

# Bootstrap Bandwidth Selection Methods for Local Linear Jump Detector

Dongryeon Park<sup>1,a</sup>

<sup>a</sup>Department of Applied Statistics, Hanshin University

---

## Abstract

Local linear jump detection in a discontinuous regression function involves the choice of the bandwidth and the performance of a local linear jump detector depends heavily on the choice of the bandwidth. However, little attention has been paid to this important issue. In this paper we propose two fully data adaptive bandwidth selection methods for a local linear jump detector. The performance of the proposed methods are investigated through a simulation study.

Keywords: Bandwidth for jump detector, bootstrap, local linear jump detector.

---

## 1. Introduction

Suppose we want to estimate a discontinuous regression function  $m$  nonparametrically using a sample of  $n$  data  $\{(x_i, Y_i), i = 1, \dots, n\}$  generated from

$$Y_i = m(x_i) + \epsilon_i, \quad i = 1, \dots, n, \quad (1.1)$$

where  $\epsilon_i$ 's are independent and identically distributed with mean 0 and finite variance  $\sigma^2$ . In this paper, we deal with a fixed design case only and assume that  $x_i$ 's are evenly spaced from 0 to 1. We also assume that the discontinuous regression function  $m$  is expressed by

$$m(x) = f(x) + \sum_{j=1}^p d_j I(x > s_j), \quad (1.2)$$

where  $f$  is a continuous function in the entire design interval,  $p$  is the number of jump points,  $S = \{s_j, j = 1, \dots, p\}$  are the jump positions, and  $\{d_j, j = 1, \dots, p\}$  are jump magnitudes.

There are two approaches to the nonparametric estimation of the discontinuous regression function. The first approach (usually called the direct approach or jump-preserving smoothing) estimates the regression function directly without detecting the jumps explicitly. The second approach (usually called the indirect approach) first estimates the locations of jump points using one of the various jump detection procedures and then estimates each smooth parts of the regression function by separately applying a traditional smoothing technique. Refer to Gijbels *et al.* (2007), and Park (2009a), for the literature on direct and indirect estimation methods.

In this paper we focus on the second approach (indirect approach). The major issue of the second approach is to obtain good estimates of jump locations. Many jump detectors are based on a kernel

---

This work was supported by a Hanshin University Research Grant.

<sup>1</sup> Professor, Department of Applied Statistics, Hanshin University, 441 Yangsan-Dong, Osan, Kyunggi-Do 447-791, Korea.  
E-mail: drpark@hs.ac.kr.

or local polynomial method and the performance of the jump detector heavily depends on the choice of the bandwidth. The bandwidth selection for the jump detector is an important issue for the second approach; however, only limited research has been done. Gijbels and Goderniaux (2004), has discussed the bootstrap bandwidth selection for the jump detection procedure of Gijbels *et al.* (1999), but their method is rather complicated. Zhang *et al.* (2009), proposed the bootstrap bandwidth selection for the jump detector which is based on local polynomial kernel smoothing. They modified the procedure of Gijbels and Goderniaux (2004), in several aspects, so they argued that the performance of their method should be superior to the method of Gijbels and Goderniaux (2004). Park (2009b), also proposed the bootstrap bandwidth selection for the case of  $p = 1$  by minimizing penalized mean squared error.

In this paper, we propose the bootstrap bandwidth selection methods for the case when  $p$  is known and  $p \geq 1$ . We also compare the performance of the proposed methods with existing methods through a simulation study.

## 2. Bandwidth Selection

### 2.1. Local linear jump detector

The common procedure of the local smoothing jump detector is to fit a smooth curve nonparametrically to the left and to the right of a given point, and the difference between these two fits contains the basic information about the jump discontinuity. There are various local smoothing jump detectors, but in this paper we consider the local linear jump detector proposed by Park (2009a). To state the jump detection procedure, the following notations will be used. Set

$$w_{j_1 j_2}^{(l)} = \sum_{i=1}^n (x_i - x)^{j_1} K_l^{j_2} \left( \frac{x_i - x}{g} \right), \quad \text{for } j_1 = 0, 1, 2 \text{ and } j_2, l = 1, 2.$$

$g$  is a smoothing parameter and  $K_1$  and  $K_2$  are kernel functions. Here the support of  $K_1$  and  $K_2$  are  $[0, 1]$  and  $[-1, 0]$ , respectively.

For  $x \in [\delta, 1 - \delta]$  where  $\delta$  is an arbitrarily small positive constant, the jump detector  $T(x)$  is defined by standardizing the difference as follows.

$$T(x) = \frac{\hat{m}_+(x) - \hat{m}_-(x)}{\sqrt{\text{Var}(\hat{m}_+(x) - \hat{m}_-(x))}} \quad (2.1)$$

where

$$\hat{m}_+(x) = \sum_{i=1}^n \frac{w_{21}^{(1)} - w_{11}^{(1)}(x_i - x)}{w_{01}^{(1)} w_{21}^{(1)} - (w_{11}^{(1)})^2} Y_i K_1 \left( \frac{x_i - x}{g} \right),$$

$$\hat{m}_-(x) = \sum_{i=1}^n \frac{w_{21}^{(2)} - w_{11}^{(2)}(x_i - x)}{w_{01}^{(2)} w_{21}^{(2)} - (w_{11}^{(2)})^2} Y_i K_2 \left( \frac{x_i - x}{g} \right),$$

and

$$\text{Var}(\hat{m}_+(x) - \hat{m}_-(x)) = V_1 + V_2$$

where

$$V_i = \frac{\sigma^2 \left[ (w_{21}^{(i)})^2 w_{02}^{(i)} - 2w_{21}^{(i)} w_{11}^{(i)} w_{12}^{(i)} + (w_{11}^{(i)})^2 w_{22}^{(i)} \right]}{\left( w_{01}^{(i)} w_{21}^{(i)} - (w_{11}^{(i)})^2 \right)^2}, \quad i = 1, 2.$$

When the number of jumps  $p$  is known, jump locations can be estimated as follows. Let  $\tilde{s}_j$  be the maximizer of  $|T(x)|$  over the set  $A_j$  where

$$A_j = [\delta, 1 - \delta] - \bigcup_{k=1}^{j-1} [\tilde{s}_k - g, \tilde{s}_k + g], \tag{2.2}$$

for  $j = 1, \dots, p$ . The order statistics of  $\{\tilde{s}_j, j = 1, \dots, p\}$  are denoted as  $\tilde{s}_{(1)} < \tilde{s}_{(2)} < \dots < \tilde{s}_{(p)}$ . Then the estimated jump locations are given by  $\hat{S} = \{\hat{s}_j = \tilde{s}_{(j)}, j = 1, \dots, p\}$ , and the estimated jump magnitudes are also given by  $\hat{d}_j = T(\hat{s}_j), j = 1, \dots, p$ .

### 2.2. Existing bandwidth selection procedures

The excellent choice of the bandwidth is indispensable in the good performance of the local smoothing jump detector, but only a little work has been done. Zhang *et al.* (2009) proposed the bootstrap bandwidth selection procedure and argued that their method is an improved version of the procedure of Gijbels and Goderniaux (2004). Henceforth, we denote their method as the ZSQ method, and the bandwidth value attained by their method as  $g_{ZSQ}$ . The algorithm for getting  $g_{ZSQ}$  is as follows:

#### Algorithm of ZSQ method

**Step 1** Define new observations

$$\tilde{Y}_i = Y_i - \sum_{j=1}^p \hat{d}_j I(x_i > \hat{s}_j), \quad i = 1, 2, \dots, n.$$

Denote  $\hat{f}$  as local smoothing estimator of  $f$  in (1.2) using bandwidth  $h$  from data  $\{(x_i, \tilde{Y}_i), i = 1, \dots, n\}$ . Then, define residuals

$$\hat{\epsilon}_i = Y_i - \hat{f}(x_i) - \sum_{j=1}^p \hat{d}_j I(x_i > \hat{s}_j), \quad i = 1, \dots, n.$$

**Step 2** Obtain  $B$  batches of resampled residuals  $\{\hat{\epsilon}_i^*, i = 1, \dots, n\}$  from  $\{\hat{\epsilon}_i, i = 1, \dots, n\}$  by random selection with replacement. For the  $b^{th}$  batch of resampled residuals, define pseudo data as follows.

$$\hat{Y}_i^* = \hat{f}(x_i) + \sum_{j=1}^p \hat{d}_j I(x_i > \hat{s}_j) + \hat{\epsilon}_i^*, \quad i = 1, \dots, n.$$

**Step 3** Apply the local polynomial jump detection procedure with bandwidth  $g$  to the  $b^{th}$  pseudo data, and the set of detected jump is denoted as  $\hat{S}_b$ . Then the Hausdorff distance  $d_H(S, \hat{S}; g, h)$  is

estimated by

$$\hat{d}_H(S, \hat{S}; g, h) = \frac{1}{B} \sum_{b=1}^B d_H(\hat{S}, \hat{S}_b; g, h)$$

where  $d_H(\hat{S}, \hat{S}_b; g, h)$  is the Hausdorff distance between  $\hat{S}$  and  $\hat{S}_b$ , which is defined by

$$d_H(\hat{S}, \hat{S}_b; g, h) = \max \left\{ \sup_{s_1 \in \hat{S}} \inf_{s_2 \in \hat{S}_b} |s_1 - s_2|, \sup_{s_1 \in \hat{S}_b} \inf_{s_2 \in \hat{S}} |s_1 - s_2| \right\}.$$

**Step 4** The bandwidth  $g_{ZSQ}$  is attained by

$$\min_{g>0} \left[ \min_{h>0} \hat{d}_H(S, \hat{S}; g, h) \right].$$

Park (2009b), also proposed the bandwidth selection method for the case of  $p = 1$ . Since it is common to measure the closeness of the estimator  $\hat{s}_1$  to its target parameter  $s_1$  by  $\text{MSE}(\hat{s}_1)$ , he tried to obtain the optimal bandwidth by minimizing Monte Carlo  $\text{MSE}(\hat{s}_1)$ . In his paper, the jump point is assumed to be located inside of  $[g, 1 - g]$  and this is a kind of a universal assumption about the range of the discontinuity points. However, because of this assumption, it is always true that  $g \leq \hat{s}_1 \leq 1 - g$ , so the large value of  $g$  reduces the range of  $\hat{s}_1$  and it has a direct effect on the large reduction of the Monte Carlo  $\text{Var}(\hat{s}_1)$  and this also entails a large reduction of the Monte Carlo  $\text{MSE}(\hat{s}_1)$  as long as  $s_1 \in [g, 1 - g]$ . Therefore, the Monte Carlo  $\text{MSE}(\hat{s}_1)$  turns out to be the decreasing function of  $g$  and it is not possible to obtain the optimal bandwidth by minimizing Monte Carlo  $\text{MSE}(\hat{s}_1)$ . As one way to resolve this problem, he defined the penalized mean squared error (PMSE) in order to penalize the large value of  $g$  for computing  $\text{MSE}(\hat{s}_1)$  and proposed the bootstrap bandwidth selection method by minimizing the Monte Carlo  $\text{PMSE}(\hat{s}_1)$ .

Another way to resolve this problem is to assume that the jump points are located inside of  $[\delta, 1 - \delta]$  where  $\delta$  is an arbitrarily small positive constant. Wu and Chu (1993), also used the same assumption. Since  $\delta$  has nothing to do with  $g$ , there is no guarantee that large value of  $g$  induces large reduction of Monte Carlo  $\text{Var}(\hat{s}_1)$ . The only problem for this assumption is that the jump detector with the bandwidth  $g$  may suffer from the boundary effect at the point near each end of design interval when  $g > \delta$ . The jump detector considered in this paper, however, is based on the local linear regression, so we do not have to worry about the boundary effect problem. Under this assumption, we propose two new bandwidth selection methods in Section 2.3.

### 2.3. Bootstrap bandwidth selection method

For the jump detection procedure of (2.2), we consider two criteria for assessing the closeness of  $\hat{S}$  to  $S$ . One is the sum of  $\text{MSE}$ ,  $\sum_{j=1}^p \text{MSE}(\hat{s}_j)$  and the other is the Hausdorff distance between  $S$  and  $\hat{S}$ ,  $d_H(S, \hat{S})$ . It is common to measure the closeness of an estimator to its target parameter by  $\text{MSE}$ , so the sum of  $\text{MSE}$  is the natural extension for the multi-parameter case. The Hausdorff distance is a similarity measure between two arbitrary point sets, and is used in many fields. For example, it is used to assess the accuracy of alignment between two images.

Using these criteria we propose two bandwidth selection methods. Since the two criteria include the unknown parameter  $S$ , we need to use a bootstrap procedure to estimate both the sum of  $\text{MSE}$  and

the Hausdorff distance. We denote the method based on the sum of MSE as P-MSE, and the method based on the Hausdorff distance as P-HD. Algorithms of P-MSE and P-HD are as follows.

### Bootstrap Algorithm

#### Step 1: Common Step for the two procedures

Obtain  $B$  batches of resampled data  $\chi^* = \{(x_1^*, Y_1^*), \dots, (x_n^*, Y_n^*)\}$  from the observed sample  $\chi = \{(x_1, Y_1), \dots, (x_n, Y_n)\}$  by random selection of a pair  $(x_i, Y_i)$  with replacement. Apply the jump detection procedure (2.2) to  $b^{\text{th}}$  batch of resampled data with the bandwidth  $g$ , and the set of detected jump points is denoted as  $\hat{S}_b = \{\hat{s}_j^{(b)}, j = 1, \dots, p\}$ .

#### Step 2-1: P-MSE method

The bandwidth  $g_{MSE}$  is attained by

$$\min_{g>0} \sum_{j=1}^p \left[ \frac{1}{B-1} \sum_{b=1}^B \left( \hat{s}_j^{(b)} - \overline{\hat{s}_j^{(*)}} \right)^2 + \left( \overline{\hat{s}_j^{(*)}} - \hat{s}_j \right)^2 \right]$$

where  $\overline{\hat{s}_j^{(*)}} = \sum_{b=1}^B \hat{s}_j^{(b)} / B$  for  $j = 1, \dots, p$ .

#### Step 2-2: P-HD method

The bandwidth  $g_{HD}$  is attained by

$$\min_{g>0} \left[ \frac{1}{B} \sum_{b=1}^B d_H(\hat{S}, \hat{S}_b; g) \right].$$

In the above algorithm, Step 1 is about constructing  $B$  batches of  $\hat{S}_b$ , and this is a common step for two bandwidth selection procedures. Step 2-1 describes the first bandwidth selection method. In this method, we select the value of  $g$  at which the bootstrap estimate of  $\sum_{j=1}^p \text{MSE}(\hat{s}_j)$  is minimal, and we denote this value as  $g_{MSE}$ . Step 2-2 describes the second bandwidth selection method. We select value of  $g$  at which the bootstrap estimate of  $d_H(S, \hat{S})$  is minimal, and we denote this value as  $g_{HD}$ .

### 3. Simulation Study

A simulation study was conducted to evaluate the finite sample properties of the bootstrap bandwidth selection methods proposed in Section 2.3. We also considered the bandwidth selection method of Zhang *et al.* (2009) to compare the performance with the proposed methods. We considered the following discontinuous regression models:

$$m_1(x) = \begin{cases} 2/3 - 2x, & x \leq 0.3, \\ 1, & 0.3 < x \leq 0.7, \\ -2(x - 2/3)(x - 2), & x > 0.7, \end{cases}$$

$$m_2(x) = \sin(4\pi x)I(x \leq 0.3) - \sin(4\pi x)I(0.3 < x \leq 0.7) + \sin(4\pi x)I(x > 0.7),$$

$$m_3(x) = x^2 + .5I(x > s_j), \quad j = 1, 2,$$

$$m_4(x) = \sin(2\pi x) + I(x > s_j), \quad j = 1, 2,$$

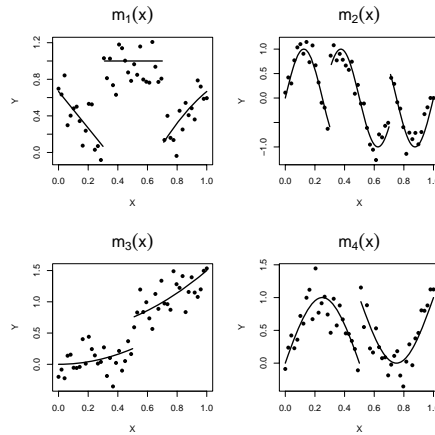


Figure 1: The true discontinuous regression functions and the typical data set of each model

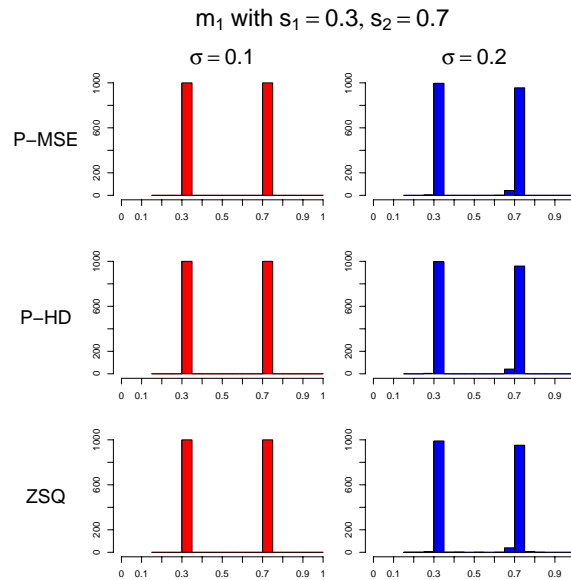


Figure 2: Histogram of  $\hat{s}_1$  and  $\hat{s}_2$  for  $m_1$

where  $s_1 = 0.5$  and  $s_2 = 0.7$ . The design points,  $x_i$ 's were evenly spaced from 0 to 1 with a sample size of  $n = 50$ , and  $Y_i$ 's were generated from model (1.1) with  $\epsilon_i \sim N(0, \sigma^2)$  where  $\sigma = 0.1, 0.2$ . When  $\sigma = 0.2$ , the four true discontinuous regression functions and the typical data set of each model are shown in Figure 1.

In all studies we considered 1000 simulation samples and the number of bootstrap replicates was  $B = 100$ . For the kernel function of  $T(x)$  in (2.1), we chose  $K_1(x) = 1.5(1 - x^2)I_{[0,1]}(x)$  and  $K_2(x) = K_1(-x)$  for all  $x$ . For each sample, the values of  $g_{MSE}$ ,  $g_{HD}$ , and  $g_{ZSQ}$  were selected, and the locations of the jump positions were estimated by the jump detection procedure of (2.2) with  $\delta = 0.1$ . The selection of the bandwidth value for  $g_{ZSQ}$  was done by the function *bds* of the package *polydect* in R. The usage of the function *bds* is as follows: *bds*( $g, B, h, X, Y, x, order$ ) where  $g$  is the bandwidth

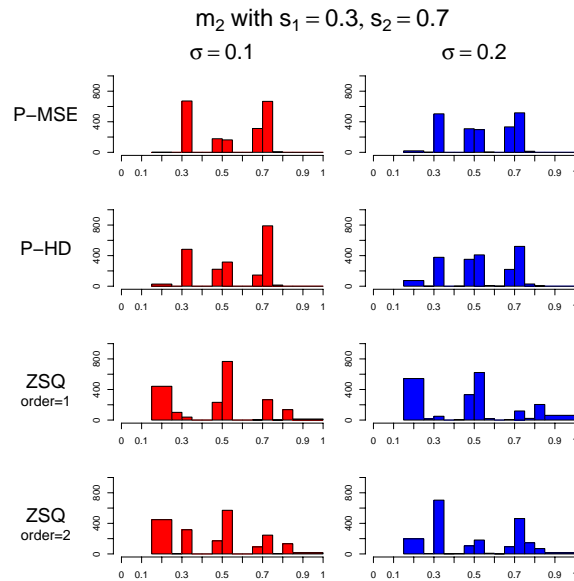


Figure 3: Histogram of  $\hat{s}_1$  and  $\hat{s}_2$  for  $m_2$

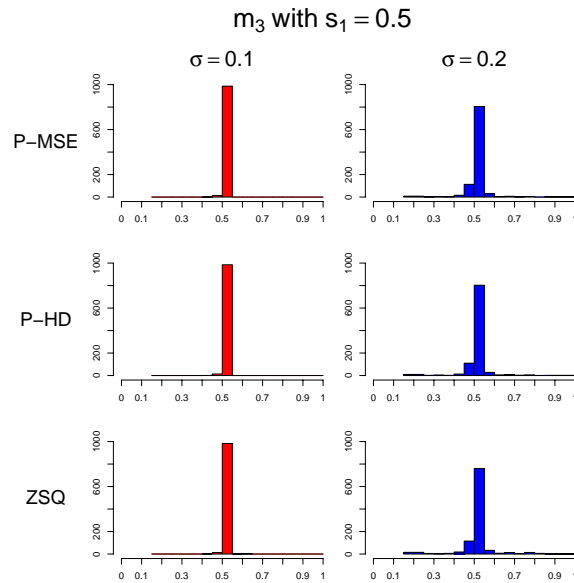


Figure 4: Histogram of  $\hat{s}_1$  for  $m_3$  with  $s_1 = 0.5$

value for detecting jump locations,  $B$  indicates the times of Bootstrap,  $h$  is the bandwidth value to estimate the function  $f$  with jump removed,  $X$  is the design points,  $Y$  is the observed data points,  $x$  is a vector that corresponds to the interval  $[\delta, 1 - \delta]$ , and the *order* specifies the order of the local polynomial estimator used. This function returns the value indicating the average of the Hausdorff distance between the detected and true set of jump positions, while the detection is repeated by  $B$

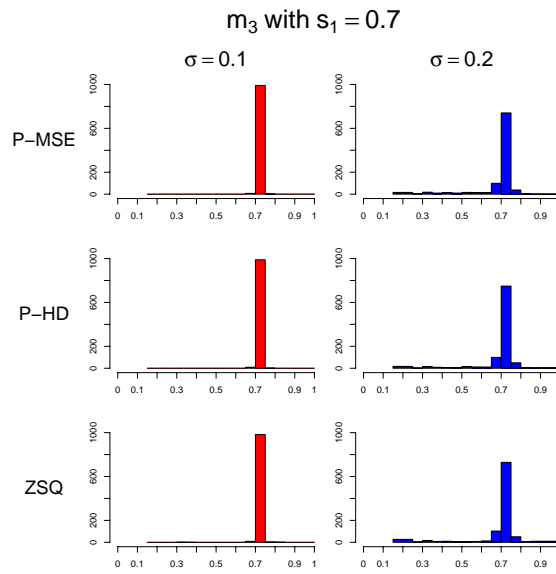


Figure 5: Histogram of  $\hat{\delta}_1$  for  $m_3$  with  $s_1 = 0.7$

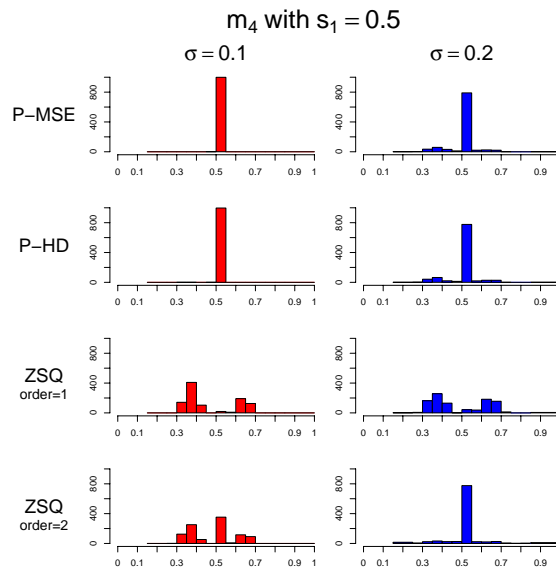


Figure 6: Histogram of  $\hat{\delta}_1$  for  $m_4$  with  $s_1 = 0.5$

times.

The simulation results are summarized from Figure 2 to Figure 7. In these figures, the distributions of  $\hat{\delta}_j$  out of 1000 simulation samples are presented by the histograms. The means and standard deviations of  $\hat{\delta}_j$  are also listed in Table 1. The distribution of the attained bandwidth values is also an important result of the simulation. In Table 2, the means and standard deviations of the bandwidths are listed.



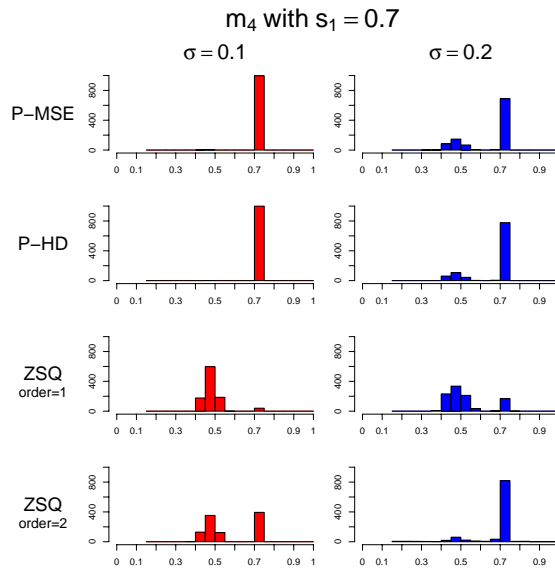


Figure 7: Histogram of  $\hat{s}_1$  for  $m_4$  with  $s_1 = 0.7$

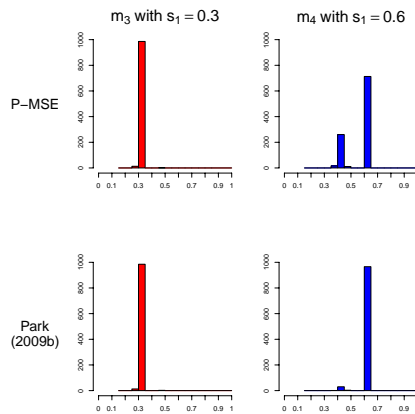


Figure 8: Histogram of  $\hat{s}_1$  for  $m_3$  with  $s_1 = 0.3$  and for  $m_4$  with  $s_1 = 0.6$

The function  $m_1$  and  $m_2$  are the two jump points models. As we can see in Figure 2, the distributions of  $\hat{s}_j$  of three bandwidth selection methods appear almost identical. In Table 1, we can see that the mean values of  $\hat{s}_j$  of three bandwidth selection methods for  $m_1$  are the same, but the ZSQ method has larger standard deviation. Note that the distributions of the bandwidths are almost identical for  $m_1$  in Table 2.

In  $m_2$ , we find out the problem of the ZSQ method. In Figure 3, the histograms of  $\hat{s}_j$  for the ZSQ method with the jump detection procedure of (2.2) are placed in the third row. Note that the jump detection procedure of (2.2) is based on the local linear regression, and its order is annotated in the figure. From these histograms, we can notice that the ZSQ method with the local linear jump detector does not work at all. It is unclear why the ZSQ method with the local linear jump detector produced the unexpected inferior results in  $m_2$ , but we can guess that the large curvature of  $m_2$  may cause the

Table 1: Means and standard deviations of  $\hat{s}_j$  in 1000 simulation samples. In both  $m_2$  and  $m_4$ , the results of the ZSQ method are for the case of  $order = 2$  in the function  $bds$ .

		$s_j$	$m_1$	$m_1$	$m_2$	$m_2$	$m_3$	$m_3$	$m_4$	$m_4$
			0.3	0.7	0.3	0.7	0.5	0.7	0.5	0.7
$\sigma = .1$	mean ( $\hat{s}_j$ )	P-MSE	.306	.714	.370	.705	.509	.714	.510	.713
		P-HD	.306	.714	.398	.702	.509	.714	.509	.714
		ZSQ	.306	.714	.297	.624	.510	.713	.477	.571
	sd ( $\hat{s}_j$ )	P-MSE	0	0	.003	.090	.025	.003	.001	.015
		P-HD	0	0	.003	.102	.044	.003	.007	.007
		ZSQ	0	0	.010	.126	.126	.021	.106	.117
$\sigma = .2$	mean ( $\hat{s}_j$ )	P-MSE	.306	.713	.396	.679	.508	.682	.500	.642
		P-HD	.306	.713	.403	.666	.509	.689	.500	.662
		ZSQ	.307	.712	.309	.678	.511	.684	.507	.681
	sd ( $\hat{s}_j$ )	P-MSE	.003	.005	.098	.069	.049	.106	.060	.110
		P-HD	.003	.004	.109	.084	.054	.101	.064	.099
		ZSQ	.025	.021	.094	.118	.070	.117	.083	.093

Table 2: Means and standard deviations of attained bandwidths in 1000 simulation samples. In both  $m_2$  and  $m_4$ , the results of  $g_{ZSQ}$  are for the case of  $order = 2$  in the function  $bds$ .

			$m_1$	$m_2$	$m_3$	$m_3$	$m_4$	$m_4$
					( $s_1 = .5$ )	( $s_1 = .7$ )	( $s_1 = .5$ )	( $s_1 = .7$ )
$\sigma = .1$	mean	$g_{MSE}$	.196	.127	.194	.191	.087	.104
		$g_{HD}$	.195	.135	.193	.187	.084	.103
		$g_{ZSQ}$	.234	.201	.219	.215	.172	.180
	sd	$g_{MSE}$	.037	.047	.041	.040	.021	.024
		$g_{HD}$	.039	.061	.041	.041	.019	.024
		$g_{ZSQ}$	.026	.065	.056	.057	.080	.082
$\sigma = .2$	mean	$g_{MSE}$	.223	.155	.213	.207	.117	.151
		$g_{HD}$	.221	.166	.213	.208	.113	.137
		$g_{ZSQ}$	.238	.108	.213	.213	.074	.073
	sd	$g_{MSE}$	.025	.039	.046	.051	.059	.063
		$g_{HD}$	.027	.046	.047	.048	.059	.056
		$g_{ZSQ}$	.039	.088	.061	.060	.057	.058

poor performance. Thus, we include the ZSQ method with the local quadratic jump detector in the simulation study for  $m_2$ , and the results are displayed in the bottom row of Figure 3. For  $\sigma = 0.2$  case, the ZQS method with the quadratic detector shows significantly improved performance; however, it still does not work for the  $\sigma = 0.1$  case. This is also a very surprising result since in the case of  $\sigma = 0.1$  it is easier to detect the jump points than the case of  $\sigma = 0.2$ . In Table 1 and Table 2, only the results for the ZSQ method with the local quadratic jump detector are displayed for  $m_2$ .

The function  $m_3$  and  $m_4$  are the one jump point models. For  $m_3$ , we can notice that the performance of three bandwidth selection methods are almost identical. The distributions of both  $\hat{s}_j$  and the attained bandwidths look alike. In  $m_4$ , we find exactly same problem of the ZSQ method as in  $m_2$ . As we can see in Figure 6 and Figure 7, the ZSQ method with the local linear jump detector does not work at all; in addition, the ZSQ method with the local quadratic jump detector only shows significantly improved results only for the case of  $\sigma = 0.2$ .

The bandwidth selection procedure proposed in Park (2009b), selects the bandwidth value by minimizing the following penalized mean squared error

$$PMSE = \log \left[ E (\hat{s}_1 - s_1)^2 \right] + \frac{10}{\log(n)(1 - 2g)},$$

so it seems that this method is comparable with the P-MSE method for the single jump point model.

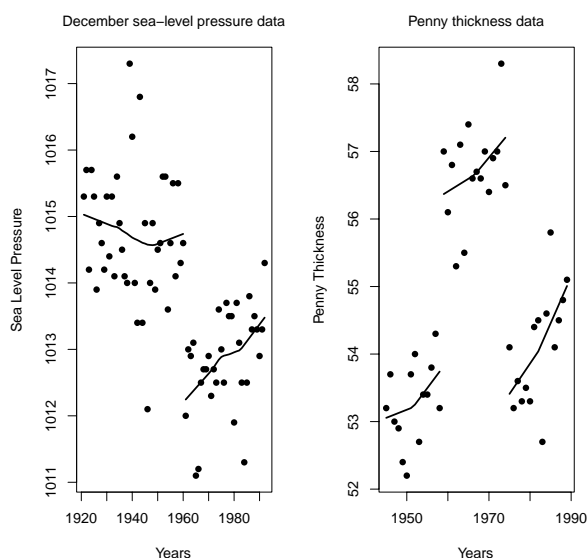


Figure 9: The December sea-level pressure data and the penny thickness data. The solid curves represent non-parametric regression estimates.

However, the method in Park (2009b), assumes that the jump point is located inside of  $[g, 1 - g]$ , which is not true in this paper. Even though the direct comparison between two methods may not have much meaning, we may have a chance to understand which assumption is more appropriate through the simulation study that compares the performance of the two methods. For  $m_3$  with  $s_1 = 0.3$  and  $m_4$  with  $s_1 = 0.6$ , we conducted the simulation study. In both models, we considered only  $\sigma = 0.1$  case. The simulation results are summarized in Figure 8. For  $m_3$  with  $s_1 = 0.3$  case, two methods show almost identical performance; however, P-MSE shows inferior results in  $m_4$  with  $s_1 = 0.6$ . We do not want to make any premature conclusion based on this simulation study due to the need for a thorough investigation that compares the performance of the two methods.

#### 4. Applications

In this section, we apply the jump detection procedure of (2.2) with the proposed bandwidth selection methods to real data sets. The first data is the December sea-level pressure from 1921–1992 in Bombay, which is shown in the left panel of Figure 9. It is well known that the sea-level pressure has a discontinuity point around 1960 and 1961. This data set is given in Qiu (2005). The bandwidth value chosen by P-MSE, P-HD, and ZSQ is  $g = 18$ . All of them choose the same bandwidth value, and the jump detection procedure (2.2) with  $g = 18$  produces  $\hat{s}_1 = 1961$ . After we estimate the jump position, we divide whole interval into two parts and then estimate the discontinuous regression function by applying a local linear regression model to each smooth part separately. The fitted curves are presented in Figure 9.

The second data set is the penny thickness data, which is available from Scott (1992), and shown in the right panel of Figure 9. This data consists of thickness in mils of a sample of 90 U.S. Lincoln pennies minted from 1945 to 1989. Since two pennies were measured for each year, we replace the original data by the mean of each pair per year. It is known that penny thickness measure has two jumps around 1959 and 1975. For this data set, three bandwidth selection methods also yield the

same bandwidth value,  $g = 8$ . With this bandwidth value, the jump detection procedure (2.2) produces  $\hat{s}_1 = 1959$  and  $\hat{s}_2 = 1975$ . Gijbels and Goderniaux (2004), treated this data set and selected  $g = 5$  by their bootstrap bandwidth selection method. The bandwidth value of  $g = 5$  gives the same jump points estimates. The discontinuous nonparametric regression function is estimated and displayed in Figure 9.

## 5. Conclusion

When we apply the local linear jump detection procedure for the estimation of the locations of the jump points of the unknown regression function, we have to select the bandwidth value for the jump detector; it is clear that the performance of the jump detector depends heavily on the choice of the bandwidth. However, limited attention has been paid to this important issue of the choice of the bandwidth. In this paper, we proposed two fully data adaptive bandwidth selection methods for local linear jump detector. Through the simulation study we investigated the performance of the proposed methods. The performance of two proposed methods are almost identical, and it seems to be verified that the proposed methods are competitive methods to select the bandwidth in practice.

To use the ZSQ method, there are actually two bandwidths involved, one is for detection and the other is to estimate the continuous part. Thus we need to try different combinations of the two and find a best combination. However, the proposed methods need only one bandwidth for detection that provides much convenience in practice. Moreover, it seems that the ZSQ method is very sensitive to the amount of curvature of the regression function, whereas proposed methods show consistent results for all cases we have considered in the simulation. This is a desirable property of the bandwidth selector since we usually have no idea about the amount of curvature of the unknown regression function in practice.

## References

- Gijbels, I. and Goderniaux, A. C. (2004). Bandwidth selection for change point estimation in nonparametric regression, *Technometrics*, **46**, 76–86.
- Gijbels, I., Hall, P. and Kneip, A. (1999). On the estimation of jump points in smooth curves, *The Annals of the Institute of Statistical Mathematics*, **51**, 231–251.
- Gijbels, I., Lambert, A. and Qiu, P. (2007). Jump-preserving regression and smoothing using local linear fitting: A compromise, *Annals of the Institute of Statistical Mathematics*, **59**, 235–272.
- Park, D. (2009a). Comparison of jump-preserving smoothing and smoothing based on jump detector, *Communications of the Korean Statistical Society*, **16**, 519–528.
- Park, D. (2009b). Bandwidth selection for local smoothing jump detector, *Communications of the Korean Statistical Society*, **16**, 1047–1054.
- Qiu, P. (2005). *Image Processing and Jump Regression Analysis*, John Wiley & Sons, New Jersey.
- Scott, D. W. (1992). *Multivariate Density Estimation: Theory, Practice and Visualization*, Wiley-Interscience, New York.
- Zhang, B., Su, Z. and Qiu, P. (2009). On jump detection in regression curves using local polynomial kernel estimation, *Pakistan Journal of Statistics*, **25**, 505–528.
- Wu, J. S. and Chu, C. K. (1993). Kernel type estimators of jump points and values of a regression function, *The Annals of Statistics*, **21**, 1545–1566.