Performance of Bandwidth Selection Rules for the Local Linear Regression

Kang-Sun Lee[1] and R. Carter Hill

Department of Economics, Louisiana State University, Baton Rouge, LA 70803, U.S.A.

Summary

The choice of bandwidth is crucial in the nonparametric estimation procedure. A number of methods to choose the associated bandwidth have been developed. In this paper we studied three existing bandwidth selectors for local linear regression with different design matrix characteristics. The performances illustrate that although there is no uniformly dominating rule, the variable bandwidth selector is superior to the other bandwidth selectors in highly skewed data or when the complicated functional form is.

Keywords: Rule-of-Thumb bandwidth, Least-squares cross validation bandwidth, Variable bandwidth selectors, Local linear regression estimate, Relative efficiency, Epanechnikov kernel

1. Introduction

Local linear regression estimation uses a random sample $\{x_i, y_i\}$ $i = 1, \ldots, n$ to estimate the curve $\hat{y}(x) = \hat{m}(x)$ by minimizing

$$(1.1)\qquad \sum_{i=1}^{n}\{y_i - \beta_0 - \beta_1(x_i - x)\}^2 K_h(x_i - x),$$

where $K_h(x_i - x) = K[(x_i - x)/h]$, K is called the kernel function and h is called the bandwidth. If the regression function $m(x)$ is approximated locally by a linear Taylor's expansion in a neighborhood of x, then the local linear regression estimator performs a weighted regression of y_i against $z'_i = (1, (x_i - x))$ using weights $w_i^{1/2} = \{K[(x_i - x)/h]\}^{1/2}$. The local linear regression estimator is obtained by fitting local straight lines. An interesting collection of effective data analysis carried out by this simple and intuitive estimator is given in Fan and Gijbels (1996). Like every kernel-type estimator, the bandwidth selection in the local linear regression estimation is important. When h is too small, the resulting curve is too wiggly,

[1] Corresponding author, e-mail: kslee@lsu.edu

reflecting too much of the sampling variability. When h is too large, the resulting estimate tends to smooth away important features. For this reason, data-driven choice of h has been a key issue of the kernel type nonparametric estimation. The general criterion of the bandwidth selection is Mean Integrated Squared Error defined by;

$$MISE = \int MSE(x)dx ,$$

where

$$MSE = E[\{\hat{m}(x) - m(x)\}^2 \mid x] \text{ and } x = (x_1, \ldots, x_n) .$$

Seather (1992) and Park and Turlach (1992) compared several constant bandwidth selectors using simulated and real data sets for density estimation, separately. They found that *plug-in* methods performed well when the data has the several modes as well as one-mode and usually least squares *cross-validation* undersmoothed. But when the data has the skewed and long tail, none of them fit the data well, since a global bandwidth fixed across the entire range of the data is not at all suited. They said that there is no best bandwidth selector that works in all cases. Although 'plug-in' estimators of h work well in the situation with density estimation, this 'plug-in' estimator does not been a great deal of merit for the conditional moment estimation (Pagan & Ullah (1999), and M.J. Lee (1996)). A procedure that responds to this observation is variable bandwidth estimation.

A variable bandwidth is introduced to allow for different degrees of smoothing by Brieman, Meisel and Prucell (1977), resulting in a possible reduction of estimation bias at peaked regions and a reduction of the variance at flat regions. This enhances the flexibility of the local polynomial fitting, so that it can adapt to spatially non-homogeneous curves. Fan and Gijbels (1992, 1995) used the variable bandwidth for the local linear smoothers and they argued that the variable bandwidth has theoretical advantages. Zhang and Lee (2000) showed that the Mean Integrated Squared Error (MISE) of variable bandwidth is much smaller than the cross-validation method and the theoretical optimal constant bandwidth. Lee and Solo (1999) studied bandwidth selection for the local linear regression with constant bandwidth selectors. Although they suggested the two new simple selectors, the least-square cross-validation performed better than other selectors generally.

The one thing we consider here is that the empirical performance has been judged using only the uniform or normal distributions for covariates. Sheather (1992) and Park and Turlach (1992) used the mixture of normal densities for kernel density estimation, Fan and Gijbels (1992, 1995) and Zhang and Lee (2000) used a normal or an uniform distribution for the covariates. Such choices do not represent all real situations formed with real data. For example, in the labor market the worker's experience has the log-normal distribution, i.e. highly skewed to the right-hand side.

Our goal of this chapter is the comparison of some bandwidth selection rules using different distributions of covariates. The selected bandwidth methods are rule-of-thumb method, least squares cross-validation constant bandwidth and variable bandwidth estimator. This chapter is organized as follows. In the next section, we briefly introduced three bandwidth methods. The simulation study is provided in section 3 and section 4 will gives some summary.

2. Some Bandwidth Selection Rules

2.1. Rule-of-Thumb Bandwidth

In many data analyses, one would like to get a quick idea about how large the amount of smoothing should be. A "rule-of-thumb (ROT)" bandwidth selection is very suitable in such a case. Such a rule is meant to be somewhat crude, but possesses simplicity and requires little programming effort that other methods are hard to compete with. Pudney (1993) and Ginther (1999) used the ROT bandwidth selection method for their empirical study.

With the local polynomial regression method such a crude bandwidth selector can easily be obtained as follows. Consider the asymptotically optimal constant bandwidth, which come from minimizes the asymptotic weighted Mean Integrated square error (WMISE)

$$(2.1)\qquad \begin{aligned} WMISE &= E[\int (\hat{m}(x)-m(x))^2 w(x)dx] \\ &= \int \left([Bias\{\hat{m}(x)\}]^2 + Var\{\hat{m}(x)\}\right) w(x)dx \end{aligned}$$

where

$$Bias\{\hat{m}(x_0)\} = \{\int x^2 K(x)dx\}\frac{1}{(2)!}m''h + o(h)$$

$$Var\{\hat{m}(x_0)\} = \int K^2(x)dx\frac{\sigma^2(x_0)}{f(x_0)nh} + o(\frac{1}{nh})$$

with $w \geq 0$ some weight function, leads to a theoretical optimal constant bandwidth. Using the asymptotic expression of conditional bias and variance of local linear regression estimator, an asymptotically optimal constant bandwidth is

$$(2.2)\qquad h_{opt} = n^{-1/5}C(K)\left[\frac{\int \sigma^2(x)w(x)/f(x)dx}{\int \{m''(x)\}^2 w(x)dx}\right],$$

where $C(K)$ is some constant values, m'' is the second derivative function estimation, and $f(x)$ is density function of x. Fan and Gijbels (1995) give $C(K)$ = 2.719 when the function $m(\cdot)$ itself is estimated with local linear regression. It contains the unknown quantities $\sigma^2(\cdot)$, $m''(\cdot)$ and $f(\cdot)$, which need to be estimated. The "Rule of Thumb" bandwidth fits a polynomial of order 4 globally to $m(x)$ by the parametric fit

$$(2.3)\qquad \hat{m}(x) = \hat{\alpha}_0 + \ldots + \hat{\alpha}_4 x^4 .$$

The standardized residual sum of squares from this parametric fit is denoted by $\hat{\sigma}^2$. Substitute the estimated value for the equation (2.2), and then we can obtain the rule of thumb bandwidth selector

$$(2.4)\qquad \hat{h}_{ROT} = C(K)\left[\frac{\hat{\sigma}^2\int w(x)dx}{\sum_{i=1}^{n}\{\hat{m}''(x_i)\}^2 w(x_i)}\right]^{1/5}.$$

2.2. Cross-Validation Bandwidth Selection Rule

The most widely studied bandwidth selector is least squares cross-validation (LSCV), proposed by Rudemo (1982) and Bowman (1984). There are some of applications of this method; e.g. Stock (1989), McMillen and Thorsnes (1999), Iwata et al. (1999), and Zheng (1999). The basic idea behind this cross-validation (CV) procedure is to choose h by minimizing the Integrated Squared Error (ISE) defined by $ISE = \int\{\hat{m}(x)-m(x)\}^2 dx$. Let $\hat{m}_h(\cdot)$ denote any estimate, involving a smoothing parameter h, of the regression function $m(\cdot)$. For each given i, we use data $\{(x_j, y_j), j \neq i\}$ to build a regression function $\hat{m}_{h,-i}(\cdot)$ and then validate the model by examining the prediction error $y_i - \hat{m}_{h,-i}(x_i)$. The least squares cross-validation technique uses the weighted average of squared errors

$$(2.5)\qquad CV(h) = n^{-1}\sum_{i=1}^{n}\{y_i - \hat{m}_{h,-i}(x_i)\}^2 w(x_i),$$

as an overall measure of the effectiveness of the estimation scheme $\hat{m}_{h,-i}(\cdot)$ where $w(x_i)$ is some positive function. From (2.5), the expression $\hat{m}_{h,-i}(x_i)$ is the 'leave-one-out' estimator of (1.1) omitting the i^{th} observation. The least squares cross-validation bandwidth selector is the one that minimizes (2.5). The method to find the minimum of (2.5) is the grid search method. Find the all of CV for the grid sets of *h* values. The least squares cross-validation bandwidth is

$$\hat{h}_{cvls} = \arg\min_{h}[CV(h)]\,. \tag{2.6}$$

2.3. Variable Bandwidth Selection Rule

The concept of the variable bandwidth was introduced by Breiman, Meisel and Prucell (1977) in the density estimation context. Instead of (1.1), the local linear regression estimator is obtained by minimizing

$$\sum_{i=1}^{n}\{y_i - \beta_0 - \beta_1(x_i - x)\}^2 \alpha(x_i) K_h\left(\frac{x_i - x}{h}\alpha(x_i)\right), \tag{2.7}$$

with respect to β_0 and β_1, where $\alpha(\cdot)$ is some nonnegative function reflecting the variable amount of smoothing at each data point. The optimal variable bandwidth is the same method that minimizes WMISE with respect to *h*, except that the variable bandwidth has the varying term $\alpha(\cdot)$ to be chosen. Fan and Gijbels (1992) suggested that an optimal choice of $\alpha(\cdot)$ is proportional to $f_x^{1/5}(\cdot)$, where f_x is marginal distribution of *x*, and this is precisely how an ideal variable kernel smoother should behave. The optimal variable bandwidth is defined by

$$\hat{h}_v = h_{opt} / \alpha(x_i) = h_{opt} / f_x^{1/5}, \tag{2.8}$$

where h_{opt} is the optimal constant bandwidth, and f_x is the marginal density function of *x*.

3. A Simulation Study

3.1. Random Number Generating Method and Simulating Function

A simulation study is conducted to evaluate the practical performance of the proposed bandwidth schemes; Rule-of-Thumb bandwidth (ROT), Cross-Validation bandwidth and Variable bandwidth. Four test functions are used:

1: $m(x) = 0.4x + 1$ $\qquad x \in [0,1]$,

2: $m(x) = 0.3 + 4x - 3x^2$ $\qquad x \in [0,1]$,

3: $m(x) = x + 2\exp(-16x^2)$ $\qquad x \in [0,1]$,

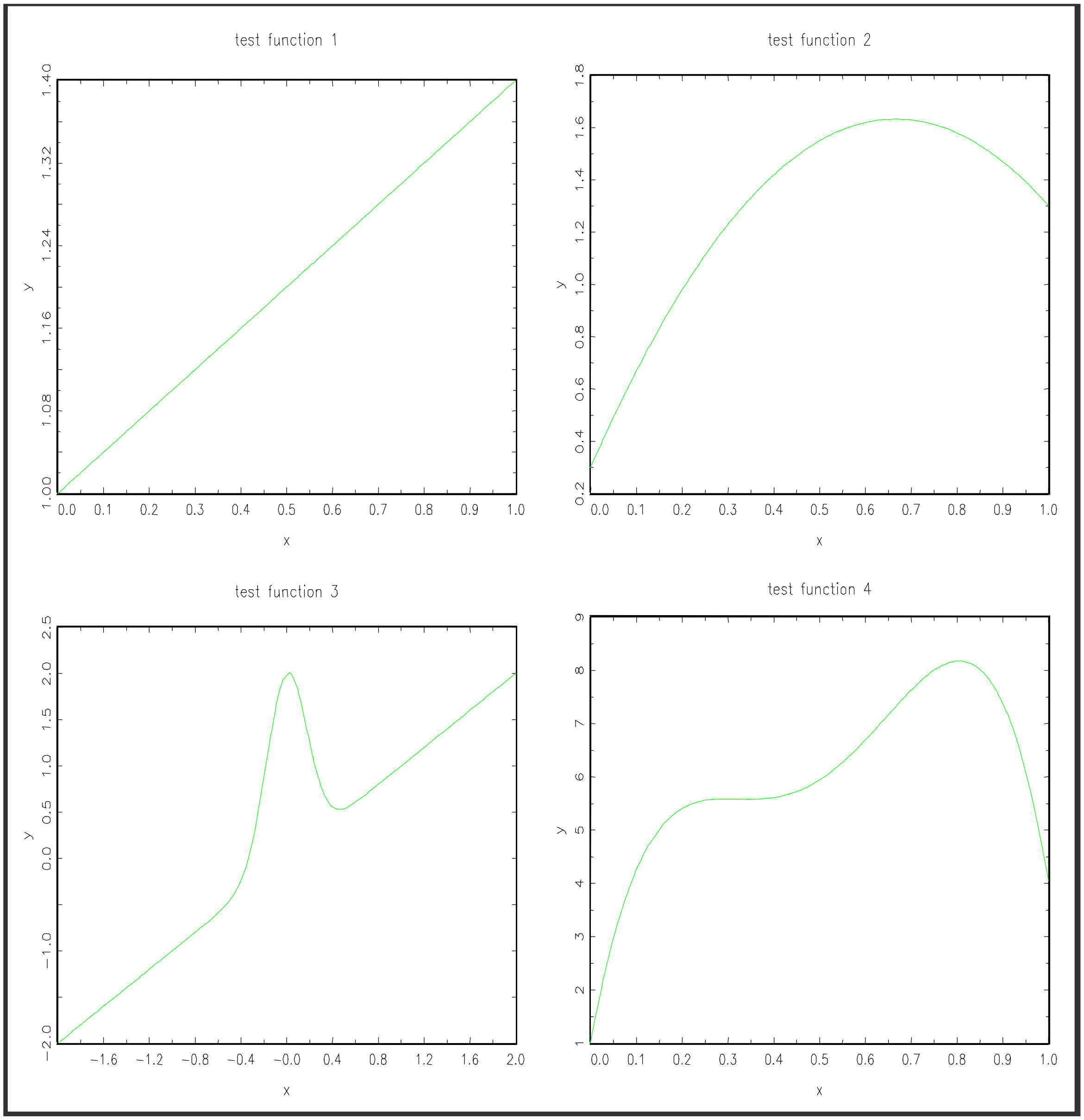
4: $m(x) = 1 + 48x - 180x^2 = 145x^4$ $\qquad x \in [0,1]$,

Let *x* and *y* be the two random variables whose relationship can be modeled as

$$y = m(x) + \sigma(x)\varepsilon \qquad E(\varepsilon) = 0, \quad \operatorname{var}(\varepsilon) = 1, \tag{2.9}$$

where *x* and ε are independent.

Figure 1 Plots of Test functions



The test functions 1 and 3 are used by Fan and Gijbels (1995), their covariates are generated from a normal distribution for test function 1 and from a uniform distribution for test function 3. The test functions 2 and 4 are quadratic and quartic functions that are chosen arbitrarily. The reason of choosing of the test functions 2 and 4 is that the quadratic and quartic functions are often used in econometric modeling, for example, the estimation of wage equation in labor economics. These four test functions are plotted in Figure 1.

Three signal-to-noise ratios (s/n) and three design densities were used. Here signal-to-noise is defined to be the variance of the function divided by the variance of the noise: $s/n = \mathrm{var}(m)/\sigma^2$. The three s/n were: low=2, medium=4 and high=8, and the three design densities were the uniform density, normal density, and gamma density. Normal random errors were used for all test function. For each test function, the distribution of error terms follow: $\sigma_1 = 0.15$, $\sigma_3 = 0.75$, $\sigma_4 = 0.25(\max m - \min m)$, and

$\sigma_2 = 0.3$, where subscription denotes each of the test functions. We use sample sizes n = 200, and 400 and number of replications in the simulation is 1000. This formatting is similar to the previous researchers' setup. In each of the examples we use the Epanechnikov kernel $K(u) = 0.75(|1-u^2|_+)$.

For the least square cross-validation (LSCV) procedure, the estimated curves are evaluated in grid points x_j, $j = 1, \ldots, n_{grid}$. So the integral involved in the methodology are implemented as averages over appropriate grid points. The grid points are used in arithmetic type, i.e. $h_i = C * h_{\min}$, where $h_{\min}$ denotes the first grid point and C is the grid span. We start from $h = h_{\min}$, keeping h by factor C and compute $M(h)$ at these geometric grid points. We stop when the function values $M(h)$ increase consecutively a certain number of times or when $h > h_{\max}$. Then we choose the minimizer of $M(h)$ as the grid point having the smallest computed $M(h)$ value. In our implementation we took $h_{\min} = (x_{(n)} - x_{(1)})/n$, $h_{\max} = (x_{(n)} - x_{(1)})/2$, and C = .1 where $x_{(1)} = \min_i x_i$, and $x_{(n)} = \max_i x_i$. For the variable bandwidth, we use the LSCV bandwidth for the pilot bandwidth, h_{opt} in equation (3.9), and the density function f_x based on this pilot bandwidth.

3.2. Results of the Simulation Study

We conduct a simulation study to evaluate and compare each of the bandwidth selectors. Tables below show WMISE of variable bandwidth and relative of efficiency of estimator of LSCV and ROT bandwidth selector for each test function for the different distribution covariates. The ratio of efficiency is computed similar to Fan's method (Fan, 1992):

$$rf_1 = \left(\frac{WMISE \text{ of the estimator with LSCV bandwidth}}{WMISE \text{ of the estimator with Variable bandwidth}} \right),$$

and

$$rf_2 = \left(\frac{WMISE \text{ of the estimator with ROT bandwidth}}{WMISE \text{ of the estimator with Variable bandwidth}} \right).$$

The weighted mean integrated squared error is defined in our simulation by:

$$WMISE = E\left[\int (\hat{m}(x) - m(x))^2 w(x) dx\right],$$

where $w(x)$ is sample density function from each bandwidth selectors. To get *WMISE*, we calculate weighted mean squared error, $E\left[(\hat{m}(x) - m(x))^2 w(x)\right]$, for each iteration, and sum them by number of iteration.

3.2.1 Uniform Random Design

Table 1 represents the uniform density design for three signal-to-noise ratios (s/n), low = 2, medium = 4 and high = 8. For the test function 1, the variable bandwidth estimator slightly dominates the LSCV bandwidth estimator in the efficiency respect. The relative efficiency ratios of LSCV (rf_1) are not significantly different from one. Also, variable bandwidth estimator dominates the ROT bandwidth estimator with an exception. ROT bandwidth estimator is significantly efficient than variable bandwidth estimator, when n = 400 and s/n = 4.

Table 1. *Relative efficiency ratios* of Bandwidth selection for the Uniform design

B/W	s/n	WMISE of Variable bandwidth		Relative Efficiency of LSCV (rf_1)		Relative Efficiency of ROT (rf_2)	
		N=200	N=400	N=200	N=400	N=200	N=400
Fn 1	2	5.86E-05	4.21E-04	1.22 (0.85)	1.38 (2.04)	1.48 (1.65)	2.71 (7.98)
	4	2.92E-05	1.33E-04	1.17 (0.85)	0.87 (-0.71)	1.54 (2.43)	0.66 (-2.26)
	8	1.77E-05	4.70E-05	1.35 (1.05)	1.32 (1.82)	5.16 (5.38)	2.98 (7.72)
Fn 2	2	0.01	0.04	1.49 (2.16)	2.29 (8.81)	1.46 (2.07)	4.10 (21.16)
	4	3.17E-03	0.04	1.15 (0.55)	2.34 (8.75)	1.16 (0.59)	3.11 (12.85)
	8	3.64E-03	0.02	2.10 (3.29)	2.75 (9.70)	1.58 (1.75)	1.32 (1.78)
Fn 3	2	0.03	0.06	0.90 (-0.31)	0.96 (-0.29)	476.70 (3.74)	1686.38 (6.74)
	4	0.01	0.03	0.91 (-0.33)	0.92 (-0.42)	1809.98 (4.02)	3141.65 (7.54)
	8	0.01	0.02	0.90 (-0.26)	0.93 (-0.52)	253.00 (3.49)	4122.80 (7.23)
Fn 4	2	1.40	0.26	1.46 (1.98)	1.23 (1.29)	0.96 (-0.12)	139.06 (787.45)
	4	0.80	13.71	1.31 (0.90)	1.19 (4.80)	0.87 (-0.39)	1.19 (0.83)
	8	0.63	7.64	1.92 (2.69)	2.38 (5.00)	0.70 (-0.88)	0.57 (-1.55)

NOTE: The numbers of parenthesis are t-value of $H_0 : rf_i = 1, t = (rf_i - 1) / s.e.(rf_i)$, where $s.e(rf_i) = \sqrt{\text{var}(WMISE_{\text{var}}) + \text{var}(WMISE_j)}$, with *i*=1,2 and *j* = LSCV and ROT.

For the test function 2, the relative efficiencies of LSCV and ROT bandwidth estimators for the variable bandwidth estimator range from about 1.15 to 2.75, and about 1.16 to 4.10, respectively. The variable bandwidth estimator strictly dominates the LSCV and ROT bandwidth estimators for test function 2 in the uniform design.

For the test function 3, variable bandwidth estimator and LSCV bandwidth estimator dose not different statistically, although LSCV bandwidth estimator dominates variable bandwidth estimator. ROT bandwidth estimator has large number of relative efficiency ratio to variable bandwidth (rf_2) for all designs. ROT bandwidth estimator is not a good estimator for the 'humped shape' functional function, since the ROT bandwidth estimator fits usually over-smooth for the 'humped' part.

For the test function 4, variable bandwidth estimator is more efficient than LSCV bandwidth estimator. The relative efficiency gain of variable bandwidth estimator is statistically significant when s/n ratio is larger. For *n* = 400 and s/n = 2 in our simulation, ROT bandwidth estimator has the larger WMISE than variable and LSCV bandwidth estimators. It comes from a situation that variable and LSCV bandwidth estimators have usually under-smoothed fits but there do not much wiggly form for small signal-to-noise ratio and ROT bandwidth estimator depends on variance of error term $\sigma_4 = 0.25(\max m - \min m)$ which represents wider bandwidth than other bandwidth resulting large variance fitting. The relative efficiency of ROT bandwidth estimator for variable bandwidth estimator is increased when signal-to-noise ratio (s/n) is increasing but there are not significantly different.

For the uniform design, we do not have a uniformly dominating bandwidth selection rule. Variable bandwidth estimator has more efficiency gain in the high signal-to-noise ratio design (s/n=8) than low signal-to-noise ratio (s/n=2) and larger number of observation in most cases of our simulation.

Table 2. *Relative efficiency ratios* of Bandwidth selection for the Normal design

B/W	S/n	WMISE of Variable bandwidth		Relative Efficiency of LSCV (rf_1)		Relative Efficiency of ROT (rf_2)	
		N=200	N=400	N=200	N=400	N=200	N=400
Fn 1	2	6.31E-05	1.68E-05	1.09 (0.30)	1.03 (0.35)	1.42 (1.64)	1.03 (0.37)
	4	5.01E-05	9.79E-05	1.14 (0.44)	1.03 (0.27)	3.66 (12.26)	1.28 (4.90)
	8	2.56E-05	8.56E-05	1.24 (0.60)	1.10 (0.51)	1.81 (1.85)	2.55 (5.00)
Fn 2	2	3.19E-04	0.03	0.73 (-1.38)	2.43 (6.65)	0.74 (-1.35)	2.93 (8.98)
	4	0.01	0.04	1.56 (1.80)	3.68 (13.36)	1.84 (2.70)	2.73 (8.61)
	8	2.38E-04	0.02	1.16 (0.40)	2.99 (12.06)	1.01 (0.02)	1.06 (0.35)
Fn 3	2	0.02	0.04	1.30 (0.64)	1.38 (1.77)	1961.44 (4.32)	6119.27 (9.63)
	4	0.15	0.02	2.37 (1.49)	1.21 (1.23)	52.66 (4.24)	14500.97 (8.59)
	8	0.09	1.87	2.99 (2.88)	3.32 (4.72)	150.75 (4.90)	97.56 (9.16)
Fn 4	2	1.03	0.26	1.33 (1.10)	1.25 (1.15)	1.00 (0.01)	29.08 (128.54)
	4	2.06	11.17	1.50 (1.15)	2.68 (7.78)	1.07 (0.17)	0.86 (-0.63)
	8	0.47	7.75	1.15 (0.48)	2.50 (5.87)	0.88 (-0.38)	0.43 (-2.23)

NOTE: The numbers of parenthesis are t-value of $H_0 : rf_i = 1$, $t = (rf_i - 1) / s.e.(rf_i)$, where $s.e(rf_i) = \sqrt{\mathrm{var}(WMISE_{\mathrm{var}}) + \mathrm{var}(WMISE_j)}$, with i=1,2 and j = LSCV and ROT.

3.2.2. Normal Random Design

Table 2 shows the results of each bandwidth selection for the normal density design. For the test function 1, the relative efficiency ratio of LSCV bandwidth estimator for variable bandwidth (rf_1) is near 1 and there are not significantly different. The variable bandwidth estimator weakly dominates ROT bandwidth estimator. The relative efficiency of variable bandwidth estimator is increased when n is increasing and s/n is larger. For the test function 2, variable bandwidth estimator is not significantly different from LSCV bandwidth estimator when n = 200. For larger observation, the variable bandwidth estimator has the large relative efficiency compare to the constant bandwidth estimator.

For test function 3, the variable bandwidth estimator has the relative efficiency for the LSCV bandwidth estimator and it is statistically significant when s/n = 8. ROT bandwidth estimator acts the same as in the uniform design.

For the test function 4, variable bandwidth estimator and LSCV bandwidth estimator are not different statistically when n = 200. When n = 400, the relative efficiency of variable bandwidth estimator for LSCV bandwidth estimator is statistically significant as s/n is large. The relative efficiency of ROT is increasing when s/n ratio is larger. It represents that variable and LSCV bandwidth estimators have large biased in more widely scattered data design since they have smaller bandwidth than ROT bandwidth estimator relatively and this small bandwidth gives wiggly fitting for the both tail parts.

Like the uniform density design, there is no uniformly dominating bandwidth selector for the normal density design. The variable bandwidth estimator is more relative efficient when s/n is larger than LSCV bandwidth estimator.

From the results of the uniform and normal density designs, there is no absolute dominating bandwidth selector. A different efficiency selector is selected for the different situations. However, for those two density designs, variable bandwidth estimator performs well.

Table 3. *Relative efficiency ratios* of Bandwidth selection for the Gamma design

B/W	s/n	WMISE of Variable bandwidth		Relative Efficiency of LSCV (rf_1)		Relative Efficiency of ROT (rf_2)	
		N=200	N=400	N=200	N=400	N=200	N=400
Fn 1	2	2.29E-05	5.54E-05	0.97 (-0.15)	0.93 (-0.42)	0.84 (-0.82)	0.29 (-5.71)
	4	9.92E-05	1.14E-04	1.31 (1.37)	0.53 (-6.77)	2.01 (3.75)	0.47 (-8.69)
	8	4.35E-05	3.79E-05	0.57 (-1.21)	1.74 (2.84)	1.07 (1.37)	1.10 (0.64)
Fn 2	2	0.07	0.02	1.05 (0.19)	0.95 (-1.14)	1.05 (0.68)	0.75 (-1.14)
	4	0.02	0.04	1.21 (1.13)	0.86 (-1.19)	3.26 (4.38)	0.84 (-1.31)
	8	0.01	0.01	0.35 (-2.02)	0.82 (-1.59)	1.11 (1.13)	1.09 (0.82)
Fn 3	2	0.03	0.07	1.49 (1.36)	1.66 (3.26)	285.42 (2.96)	512.71 (6.32)
	4	0.02	0.03	1.52 (1.43)	1.42 (2.38)	670.94 (5.05)	860.57 (7.47)
	8	0.01	0.01	1.42 (1.28)	1.65 (3.58)	808.94 (4.31)	3197.55 (8.08)
Fn 4	2	22.43	8.24	1.17 (0.57)	2.01 (2.66)	1.17 (0.58)	2.01 (2.74)
	4	5.07	16.56	1.43 (0.68)	1.52 (2.56)	1.10 (0.69)	1.09 (0.44)
	8	2.22	1.08	0.19 (-2.91)	2.66 (8.89)	0.92 (-0.28)	0.98 (-0.12)

NOTE: The numbers of parenthesis are t-value of $H_0 : rf_i = 1, t = (rf_i - 1) / s.e.(rf_i)$, where $s.e(rf_i) = \sqrt{\mathrm{var}(WMISE_{\mathrm{var}}) + \mathrm{var}(WMISE_j)}$, with i=1,2 and j = LSCV and ROT.

3.2.3. Gamma Random Design

Table 3 shows the results of WMISE of variable bandwidth estimator and relative efficiency ratios when the covariates are generated from skewed distribution function. For the test functions 1 and 2, there is no dominating bandwidth selection rule. In our simulation, the constant bandwidth estimator has the greatest relative efficiency in several cases. However, variable bandwidth estimator has improved relative efficiency when s/n is larger, and variable and constant bandwidth estimators are not significantly different in efficiency respect in most of cases. For small s/n ratio, the reason that the relative efficiency gain of ROT bandwidth estimator is large is that variable and LSCV bandwidth estimators undersmooth for the skewed tail part.

For the test function 3, variable bandwidth estimator is strictly dominate constant bandwidth estimator, although it is not significantly different from LSCV bandwidth estimator when n = 200. ROT bandwidth estimator fits over-smoothly as in uniform and normal designs.

For the test function 4, the efficiency gain of variable bandwidth for the LSCV bandwidth estimator is much larger when s/n is larger and n = 400. For n =200 and larger s/n ratio, the variable bandwidth is less efficient than constant bandwidth estimator, which means that variable bandwidth estimator has too under-smoothing (too wiggly) estimator.

For all test function, relative efficiency ratios are larger when the signal-to-noise ratio is higher and n =400. Also, like other designs, the efficiency gain of ROT bandwidth estimator is increasing when s/n ratio is larger.

4. Summary

In this paper we surveyed three existing bandwidth selectors for local linear regression with different density design. All selectors were empirically assessed by means of a simulation study. Numerical results demonstrate that the variable bandwidth estimator compare favorably to selectors.

Numerical results suggest that the LSCV selector performed well in the simple functional form and uniform or normal density design. This observation agrees with the study reported in Lee and Solo (1999).

There is no bandwidth selector performed uniformly the best in our uniform and normal design simulation. There are a few important empirical results:

1. LSCV bandwidth selector superior to ROT bandwidth selector in most designs.

2. ROT bandwidth estimator fits over-smoothly for the ‘humped’ part over all cases.

3. The constant bandwidth estimator seems enough to use the simple linear functional form regardless any random design.

4. The more complicated functional form, variable bandwidth estimator performs better than other bandwidth selector in our simulation.

5. For more observation, the relative efficiency gain of variable bandwidth estimator increases in the skewed data.

The variable bandwidth selector performs well in almost everywhere in our simulation with some exceptions. When the data are highly skewed or the functional form is very complicated in large data set, the variable bandwidth selector is superior to the other bandwidth selectors.

References

Biwman, A. (1984), An alternative method of cross-validation for the smoothing of density estimates, *Biometrika*, 71, 353-360.

Breiman, L., Meise, W. and E. Purcell (1977), Variable kernel estimates of multivariate densities. *Technometrics*, 19, 135-144.

Fan, J. (1992), Design-adaptive Nonparametric regression, *Journal of the American Statistical Association*, 87, 998-1004.

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

__________________(1995), Data-driven bandwidth selection in local polynomial fitting: Variable bandwidth and spatial adaptation, *Journal of the American Statistical Association*, 86, 643-652.

__________________ (1996), *Local Polynomial Modeling and Its Applications*, Chapman and Hall, London.

Ginther, D. K. (1999), A nonparametric analysis of the U.S. male earnings distribution, in *Applying Kernel and Nonparametric Estimation to Economic Topics*, edited by Fomby, T. B. and R. C. Hill, 275-308.

Iwata, S. H. Murao, and Q. Wang (1999), Nonparametric assessment of the effects of neighborhood land uses on residential house values, in *Applying Kernel and Nonparametric Estimation to Economic Topics*, edited by Fomby, T. B. and R. C. Hill, 201-228.

Horowitz, J. L. (1997), Bootstrap methods in econometrics: Theory and numerical Performance, in *Advances in Economics and Econometrics : Theory and Applications, Vol. III.*

Lee, Myoung-Jae (1996), *Methods of Moments and Semiparametric Econometrics for Limited Dependent Variable Models*, Springer-Verlag, New York.

Lee, T. C.M. and V. Solo (1999), Bandwidth selection for local linear regression: A simulation study, *Computational Statistics*, 14, 515-532.

McMillen, D.P. and P. Thorsnes (1999), The reaction of housing prices to information on superfund sites: A semiparametric analysis of the Tacoma, Washington Market, in *Applying Kernel and Nonparametric Estimation to Economic Topics*, edited by Fomby, T. B. and R. C. Hill, 229-258.

Park, B.U. and B.A. Turlach (1992), Practical Performance of Several Data Driven Bandwidth Selectors, *Computational Statistics*, 7, 251-270.

Pagan, A. and A. Ullah (1999), *Nonparametric Econometrics*, Cambridge University Press, London.

Pudney, S. (1993), Income and wealth inequality and the life cycle: A nonparametric analsysi for China, *Journal of Applied Econometrics*, 8, 249-276.

Rudemo, M. (1982), Empirical choice of histograms and kernel density estimators, *Scandinavian Journal of Statistics*, 9, 65-78.

Ruppert, D., S.J. Sheather, and M. P. Wand (1995), An effective bandwidth selector for local least squares regression, *Journal of the American Statistical Association*, 90, 1257-1270.

Seather, S. J. (1992), The Performance of Six Popular Bandwidth Selection Methods on Some Real Data Sets, *Computational Statistics*, 7, 225-250.

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

Stock, James H. (1989), Nonparametric Policy Analysis, *Journal of the American Statistical Association*, 84, 567-575.

Wand, M. P. and M. C. Jones (1995), *Kernel Smoothing*, Chapman and Hall, London.

Zhang, W. and S. Y. Lee (2000), Variable bandwidth selection in varying-coefficient Models, *Journal of Multivariate Analysis*, 74, 116-134.