data Analysis. Chapman and Hall, London.

Simonoff, J.S., 1996. Smoothing Methods in Statistics. Springer, New York.

Staniswalis, J.G. , 1989. Local bandwidth selection for kernel estimators. Journal of the

American Statistical Association, 84, 284-288.
Stone, C.J., 1977. Consistent nonparametric regression. Annals of Statistics, 5, 595-645.
Wand, M.P., Jones, M.C., 1995. Kernel Smoothing. Chapman and Hall, London.