Süleyman Demirel Üniversitesi
Fen Bilimleri Enstitüsü Dergisi
Cilt 22, Sayı 2, 511-520, 2018

Süleyman Demirel University
Journal of Natural and Applied Sciences
Volume 22, Issue 2, 511-520, 2018

DOI: 10.19113/sdufbed.70996

On the Adaptive Nadaraya-Watson Kernel Estimator for the Discontinuity in the Presence of Jump Size

Muhammad HANIF[1], **Shabnam SHAHZADI**[2], **Usman SHAHZAD**[3], **Nursel KOYUNCU**[*4]

[1,2,3]PMAS-Arid Agriculture University, Department of Mathematics and Statistics, Rawalpindi, PAKISTAN
[4]Hacettepe University, Faculty of Science, Department of Statistics, Ankara, TURKEY

(Alınış / Received: 01.08.2017, Kabul / Accepted: 30.04.2018, Online Yayınlanma / Published Online: 12.06.2018)

Keywords
Adaptive Nadaraya Watson,
Jump size,
Mean Integrated Square Error,
Rate of convergence.

Abstract: In this paper, we studied an Adaptive Nadaraya Watson kernel estimator to check the bias effect on both side of the discontinuity in the presence of jump size for regression discontinuity model. We have proposed the modified Adaptive Nadaraya Watson kernel estimator and derived its normality and variance. We have also compared with the asymptotic normality of the Mean Integrated Square Error (MISE) of Adaptive Nadaraya Watson kernel estimator and Nadaraya Watson kernel estimator. The results obtained from the simulation study have showed that Adaptive Nadaraya Watson estimator has better performance than the Nadaraya Watson Kernel estimator.

Sıçrama Büyüklüğünde Süreksizlik Olması Durumunda Uyarlanabilir Nadaraya-Watson Kernel Tahmin Edicisi

Anahtar Kelimeler
Uyarlanabilir Nadaraya Watson,
Sıçrama büyüklüğü,
Butunlesik Hata Kareleri Ortalaması,
Yakınsama oranı

Özet: Bu çalışmada, süreksiz regresyon modeli için sıçrama büyüklüğünün varlığında süreksizliğin her iki tarafı üzerine yan etkisini kontrol etmek için Uyarlanabilir Nadaraya Watson kernel tahmin edicisi çalıştık. Modifiye edilmiş Uyarlanabilir Nadaraya Watson kernel tahmin edicisi önerdik ve bunun normalliği ve varyansını çıkarsadık. Ayni zamanda, Uyarlanabilir Nadaraya Watson kernel tahmin edicisi ve Nadaraya Watson kernel tahmin edicilerinin butunlesik hata kareler ortalaması ile karsılaştırma yaptık. Benzetim çalışması sonucunda Uyarlamalı Nadaraya Watson tahmin edicisinin, Nadaraya Watson tahmin edicisinden daha iyi performans gösterdiği görüldü.

1. Introduction

Nonparametrics: For last few years, nonparametric regression has become important tool for data smoothing. Most commonly used estimates of nonparametric regression functions including kernel estimates based on smoothness of the regression functions. In many applications, the function is to be estimated as discontinuities or threshold points. Gao, et al.[1], for example, when studying the impact of advertising, the time at which this action takes place impact could effectively be modeled by the location of a jump point and the magnitude of the effect of this action is measured by the jump size. If we ignore the jump point we make serious error in order to draw the inference about the processes under study. Similarly also see Yin [2] for estimating locations of discontinuity points of the regression function.

Regression discontinuity design becomes one of the useful designs when there is threshold point in the treatment or in the probability assignment. The treatment assignment under the weak smoothness becomes random near the threshold point. Regression Discontinuity model is mostly used only for the information that is very close to threshold point. In Regression Discontinuity model lacks of smoothness is not only the problem but the size of discontinuity is also important. It is useful to estimate the conditional expectation of the boundary points to check the difference of results given by boundary estimation (see Porter [3]).

Demir and Toktamis [4] considered Adaptive Nadaraya Watson kernel estimator to estimate the regression function. Our paper is also composed of Adaptive Nadaraya Watson as a jump size instead of Nadaraya Watson. Because the main drawback of the Nadaraya Watson estimator used as jump size in regression discontinuity model has poor asymptotic bias behavior whereas Adaptive Nadaraya Watson overcomes this problem.

Variance estimation of regression discontinuity was constructed by Hardle [5]. Pagan and Ullah [6] also gave the consistent estimation of variance for the left hand limit i.e. σ^{2-} and right hand limit i.e. σ^{2+} of the derivatives. Similarly Silverman [7] gave density of the discontinuity of the jump size, $f_0(\bar{u})$ which can be estimated consistently

* *Corresponding author: nkoyuncu@hacettepe.edu.tr*

by kernel density estimation.

In this paper, we have proposed Adaptive Nadaraya Watson kernel estimator for the estimation of σ^{2+} , σ^{2-} for jump size in regression discontinuity model. Further, we have derived the asymptotic properties of σ^{2+} , σ^{2-}. Marginal and hence optimal rate of convergence is derived theoretically to estimate the treatment effect in regression discontinuity by using Stone's [8] definition. The properties such as MSE, MISE and bias effect on the left and right hand side of the discontinuity of regression discontinuity model are also derived. The performance of proposed regression discontinuity model is compared with the existing once numerically.

The rest of the paper is organized as follows: Section 2 includes Nadaraya Watson kernel estimator as a jump size in the regression discontinuity model and technical lemmas are discussed. In section 3, proposed work based on Adaptive Nadaraya Watson kernel estimator and its assumptions are discussed. Simulation study is performed in Section 4. Real life data is used to see performance of proposed estimator in Section 5. The derivations of the Section 3 and Section 4 are available in Appendix .

2. Nadaraya Watson Estimator and Technical Lemmas

As the previous work is based on Nadaraya Watson kernel estimator used as a jump size in regression discontinuity model. The connection between the regression discontinuity model and the treatment effect have been discussed in Hahn et al.[9]. Trochim [10] distinguished among the two dissimilar connections between the Regression Discontinuous model, that depends upon the treatment assignment connected with the observed variable. The treatment assignment given by the indicator variables $d\varepsilon(0,1)$ are defined as;

$$\begin{cases} 1 \;\; if u \;\; > \bar{u} \\ 0 \;\; if u \;\; < \bar{u}. \end{cases}$$

where u is observed variable with the known threshold point $\bar{u}$. Let y_0 and y_1 be the potential outcomes parallel to the two treatment assignments, and as usual $Y = dy_1 + (1-d)y_0$ is the observed outcomes. By using the smoothness assumption that $E(\frac{y_i}{u})$ is continuous at u for $j = 0,1$. So the expected casual effect of the treatment effect can be identified at the discontinuity point.

$$\alpha = \lim_{u\downarrow\bar{u}} E(y/u) - \lim_{u\uparrow\bar{u}} E(y/u) \tag{2.1}$$

$$\hat{\alpha}(\bar{u}) = \frac{\frac{1}{n}\sum_{i=1}^{n} K_h(\bar{u}-U_i)y_i d_i}{\frac{1}{n}\sum_{j=1}^{n} K_h(\bar{u}-U_j)d_j} - \frac{\frac{1}{n}\sum_{i=1}^{n} K_h(\bar{u}-U_i)y_i(1-d_i)}{\frac{1}{n}\sum_{j=1}^{n} K_h(\bar{u}-U_j)(1-d_j)}$$

Following Lammas are useful in obtaining the main results of our paper (see Porter [3]).

Lemma A_1. Suppose the kernel k is bounded, symmetric, zero outside of a bounded set M and Lipschitz, and f_0 is continuous on N_1. If $\frac{nh}{lnn} \to \infty$, then $sup_{x\varepsilon N_{0-}}|\hat{f_+}(x) - E\hat{f_+}(x)| = O_p(\sqrt{\frac{lnn}{nh}})$,$sup_{x\varepsilon N_{0+}|\hat{f_-}(x) - E\hat{f_-}(u)|} = O_p\sqrt{\frac{lnn}{nh}}$, and $sup_{x\varepsilon N_0}|\hat{f}(x) - E\hat{f}(x)| = O_p(\sqrt{\frac{lnn}{nh}})$

Lemma A_2. Suppose the kernel k is bounded, symmetric, zero outside of a bound set M and Lipschitz, f_0 is a continuous on N_1,$sup_{x\varepsilon N_1} E[|y|^{2+\xi}/x] < \infty$ for some $\xi > 0$ and $n^{\frac{\xi}{(2+\xi)h}}/(lnn) \to \infty$. The $sup_{x\varepsilon N_0}|\tilde{r}(x) - E\tilde{r}(x)| = O_p(\sqrt{\frac{lnn}{nh}})$.

Lemma A_3 Suppose the kernel k is a bounded, symmetric, zero outside of a bounded set $[-M,M]$. On N_1, f_0 is continuously differentiable for $x \neq \bar{x}$, $x\varepsilon N_1$, and m is continuous at $\bar{x}$ with finite right and left hand derivatives. Then $sup_{u\varepsilon[\bar{x},\bar{x}+Mh]}|E\tilde{r}(x) - r(x)| = O(h)$

3. Proposed Estimator and Assumptions

The main purpose of our study is to minimize the bias effect on the left and right side of the discontinuity by using Adaptive Nadaraya Watson Kernel estimator as a jump size. We also obtained the optimal rate of convergence through simulation study.
Consider the random-design regression model given by;

$$Y_i = m(u_i) + \varepsilon_i \tag{3.1}$$

where m is an unknown regression function with compact interval $[0,1]$ and ε_i is the observation error which is independently identically distributed with mean 0 and variance σ^2. In the discontinuity model we have a cut off point existed for m function whereas cut off point $(\bar{u})$ is $0 < \bar{u} < 1$. Usually regression function is defined as;

$$y = m(u) + \alpha d + \varepsilon$$

and

$$0 < u < 1 \tag{3.2}$$

Here, $m(u)$ represents a continuous function defined on $[0,1]$. α=jump size and can be defined as

$$\alpha(\bar{u}) = m_2(\bar{u}) - m_1(\bar{u}) \tag{3.3}$$

where $\bar{u}$= jump size at cut off point
Basically the jump size at the possible cut off point $\bar{u}$ is $m_1(\bar{u}) = \lim_{u\downarrow\bar{u}} m(u)$ is at the right of the discontinuity curve and $m_2(\bar{u}) = \lim_{u\uparrow\bar{u}} m(u)$ is at the left side of the discontinuity curve.

$$\begin{aligned} \alpha &= E[y_1 - y_0/u] \\ &= E[y/\bar{u}] - E[y_0/\bar{u}] \\ &= \lim_{u\downarrow\bar{u}} E[y_1/\bar{u}] - \lim_{u\uparrow\bar{u}} E[y_0/\bar{u}] \\ &\lim_{u\downarrow\bar{u}} E[y/\bar{u}] - \lim_{u\uparrow\bar{u}} E[y/\bar{u}] \end{aligned}$$

where

$$\hat{\alpha}(\bar{u}) = \frac{\frac{1}{n}\sum_{i=1}^{n} K_{h\lambda_i^*}(\bar{u}-U_i)y_i d_i}{\frac{1}{n}\sum_{j=1}^{n} K_{h\lambda_j^*}(\bar{u}-U_j)d_j} - \frac{\frac{1}{n}\sum_{i=1}^{n} K_{h\lambda_i^*}(\bar{u}-U_i)y_i(1-d_i)}{\frac{1}{n}\sum_{j=1}^{n} K_{h\lambda_j^*}(\bar{u}-U_j)(1-d_j)}$$

$$\lambda_i^* = [\frac{\hat{f}(u)}{\bar{u}}]^{-0.5}$$

and $d = 1(u_i \geq \bar{u})$, $I(A)$ is an indicator for event A. u_i represents a random variable and d is an indicator whereas $h\lambda_i^*$ shows adaptive bandwidth that is used to control the size of local neighborhood on average. First term in the $\hat{\alpha}(\bar{u})$ is a weighted average depends on the distance from the discontinuity $(\bar{u} - U_i)$ whereas α is the jump size of the discontinuity model.
The cut-off point $\bar{u}$ provides a chance to observe the average difference in the potential outcomes from the points on either side of the discontinuity. The main point of our estimation is that in all cases, the casual effect is found out from any of the expression that is only involve in the size of the discontinuity in the conditional expectation.
Necessary assumptions required to derive the limiting distribution for the estimator are:

Assumption 1. Choice for the kernel estimator:
$a)$ Kernel estimator is symmetrically bounded, Lipchitz function and bounded .

$$\int_{-\infty}^{+\infty} k(v)dv = 1$$

$b)$ For any positive integer, $\int k(v)v^j dv = 0 \quad 1 < j < r-1$ and $\quad r \geq 3$
Let's suppose q_o be the any marginal density function of y and $m(y)$ denotes the conditional expectation of z given y minus discontinuity. So, $m(y) = E(z/y) - \alpha 1[y \geq \bar{y}]$.

Assumption 2.
$a)$ Smoothness on any side of the discontinuity for compact interval M of $\bar{y}$ is with $M \subset (\bar{y})$, but it allows for unequal left and right side derivatives of m. Let's q_o is l_q times continuously differentiable and it is bounded away from zero, $m(y)$ is l_m times continuously differentiable for $\bar{y} \in M/\bar{y}$.
$b)$ Results of the limiting distribution have no effect on the Adaptive Nadaraya Watson estimator, but play an important role in the asymptotic biasness of the subsequent estimator. Whereas left and right hand side of the discontinuity of m to order l_q are equal at cut off point$(\bar{y})$.

Assumption 3.
$a)$

$$E(\varepsilon/(y,d)) = 0$$

$b)$ $\sigma^2(y) = E(\varepsilon^2/y)$ is continuous for $y \neq (\bar{y})$ and left and right hand side of the limits exist at $\bar{y}$.
$c)$ For some $\xi > 0, E(|\varepsilon|^{2+\xi} y)$ is uniformly bounded in a compact interval M.
$d)$ The marginal density $f(y)$ of y is continues on the compact interval M.
To estimate the Adaptive Nadaraya Watson kernel estimator, we use the estimators $\hat{f}_r(u)$ and $\hat{f}_r(u,y)$ of the density function to estimate the regression function. We obtain the adaptive Nadaraya Watson kernel estimator with varying bandwidths as follows

$$\hat{m}_{ANW}(u) = \int \frac{y\hat{f}_r(u,y)}{\hat{f}_r(u)} dy$$

$$= \frac{\sum_{i=1}^n \frac{Y_i}{\lambda_i^*} K(\frac{u-U_i}{\lambda_i^*})}{\sum_{i=1}^n \frac{1}{\lambda_i^*} K(\frac{u-U_i}{h\lambda_i^*})}$$

$$\lambda_i^* = [\frac{\hat{f}(U_i)}{\bar{u}}]^{-\alpha}$$

and

$$\alpha = 0.5$$

This idea is formalized in the following theorem, which is based on the limiting distribution of the adaptive Nadaraya Watson estimator at a boundary point.

Theorem 1. By using assumption 1(a), 2(a) and assumption 3 holds l_q with any positive integer and l_r with any negative integer. If $h\lambda_i^* \to 0, nh\lambda_i^* \to \infty$ and $h\lambda_i^* \sqrt{nh} \to p$ where $0 \leq p \leq \infty$

$$\sqrt{nh\lambda_i^*}(\hat{\alpha} - \alpha) \xrightarrow{d} (p2k_1(0)(m'^{+}(\bar{u}) + m'^{-}(\bar{u})), 4\delta_0 \frac{\sigma^{2+}(\bar{u}) + \sigma^{2-}(\bar{u})}{f_0(\bar{u})})$$

Remarks.
Observations used in the first and second terms of the difference defining the $\hat{\alpha}$ are independent. For Adaptive Nadaraya Watson estimator at its interior point shows that its asymptotic bias is written in a form that underscores its dependence on the rate with adaptive bandwidth approaches to zero. When $p = 0$ we are "under smoothness" and the asymptotic bias is zero. When $h\lambda_i^* \sim (n^{-4/5})$ and, $\hat{\alpha}$ achieves its fastest rate of convergence at $(n^{-4/5})$ than the asymptotic bias term is considered. From this, we see that the bias of $\hat{\alpha}$ is of order $o(h\lambda_i^*)$. Whereas, from the theorem we find that higher-order bias-reducing kernels do not affect the order of asymptotic bias. Also, left and right hand derivates of the cut-point are equal or not equal do not affect the order of the asymptotic bias.

Theorem 2. Suppose Assumption 1(a), 2(a), and 3 with l_q and l_m are positive integers and $\xi \geq 2$. If $\frac{\sqrt{nh\lambda_i^*}}{lnn} \to \infty$and $\hat{\alpha} \to^p \alpha$, then $\hat{\delta}^{2+}(\bar{y}) \to^p \delta^{2+}(\bar{y})$and $\hat{\delta}^{2-}(\bar{y}) \to^p \delta^{2-}(\bar{y})$.

4. Simulation Study

We have performed practical strategy for Adaptive Nadaraya Watson and Nadaraya Watson kernel estimator to provide simulation evidence for finite sample performance. Hence our objective is to estimate the discontinuity at particular point, for that we use unbiased cross validation which was proposed by Hall and Schucany [11] for the density estimation. Simulation study is used to compare the performance of the Nadaraya Watson kernel estimator and Adaptive Nadaraya Watson kernel estimators. For the simulation study we consider the regression discontinuity function;

$$y = m(u) + \alpha d + \varepsilon$$

$$m(u) = 1 - u_i + exp^{(-200*(u_i-0.5)^2)}$$

$$\hat{\alpha}(\bar{u}) = \frac{\frac{1}{n}\sum_{i=1}^{n} K_{h\lambda_i^*}(\bar{u}-u_i)y_i d_i}{\frac{1}{n}\sum_{j=1}^{n} K_{h\lambda_j^*}(\bar{u}-u_j)d_j} - \frac{\frac{1}{n}\sum_{i=1}^{n} K_{h\lambda_i^*}(\bar{u}-u_i)y_i(1-d_i)}{\frac{1}{n}\sum_{j=1}^{n} K_{h\lambda_j^*}(\bar{u}-u_j)(1-d_j)}$$

where, u_i is uniformly distributed with interval [0,1]. Hardle [5] had given that error term ε_i is normally distributed with mean 0 and variance (0.1). We have generated sample of size 50,100,250,500,1000 for the fixed bandwidth and adaptive bandwidth. We have used the Epanechnikov and Gussain kernel density function for the simulation. For each group of simulation, we have calculated mean square error (MSE), mean integrated square error(MISE), bandwidth and jump size of the proposed model which we have considered. The number of the replication is 1000 and for the varying sample sizes.
In the Table 2 Adaptive Bandwidth, Jump Size, MSE and MISE are minimized as compare to Table 1 as we increase the sample size and we can also see that the rate of convergence of Adaptive Nadaraya Watson is faster than Nadaraya Watson estimator.

5. Real Life Data

Data is taken from Gross Domestic Product, Current Prices. Values are based upon GDP in national currency converted to U.S. dollars using market exchange rates (yearly average). Exchanges rate projections are provided by country econometrics for the group of other emerging market and developing countries. Exchanges rates of advanced economics are established in the WEO assumptions for each WEO exercise. The data is summarized in Table3. (Source of Data:International Monetary Fund, World Economic Outlook Database, April 2015)

6. Conclusion

In this study, we concluded that the results obtained from the Adaptive Nadaraya Watson Kernel estimator gives better result than the Nadaraya Watson estimator used in discontinuity model. We showed it by finding Mean Square Error and Rate of convergence. From Table 4 and 5 we can see that the MSE of proposed estimator is less than the MSE of existing estimator. We can also check the rate of convergence of proposed Adaptive Nadaraya Watson estimator is faster than existing Nadaraya Watson estimator.

References

[1] Gao, J., Pettitt, A.N., Wolff, R.C.L. 1998. Local Linear Kernel Estimation for Discontinuous Nonparametric Regression Function. Communication in Statistics - Theory and Methods, 27(12), 2871-2894.

[2] Yin, Y.Q. 1988. Detection of the Number, Locations and Magnitudes of jumps. Communication in Statistics. Stochastic Models, 4, 445-455.

[3] Porter, J. 1998. Estimation of Regression Discontinuities. Seminar Notes.

[4] Demir, S., Toktamis, O. 2010. On the Adaptive Nadaraya Watson Kernel Regression Estimators. Hacettepe Journal of Mathematics and Statistics, 39(3), 429-437.

[5] Hardle, W. 1990. Applied Nonparametric Regression. Cambridge University Press.

[6] Pangan, A., Ullah, A. 1999. Nonparametric Econometrics. Cambridge University Press, New York.

[7] Silverman, B. 1992. Density Estimation for Statistics and Data Analysis. Chapman and Hall, New york.

[8] Stone, C. 1980. Optimal Rates of Convergence for Nonparametric Estimators. The Annals of Statistics, 8, 1348-1360.

[9] Hahn, J., Todd, P., Klaauw, V.W. 2001. Identification and Estimation of Treatment Effects with a Regression Discontinuity Design. Econometrica , 69(1), 201-209.

[10] Trochim, W. 1984. Research Design for Program Evaluation; The Regression Discontinuity Approach.

[11] Hall, P., Schucany W. R. 1989. A Local Cross-validation Algorithm. Statistics and Probability Letters, 8, 109-117.

Appendix

Proof. To estimate the Adaptive Nadaraya Watson estimator we express $m(\bar{u})$ in term of probability density function $pdf\ f(u,y)$. As we have

$$f(x) = \frac{1}{nh\lambda_i^*}K(\frac{u-U_i}{h\lambda_i^*}) \quad (5.1)$$

whereas;

$$\lambda_i^* = [\frac{\hat{f}(u)}{\bar{x}}]^{-0.5}$$

Where α is sensitivity parameter which varies between $(0,1)$ or we can write it as $0 < \alpha < 1$.Here we take $\alpha = 0.5$ As, we have

$$m(\bar{u}) = E[Y/U_{=u}] = \int_{-\infty}^{+\infty} yf(y/u)dy = \frac{\int_{-\infty}^{+\infty} yf(u,y)dy}{\int_{-\infty}^{+\infty} f(u)du}$$

$$\hat{m}(\bar{u}) = \frac{\int_{-\infty}^{+\infty} y\hat{f}(u,y)dy}{\int_{-\infty}^{+\infty} \hat{f}(u)du}$$

$$= \frac{1}{\hat{f}(u)}\int_{-\infty}^{+\infty} y\hat{f}(u,y)dy$$
$$= \frac{1}{\hat{f}(u)}\int_{-\infty}^{+\infty} y\frac{1}{n}\sum_{i=1}^{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^* h\lambda_j^*}K(\frac{u-U_i}{h\lambda_i^*})K(\frac{y-Y_i}{h\lambda_i^*})dy$$
$$= \frac{1}{\hat{f}(u)}\sum_{i=1}^{n}\frac{1}{nh\lambda_i^*}k(\frac{u-U_i}{h\lambda_i^*})\int_{-\infty}^{+\infty}\frac{y}{nh\lambda_i^*}(\frac{y-Y_i}{h\lambda_i^*})dy$$

By using the transformation $t = \frac{(y-Y_i)}{h\lambda_i^*}$ which is given as;

$$\hat{m}(\bar{u}) = \frac{1}{\hat{f}(x)}\sum_{i=1}^{n}\frac{Y_i}{nh\lambda_i^*}k(\frac{u-U_i}{h\lambda_i^*})\int [h\lambda_i^* t + uY_i]K(t)dt$$

Table 1. Nadaraya Watson Estimator with Epanechikov as Density Function

sample size	Bandwidth	Jump Size	MSE	MISE	Rate Convergence
50	0.144348	0.3080614	0.5195428	0.2859086	0.007906131
100	0.06178708	0.2717062	0.4974971	0.2674444	0.003249796
250	0.01923375	0.2200431	0.4241645	0.2305969	0.00249796
500	0.00564366	0.1728534	0.40447052	0.1905969	0.001243869
1000	0.00307073	0.1471928	0.2975396	0.1605706	0.0003568771

Table 2. Adaptive Nadaraya Watson Estimator with Epanechikov as Density Function

sample size	Bandwidth	Adaptive (h)	Jump Size $\hat{\alpha}(\bar{u})$	MSE	MISE	Rate Convergence
50	0.09860128	0.01394433	0.2432769	0.4313248	0.2088481	0.01577945
100	0.05908542	0.00373689	0.2408345	0.4048373	0.1945296	0.003941348
250	0.04201298	0.00273689	0.1728505	0.4047029	0.1554121	0.002047676
500	0.0018066	0.0018066	0.1540415	0.3187659	0.12592683	0.001769833
1000	0.02674475	0.0008457432	0.0994501	0.2296714	0.0183539	0.0003615997

By using the property of kernel that is;

$$\int_{-\infty}^{+\infty} k(v)dv = 1, \int_{-\infty}^{+\infty} vk(v)dv = 0$$

Hence

$$\hat{m}(\bar{u}) = \frac{1}{nh\lambda_i^*}\sum_{i=1}^{n} k(\frac{u-U_i}{h\lambda_i^*})[h\lambda_i^* + Y_i]$$

After simplification we have;

$$= \frac{1}{\hat{f}(u)}\sum_{i=1}^{n}\frac{1}{n}k(\frac{u-U_i}{h\lambda_i^*})Y_i \qquad (5.2)$$

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{nh\lambda_i^*}k(\frac{u-U_i}{h\lambda_i^*})$$

By replacing in equation (5.1) we have;

$$\lambda_i^* = [\frac{\hat{f}(u)}{\bar{x}}]^{-0.5}$$

$$\hat{m}(\bar{u}) = \frac{\sum_{i=1}^{n}\frac{Y_i}{nh\lambda_i^*}k(\frac{u-U_i}{h\lambda_i^*})}{\sum_{i=1}^{n}\frac{1}{nh\lambda_i^*}k(\frac{u-U_i}{h\lambda_i^*})}$$

Hence, we obtain the Adaptive Nadaraya Watson kernel regression estimator as: By simplify we have;

$$\hat{m}(\bar{x}) = \frac{\sum_{i=1}^{n}\frac{Y_i}{nh\lambda_i^*}k(\frac{u-U_i}{h\lambda_i^*})}{\sum_{i=1}^{n}\frac{1}{n\lambda_i^*}k(\frac{u-U_i}{h\lambda_i^*})} \qquad (5.3)$$

Replacing it in equation (5.3) we have estimated jump size is

$$\hat{\alpha}(\bar{u}) = \frac{\frac{1}{nh\lambda_i^*}\sum_{i=1}^{n}K_{h\lambda_i^*}(\bar{u}-u_i)y_i d_i}{\frac{1}{nh\lambda_j^*}\sum_{j=1}^{n}K_{h\lambda_j^*}(\bar{u}-u_j)d_j}$$

$$-\frac{\frac{1}{nh\lambda_i^*}\sum_{i=1}^{n}K_{h\lambda_i^*}(\bar{u}-u_i)y_i(1-d_i)}{\frac{1}{nh\lambda_j^*}\sum_{j=1}^{n}K_{h\lambda_j^*}(\bar{u}-u_j)(1-d_j)}$$

Proof of Theorem 1.
Let q denote a positive generic constant and as we know M is a compact interval. Let suppose M_0 also be compact interval such that $\bar{u}\varepsilon\ int(M_0)$ and $M_0 \subset int(M)$and by using assumption 1(a), suppose we support of the kernel k is [-M,M]. Hence the observation just to the right of the discontinuity are more likely to be greater than the intercept $m(\bar{u})+\alpha$ giving upward biassed. similarly, an average of observations just to the left of the discontinuity would provide a downward biassed estimate $m(\bar{u})$.
We have

$$\hat{\alpha}(\bar{u}) = (m(\bar{u})+\alpha)\frac{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k_{h\lambda_i^*}(\bar{u}-u_i)d_i y_i}{\frac{1}{n}\sum_{j=1}^{n}\frac{1}{h\lambda_j^*}k_{h\lambda_j^*}(\bar{u}-u_j)d_j}$$

$$-(m(\bar{u}))\frac{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k_{h\lambda_i^*}(\bar{u}-u_i)(1-d_i)y_i}{\frac{1}{n}\sum_{j=1}^{n}\frac{1}{h\lambda_j^*}k_{h\lambda_j^*}(\bar{u}-u_j)(1-d_j)}$$

By replacing

$$y_i = m(u_i)+\varepsilon_i$$

And rearranging the above equation, we have

$$\hat{\alpha}-\alpha = \frac{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{u}-u_i}{h\lambda_i^*})d_i[m(u_i)-m(\bar{u})+\varepsilon_i]}{\frac{1}{n}\sum_{j=1}^{n}\frac{1}{h\lambda_j^*}k(\frac{\bar{u}-u_j}{h\lambda_j^*})d_j}$$

$$-\frac{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{u}-u_i}{h\lambda_i^*})(1-d_i)[m(u_i)-m(\bar{u})+\varepsilon_i]}{\frac{1}{n}\sum_{j=1}^{n}k\frac{1}{h\lambda_j^*}(\frac{\bar{u}-u_j}{h\lambda_j^*})(1-d_j)}$$

Multiple both sides with $\sqrt{nh\lambda_i^*}$, we have $\sqrt{nh\lambda_i^*}(\hat{\alpha}-\alpha)$;

$$\sqrt{nh\lambda_i^*}(\hat{\alpha}-\alpha) = \frac{\frac{\sqrt{nh\lambda_i^*}}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{u}-u_i}{h\lambda_i^*})d_i[m(u_i)-m(\bar{u})+\varepsilon_i]}{\frac{1}{n}\sum_{j=1}^{n}\frac{1}{h\lambda_j^*}k(\frac{\bar{u}-u_j}{h\lambda_j^*})d_j}$$

$$-\frac{\frac{\sqrt{nh\lambda_i^*}}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{u}-u_i}{h\lambda_i^*})(1-d_i)[m(u_i)-m(\bar{u})+\varepsilon_i]}{\frac{1}{n}\sum_{j=1}^{n}\frac{1}{h\lambda_j^*}k(\frac{\bar{u}-u_j}{h\lambda_j^*})(1-d_j)}$$

By simplification we have

$$\sqrt{nh\lambda_i^*}(\hat{\alpha}-\alpha) = \frac{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\sqrt{h\lambda_i^*}}k(\frac{\bar{u}-u_i}{h\lambda_i^*})d_i[m(u_i)-m(\bar{u})+\varepsilon_i]}{\frac{1}{n}\sum_{j=1}^{n}\frac{1}{h\lambda_j^*}k(\frac{\bar{u}-u_j}{h\lambda_j^*})d_j}$$

Table 3. Real Data: Year and Gross Domestic Product Current Prices

year	prices	year	prices	year	prices
1980	2.016	1981	2.309	1982	2.378
1983	2.403	1984	2.372	1985	2.423
1986	2.68	1987	2.659	1988	2.621
1989	2.879	1990	2.301	1991	1.381
1992	0.874	1993	1.513	1994	2.446
1995	2.986	1996	3.315	1997	2.36
1998	2.708	1999	3.411	2000	3.64
2001	4.065	2002	4.435	2003	5.747
2004	7.315	2005	8.158	2006	9.001
2007	10.698	2008	12.901	2009	12.093
2010	11.927	2011	12.891	2012	12.345
2013	12.916	2014	13.262		

Table 4. Nadaraya Watson Estimator With Epanechikov As Density Function (Existing Model)

	sample size	Bandwidth	Jump Size	MSE	rate convergence
Existing Model	35	0.1104331	1.786687	36.93614	0.005906131

$$-\frac{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\sqrt{h\lambda_i^*}}k(\frac{\bar{u}-u_i}{h\lambda_i^*})(1-d_i)[m(u_i)-m(\bar{u})+\varepsilon_i]}{\frac{1}{n}\sum_{j=1}^{n}\frac{1}{h\lambda_j^*}k(\frac{u_j-\bar{u}}{h\lambda_j^*})(1-d_j)}. \tag{5.4}$$

Now by taking the denominator of the first term.Show$\frac{1}{n}\sum_{j=1}^{n}\frac{1}{h\lambda_j^*}k(\frac{\bar{u}-u_j}{h\lambda_j^*})d_j \xrightarrow{p} \frac{f_0(\bar{u})}{2}$

$$=\frac{1}{n}\sum_{j=1}^{n}\frac{1}{h\lambda_j^*}k(\frac{\bar{u}-u_j}{h\lambda_j^*})d_j$$

Taking variance on both sides, we have

$$var[\frac{1}{n}\sum_{j=1}^{n}\frac{1}{h\lambda_j^*}k(\frac{\bar{u}-u_j}{h\lambda_j^*})d_j] \le \frac{1}{(nh\lambda_j^*)^2}E[k^2(\frac{u_i-\hat{u}}{h\lambda_j^*})d_j] \tag{5.5}$$

Hence by using the transformation in equation (5.4), we have

$$=\frac{1}{(n)}\int_0^M \frac{1}{(h\lambda_j^*)^2}k^2(v)f_0(\bar{u}+h\lambda_j^* v)dv$$

$$=\frac{1}{(nh\lambda_j^*)}\int_0^M k^2(v)f_0(\bar{u}+h\lambda_j^* v)dv$$

By simplification, we have

$$=O(\frac{1}{nh\lambda_j^*})=o(1)$$

Then, by using Chebyshev's Inequality the general formula is $1-1/SD^2$.

$$\frac{1}{n}\sum_{j=1}^{n}\frac{1}{h\lambda_j^*}k(\frac{\bar{u}-u_j}{h\lambda_j^*})d_j = E[\frac{1}{n}\sum_{j=1}^{n}\frac{1}{h\lambda_j^*}k(\frac{\bar{u}-u_j}{h\lambda_j^*})d_j]+o_p(1)$$

By using the transformation, we have;

$$=f_0(\bar{u})\int_0^M k(v)dv+o_p(1)$$

$$=\frac{f_0(\bar{u})}{2}+o_p(1)$$

Hence, we know that

$$\frac{1}{n}\sum_{j=1}^{n}\frac{1}{h\lambda_j^*}k(\frac{\bar{u}-u_i}{h\lambda_j^*})d_j \xrightarrow{p} \frac{f_0(\bar{u})}{2}$$

Similarly, we do with second term of equation (5.3), by taking their denominator;

$$\frac{1}{n}\sum_{j=1}^{n}\frac{1}{h\lambda_j^*}k(\frac{\bar{u}-u_j}{h\lambda_j^*})(1-d_j)$$

Taking variance on both sides, we have

$$var[\frac{1}{n}\sum_{j=1}^{n}\frac{1}{h\lambda_j^*}k(\frac{\bar{u}-u_j}{h\lambda_j^*})(1-d_j)]$$

$$\le \frac{1}{(nh\lambda_j^*)^2}E[k^2(\frac{u_i-\hat{u}}{h\lambda_j^*})(1-d_j)] \tag{5.6}$$

Hence by using the transformation we have

$$=\frac{1}{(n}\int_0^M \frac{1}{(h\lambda_j^*)^2}k^2(v)f_0(\bar{u}+h\lambda_j^* v)dv$$

$$=\frac{1}{(nh\lambda_j^*)}\int_0^M k^2(v)f_0(\bar{u}+h\lambda_j^* v)dv$$

By simplification, we have

$$=O(\frac{1}{nh\lambda_j^*})=o(1)$$

Then, by using Chebyshev's Inequality the general formula is $1-1/SD^2$.

$$\frac{1}{n}\sum_{j=1}^{n}\frac{1}{h\lambda_j^*}k(\frac{u_i-\bar{u}}{h\lambda_j^*})(1-d_j)$$

$$=E[\frac{1}{n}\sum_{j=1}^{n}\frac{1}{h\lambda_j^*}k(\frac{u_i-\bar{u}}{h\lambda_j^*})(1-d_j)]+o_p(1)$$

Table 5. Adaptive Nadaraya Watson Estimator With Epanechikov As Density Function (Proposed Model)

sample size	Bandwidth	Adaptive Bandwidth	Jump Size	MSE	rate convergence
35	0.1104331	0.0186666	-0.5111896	0.856058	0.01477945

By using the transformation

$$= f_0(\bar{u})\int_0^M k(v)dv + o_p(1)$$

$$= \frac{f_0(\bar{u})}{2} + o_p(1)$$

Hence, we know that

$$\frac{1}{n}\sum_{j=1}^{n}\frac{1}{h\lambda_j^*}k(\frac{u_i-\bar{u}}{h\lambda_j^*})(1-d_j)\xrightarrow{p}\frac{f_0(\bar{u})}{2}$$

Hence, when we have large number we use Liapunov's condition;

$$\sum_i E|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_j^*}k(\frac{u_i-\bar{u}}{h\lambda_i^*})d_i\varepsilon_i|^{2+\xi}$$

$$= \frac{1}{(nh\lambda_i^*)^{\frac{\xi}{2}}}E[|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_j^*}k(\frac{u_i-\bar{u}}{h\lambda_i^*})|^{2+\xi}]dE|\varepsilon|^{2+\xi}u$$

$$\leq \frac{1}{nh\lambda_i^*}^{\frac{\xi}{2}}[sup_{u\varepsilon M}E(|\varepsilon|^{2+\xi})u]\int_0^M k(v)^{2+\xi}f_0(\bar{u}+h\lambda_i^*v)dv$$

$$= o(1)$$

By using central limit theorem (CLT) to find out the asymptotic variance;

$$\sum_j var(\frac{1}{\sqrt{nh\lambda_i^*}}\sum_{i=1}^{n}k(\frac{u_i-\bar{u}}{h\lambda_i^*})d_i\varepsilon_i) = E[\frac{1}{h\lambda_i^*}k^2(\frac{u_i-\bar{u}}{h\lambda_i^*})d\delta^2(u)]$$

$$= \int_0^M k^2(v)\delta^2(\bar{u}+h\lambda_i^*v)dv$$

Similarly, we have

$$= \delta^{2-}(\bar{u})f_0(\bar{u})\int_0^M k^2(v)dv + o(1)$$

now by using Liapunov's CLT and after simplification, we have

$$\frac{\frac{1}{\sqrt{nh\lambda_i^*}}\sum_i\sum_{i=1}^{n}k(\frac{u_i-\bar{u}}{h\lambda_i^*})d_i\varepsilon_i}{\frac{1}{nh\lambda_i^*}\sum_j k(\frac{u_j-\bar{u}}{h\lambda_j^*})d_j\varepsilon_j} - \frac{\frac{1}{\sqrt{nh\lambda_i^*}}\sum_i\sum_{i=1}^{n}k(\frac{u_i-\bar{u}}{h\lambda_i^*})(1-d_i)\varepsilon_i}{\frac{1}{nh\lambda_i^*}\sum_j k(\frac{u_j-\bar{u}}{h\lambda_j^*})(1-d_j)\varepsilon_j}$$

$$\xrightarrow{d} N(0, 4\delta^{2+}(\bar{u}) + \delta^{2-}(\bar{u})f_0(\bar{u})\int_0^M k^2(v)dv)$$

Finally, we consider the bias of the estimator;

$$var(\frac{1}{\sqrt{nh\lambda_i^*}}\sum_{i=1}^{n}k(\frac{u_i-\bar{u}}{h\lambda_i^*})d_i[m(u_i-m(\bar{u}))]$$

$$\leq \frac{1}{h\lambda_i^*}Ek^2(\frac{u_i-\bar{u}}{h\lambda_i^*})d[m(u_i-m(\bar{u}))]^2)$$

$$\leq [sup_r\varepsilon[0,M]|m(\bar{u}+rh\lambda_i^*)-m(\bar{u})|^2]\int_0^M k^2(v)f_0(\bar{u}+h\lambda_i^*v)dv$$

$$= o(1)$$

Hence, again by using Chebyshev's Inequality;

$$\frac{\frac{1}{\sqrt{nh\lambda_i^*}}\sum_i\sum_{i=1}^{n}k(\frac{u_i-\bar{u}}{h\lambda_i^*})d_i[m(u_i)-m(\bar{u})]}{\frac{1}{nh\lambda_i^*}\sum_j k(\frac{u_j-\bar{u}}{h\lambda_j^*})d_j}$$

$$-\frac{\frac{1}{\sqrt{nh\lambda_i^*}}\sum_i\sum_{i=1}^{n}k(\frac{u_i-\bar{u}}{h\lambda_i^*})(1-d_i)[m(u_i)-m(\bar{u})]}{\frac{1}{nh\lambda_i^*}\sum_j k(\frac{u_j-\bar{u}}{h\lambda_j^*})(1-d_j)} \quad (5.7)$$

By using the values which we find out for the denominator of equation (5.3), replace into the equation (5.7), we have

$$= \frac{2}{f_0(\bar{u})}\sqrt{\frac{n}{h\lambda_i^*}}(E[k(\frac{u_i-\bar{u}}{h\lambda_i^*})d[m(u_i-m(\bar{u}))]]$$

$$-E(k(\frac{u_i-\bar{u}}{h\lambda_i^*})d[m(u_i-m(\bar{u})(1-d_i)[m(u_i-m(\bar{u}))]$$

After simplification, we have

$$= \frac{2\sqrt{nh\lambda_i^*}}{f_0(\bar{u})}h\lambda_i^*(m^{'+}(\bar{u})+m^{'-}(\bar{u}))f_0(\bar{u})\int_0^M k(v)dv + o(1)$$

$$= 2\sqrt{nh\lambda_i^*}h\lambda_i^*(m^{'+}(\bar{u})+m^{'-}(\bar{u}))\int_0^M k(v)dv + o_p(1)$$

Therefore, at the end after simplification and by using different method, we proved that:

$$\sqrt{nh\lambda_i^*}(\hat{\alpha}-\alpha)\xrightarrow{d} N(C_2k_1(0)(m^{'+}(\bar{u})+m^{'-}(\bar{u}))),$$

$$4\delta_0\delta^{2+}(\bar{u}) + \delta^{2-}(\bar{u})/f_0(\bar{u})$$

Whereas,

$$p = 2h(\lambda_i^*)^{3/2}; k_1(0) = \int_0^M k(v)dv$$

Proof of Theorem 2. In order to give importance to the optimization, we have

$$\hat{q}_+(y) = \frac{1}{n}\sum_{j=1}^{n}k_{h\lambda_j^*}(y-Y_j)d_j$$

And

$$\hat{q}_-(y) = \frac{1}{n}\sum_{j=1}^{n}k_{h\lambda_j^*}(y-Y_j)(1-d_j)$$

The kernel density estimate at y is:

$$\hat{q}_0 = \hat{q}_+(y) + \hat{q}_-(y)$$

Whereas, $\hat{q}_+(y)$ is the density estimate come from data to right of the discontinuity and similarly $\hat{q}_-(y)$ is the part left of discontinuity. Whereas,$\hat{m}(y_i) = v^i_j \hat{y}_j$ is the consistent adaptive nadaraya Watson kernel estimator of $m(y_i)$.

$$\hat{y}_j = y_i + \alpha d$$

Such that q_- is l_q times continuously differentiable and bounded away from zero. And

$$v^i_j = \frac{k_{h\lambda_i^*}(y_i - y_j)}{\sum_{r=1}^n k_{h\lambda_i^*}(y_i - y_j)}$$

We have,

$$\hat{\varepsilon}_i = y_i - \hat{m}(y_i) - d_i\hat{\alpha}$$

$$\hat{m}(y) = \frac{\frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{y-y_i}{h\lambda_i^*})(y_i - d_i\hat{\alpha})}{\hat{q}_0(y)}$$

Whereas, we have

$$\delta^{2+}(\bar{y}) = \lim y_{y\to y_0^+ var(z/y_i=y)}$$

and

$$\delta^{2-}(\bar{y}) = \lim y_{y\to y_0^+ var(z/y_i=y)}$$

$\delta^2(y_i) = var(z_i/y_i)$ is uniformly bounded near y_0 and the limits $\delta^{2+}(y_0), \delta^{2+}(y_0)$ exist are finite we have;

$$\hat{s}(y) = \frac{1}{nh\lambda_i^*}\sum_i k(\frac{y-y_i}{h\lambda_i^*})\hat{x}_i$$

and

$$s(y) = m(y)q_0(y)$$

Suppose $\hat{\alpha}$ $(jump \quad size)$ is a consistent estimator for α. Then by defining their left and right hand side of variance estimation by

$$\hat{\delta^2}^+(\bar{y}) = \frac{\frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i\hat{\varepsilon_i^2}}{\frac{1}{2}\hat{q}_0(\bar{y})}$$

and

$$\hat{\delta^2}^-(\bar{y}) = \frac{\frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i)\hat{\varepsilon_i^2}}{\frac{1}{2}\hat{q}_0(\bar{y})}$$

Let we have

$$\hat{\delta^2}^+(\bar{y}) = \frac{\frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i\hat{\varepsilon_i^2}}{\frac{1}{2}\hat{q}_0(\bar{y})} \tag{5.8}$$

Hence, we consider the numerator of the equation (5.8)

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i\hat{\varepsilon}_i^2 = \frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i\hat{\varepsilon}_i^2 \\
&+\frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i[(m(y_i) \\
&-\frac{(\frac{\frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i}{})^2}{\hat{q}_0^2}(\hat{\alpha}-\alpha)^2] \\
&+2\frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i\hat{\varepsilon}_i[(m(y)_i) - \frac{\frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i\hat{y}_i}{\hat{q}_0}(y)_i] \\
&-(\frac{\frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i(1-d_i)^2}{\hat{q}_o(y_i)})(\hat{\alpha}-\alpha)] \\
&-2\frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i\frac{\frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i)}{\hat{q}_o(y_i)} \\
&(\hat{\alpha}-\alpha)(m(y_i) - \frac{\frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i\hat{y}_i}{\hat{q}_o(y_i)})
\end{aligned} \tag{5.9}$$

Divide equation (5.9) in part $(a+b+c)$ and solve one by one;

$$\frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i\hat{\varepsilon}_i^2 + \frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i[(m(y_i)$$

$$-\frac{(\frac{\frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i}{})^2}{\hat{q}_0^2}(\hat{\alpha}-\alpha)^2]$$

By using Lemmas A_1, A_2 we have

$$sup_{y\varepsilon M_0}|d\hat{q}_+(y)| \le sup_{y\varepsilon M_0}d|\hat{q}_+(y) - E\hat{q}(y)| + sup_{y\varepsilon M_0}|d\hat{q}_+(y)|$$

$$= o_p(\frac{1}{nh\lambda_i^*})$$

now by taking part we have

$$2\frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i\hat{\varepsilon}_i[(m(y)_i) - \frac{\frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i\hat{y}_i}{\hat{q}_0}(y)_i]$$

$$-(\frac{\frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i(1-d_i))^2}{\hat{q}_o(y_i)})(\hat{\alpha}-\alpha)]$$

By using Lemma A_3

$$sup_{y\varepsilon M_0}|\hat{q}_0(y) - q_0(y)| \le sup_{y\varepsilon M_0}|\hat{q}_0(y) - q_0^+(y)| + |q_0^+(y) - q_0(y)|$$

$$= o_p(\frac{1}{nh\lambda_i^*})$$

Similarly by taking part, we have

$$2\frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i\frac{\frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i)}{\hat{q}_o(y_i)}$$

$$(\hat{\alpha}-\alpha)(m(y_i) - \frac{\frac{1}{n}\sum_{i=1}^n \frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i\hat{y}_i}{\hat{q}_o(y_i)}$$

By using lemma A_3, we have

$$sup_{\bar{y}\le y\le \bar{y}+Mh\lambda_i^*}|\hat{s}(y) - s(y)| \le sup_{y\varepsilon M_0}|\hat{s}(y) - \bar{s}(y)|$$

$$+sup_{\bar{y}\le y\le \bar{y}+Mh\lambda_i^*}|\bar{s}(y) - s(y)|$$

$$= o_p(\frac{1}{nh\lambda_i^*})$$

$$sup_{\bar{y}\le y\le \bar{y}+Mh\lambda_i^*}|m(y) - \frac{\hat{s}(y)}{\hat{q}_0(y)}| \le sup_{y\varepsilon M_0}|\frac{m(y)}{\hat{q}_0(y)}|sup_{y\varepsilon M_0}|\hat{q}_0(y) - q_0(y)|$$

$$+sup_{y\varepsilon M_0|\frac{1}{\hat{q}_0(y)}|sup_{\bar{y}\le y\le \bar{y}+Mh\lambda_i^*}|\hat{s}(y)-s(y)|}$$

$$= o_p(\frac{1}{nh\lambda_i^*})$$

Whereas we have

$$var(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i\hat{\varepsilon}_i) \leq \frac{1}{(nh\lambda_i^*)^2}E[k^2(\frac{\bar{y}-y_i}{h\lambda_i^*})d\delta^2(y)]$$

By using approximation, we have

$$= \frac{1}{nh\lambda_i^*}\int_0^M k^2(v)\delta(\bar{y}+h\lambda_i^* v)f_0(\bar{y}+h\lambda_i^* v)dv$$

$$= o(\frac{1}{nh\lambda_i^*})$$

As already we have prove they asymptotic properties of the estimator, we use the asymptotic variance that is;

$$var(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i) = o(\frac{1}{nh\lambda_i^*})$$

Hence, by using the results of chebysehev's inequality for the uniform convergence, and also that $(\hat{\alpha}-\alpha) = o_p(\frac{1}{nh\lambda_i^*})$we have the last term of the equation (5.7) equals to $o_p(\frac{1}{nh\lambda_i^*})$ Hence, by applying the application of chebyseve's inequality we prove that; we have the last term of the equation (5.8) equals to $o_p(\frac{1}{nh\lambda_i^*})$ Hence, by applying the application of chebyseve's inequality we prove that;

$$\hat{\delta}^{2+}(\bar{y}) \xrightarrow{p} \delta^{2+}(\bar{y})$$

Similarly, we do for

$$\delta\hat{^{2-}}(\bar{y}) = \frac{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i)\hat{\varepsilon_i^2}}{\frac{1}{2}\hat{q}_0(\bar{y})} \tag{5.10}$$

Hence, we consider the numerator of the equation (5.10)

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i)\hat{\varepsilon}_i^2 = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i)\hat{\varepsilon}_i^2\\
&+\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i)[(m(y_i)-\\
&\frac{(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i))^2}{\hat{q}_0^2}(\hat{\alpha}-\alpha)^2]\\
&+2\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i)\hat{\varepsilon}_i[(m(y)_i)\\
&-\frac{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i)\hat{y}_i}{\hat{q}_0}(y)_i]\\
&-(\frac{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i(1-d_i)^2}{\hat{q}_o(y_i)})(\hat{\alpha}-\alpha)]\\
&-2\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i)\frac{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i)}{\hat{q}_o(y_i)}\\
&(\hat{\alpha}-\alpha)(m(y_i)-\frac{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i)\hat{y}_i}{\hat{q}_o(y_i)})
\end{aligned} \tag{5.11}$$

Divide equation (5.11) in parts and solve one by one;

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i)\hat{\varepsilon}_i^2+$$

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i)[(m(y_i)-\frac{(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i))^2}{\hat{q}_0^2}(\hat{\alpha}-\alpha)^2]$$

By solving using Lemmas A_1, A_2 we have

$$sup_{y\varepsilon M_0}|d\hat{q}_-(y)| \leq sup_{y\varepsilon M_0}d|\hat{q}_-(y)-E\hat{q}(y)| + sup_{y\varepsilon M_0}|d\hat{q}_-(y)|$$

$$= o_p(\frac{1}{nh\lambda_i^*})$$

Know by taking part we have

$$2\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i)\hat{\varepsilon}_i[(m(y)_i)$$

$$-\frac{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i)\hat{y}_i}{\hat{q}_0}(y)_i]$$

$$-(\frac{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})d_i(1-d_i)^2}{\hat{q}_o(y_i)})(\hat{\alpha}-\alpha)]$$

By using Lemma A_3

$$sup_{y\varepsilon M_0}|\hat{q}_0(y)-q_0(y)| \leq sup_{y\varepsilon M_0}|\hat{q}_0(y)-q_0^-(y)| + |q_0^-(y)-q_0(y)|$$

$$= o_p(\frac{1}{nh\lambda_i^*})$$

Similarly by taking part, we have

$$2\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i)\frac{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i)}{\hat{q}_o(y_i))}(\hat{\alpha}-\alpha)$$

$$(m(y_i)-\frac{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i)\hat{y}_i}{\hat{q}_o(y_i)}$$

using lemma A_3, we have

$$\begin{aligned}
sup_{\bar{y}\leq y\leq \bar{y}+Mh\lambda_i^*}|\hat{s}(y)-s(y)| &\leq sup_{y\varepsilon M_0}|\hat{s}(y)-\bar{s}(y)|\\
&+sup_{\bar{y}\leq y\leq \bar{y}+Mh\lambda_i^*}|\bar{s}(y)-s(y)|
\end{aligned}$$

$$= o_p(\frac{1}{nh\lambda_i^*})$$

$$sup_{\bar{y}\leq y\leq \bar{y}+Mh\lambda_i^*}|m(y)-\frac{\hat{s}(y)}{\hat{q}_0(y)}| \leq sup_{y\varepsilon M_0}|\frac{m(y)}{\hat{q}_0(y)}|sup_{y\varepsilon M_0}|\hat{q}_0(y)-q_0(y)|$$

$$+sup_{y\varepsilon M_0|\frac{1}{\hat{q}_0(y)}|sup_{\bar{y}\leq y\leq \bar{y}+Mh\lambda_i^*}|\hat{s}(y)-s(y)|}$$

$$= o_p(\frac{1}{nh\lambda_i^*})$$

Whereas we have

$$var(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i)\hat{\varepsilon}_i) \leq \frac{1}{(nh\lambda_i^*)^2}E[k^2(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i\delta^2(y)]$$

By using approximation, we have

$$= \frac{1}{nh\lambda_i^*}\int_0^M k^2(v)\delta(\bar{y}+h\lambda_i^* v)f_0(\bar{y}+h\lambda_i^* v)dv$$

$$= o(\frac{1}{nh\lambda_i^*})$$

As already we have prove they asymptotic properties of the estimator, we use the asymptotic variance that is;

$$var(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i^*}k(\frac{\bar{y}-y_i}{h\lambda_i^*})(1-d_i)) = o(\frac{1}{nh\lambda_i^*})$$

Hence, by using the results of chebysehev's inequality for the uniform convergence, and also that

$$(\hat{\alpha}-\alpha) = o_p(\frac{1}{nh\lambda_i^*})$$

we have the last term of the equation (5.11) equals to $o_p(\frac{1}{nh\lambda_i^*})$ Hence, by applying the application of chebyseve's inequality we prove that; we have the last term of the equation (5.11) equals to $o_p(\frac{1}{nh\lambda_i^*})$ Hence, by applying the application of chebyseve's inequality we prove that;

$$\hat{\delta}^{2-}(\bar{y}) \xrightarrow{p} \delta^{2-}(\bar{y})$$

This result requires a more stringent moment for the limiting distribution to estimate . variance estimator and the above consistency theorem don't required continuity in the derivatives of m at the discontinuity.