Change-point Problems
IMS Lecture Notes – Monograph Series (Volume 23, 1994)

NONPARAMETRIC ESTIMATION OF FUNCTIONS WITH JUMP DISCONTINUITIES

BY R. L. EUBANK AND P. L. SPECKMAN
Texas A&M University University and University of Missouri-Columbia

The problem of estimating a function with a jump discontinuity in one of its derivatives is considered. A semi-parametric framework is employed to formulate the problem, and a least-squares type estimator of the jump point is proposed for this setting. The asymptotic properties of this estimator are derived, including consistency and asymptotic distribution theory.

1. Introduction. The problem we consider is that of locating the point of a cusp (or change-point of the first derivative) and the size of the change in the first derivative for an otherwise smooth function. This problem has been studied in one form or another by several authors. Wahba (1984) and Engle, Granger, Rice and Weiss (1986) were among the first to use a semiparametric approach to derive partial smoothing splines for estimating curves with cusps assuming the change-point is known. This article adapts a related technique given by Eubank and Speckman (1991) to the problem where the change-point is unknown. Our results are similar to those of Müller (1992) who uses boundary kernels to locate the change-point and estimate the size of the change. See Müller (1992) for a more complete list of references.

We consider the following model. Responses $z_{1n}, \ldots, z_{nn}$ are obtained at equally spaced design points $t_{rn} = r/n$, $r = 1, \ldots, n$. The z_{rn} and t_{rn} are related under the model

$$z_{rn} = g(t_{rn}) + \varepsilon_{rn}, r = 1, \ldots, n, \tag{1}$$

where the ε_{rn} are independent, identically distributed random variables having zero means and common variance σ^2. (Throughout the rest of the article, we will suppress the dependence of the z_{rn} and t_{rn} on n.)

The unknown regression function g is assumed to be continuous on $[0, 1]$ and twice continuously differentiable on $[0, \tau_0)$ and $(\tau_0, 1]$ for some point $\tau_0 \in$

AMS 1991 Subject Classification: Primary 62G20; Secondary 62G05

Key words and phrases: Change-point, nonparametric regression, asymptotic properties, semiparametric model.

$(0,1)$. At τ_0, g is assumed to have a jump discontinuity in its first derivative. Thus g can be written as the sum of a basis function and a smooth part

$$g(t) = \beta_0 \phi_{\tau_0}(t) + f(t),$$

where

$$\phi_\tau(t) = \begin{cases} 0, & t < \tau, \\ t - \tau, & t \geq \tau, \end{cases}$$

τ_0 is the unknown change-point, β_0 is the size of the discontinuity in the first derivative of g at τ_0, and $f(t)$ is an unknown continuously differentiable function. We assume that $f''(t)$ exists and is continuous for all $t \neq \tau_0$, and that f has right and left continuous second derivatives at τ_0, $f''(\tau_0^+) = \lim_{t \downarrow \tau_0} f''(t)$ and $f''(\tau_0^-) = \lim_{t \uparrow \tau_0} f''(t)$. We further assume that $\beta_0 \neq 0$, which excludes treatment of the problem of testing for the existence of a change point. This latter problem is different and will not be discussed here.

Our estimator $\hat{\beta}$ for β_0 is based on ideas from semiparametric estimation. This estimator and a related weighted-least squares estimator $\hat{\tau}$ of τ_0 are defined in Section 2. The main theorems are presented in Section 3, establishing consistency and asymptotic normality. Under standard assumptions for second order smoothing, we show that

$$\hat{\tau} - \tau_0 = O_p(n^{-2/5})$$

and

$$\hat{\beta} - \beta_0 = O_p(n^{-1/5}).$$

This latter rate is known to be asymptotically optimal for estimating a first derivative under the conditions we impose on f.

Our model is very similar to one considered by Müller (1992). He treated a class of problems with a single discontinuity in an arbitrary derivative. The problem addressed here is related to Müller's case $\nu = 1$. However, Müller required $f \in C^{k+\nu}[0,1]$ for some $k \geq 2$. In particular, to apply Müller's results to the problem here, we would need $f \in C^3[0,1]$. We relax the requirement of smooth higher derivatives for f because it seems natural in many applications that a response function with a discontinuity in the first derivative would also have a discontinuous second derivative. Our asymptotic results in Section 3 reflect this weakened assumption. While our results are qualitatively similar to Müller's, we have not made any finite or large sample comparisons. A more precise connection between the two methods is given at the end of the next section.

2. Kernel Smoothing and Semiparametric Estimation of the Change-point. To begin, we cast the problem in vector notation. Write

$$z = g + \varepsilon,$$

where $z = (z_1, \ldots, z_n)'$, etc. Letting $\phi_\tau = (\phi_\tau(t_1), \ldots, \phi_\tau(t_n))'$, the goal is to estimate τ_0, β_0 and

$$g = \beta_0 \phi_{\tau_0} + f.$$

The foundation for our method is a semiparametric approach to estimating β_0 found in Eubank and Speckman (1991) based on one dimensional smoothing. In the application here, we focus on second-order kernel smoothers and assume observations as in (1). The fundamental kernel smoother for this case is (cf. Eubank, 1988)

$$g_h(t) = \frac{1}{nh} \sum_{r=1}^{n} K\left(\frac{t - t_r}{h}\right) z_r,$$

where h is the bandwidth and K is a *kernel* function satisfying the second order conditions

$$K(t) = 0, |t| > 1, \tag{A1}$$

$$\int_{-1}^{1} K(t)dt = 1, \tag{A2}$$

$$\int_{-1}^{1} tK(t)dt = 0, \tag{A3}$$

and

$$\int_{-1}^{1} t^2 K(t)dt = C_0 > 0. \tag{A4}$$

Letting S denote the $n \times n$ matrix of kernel weights, i.e.

$$S = \frac{1}{nh}\left[K\left(\frac{t_i - t_j}{h}\right)\right]_{i,j=1,n},$$

and $g_h = (g_h(t_1), \ldots, g_h(t_n))'$, the smooth can be written in vector form as $g_h = Sz$.

If the change-point τ is known, β_0 can be estimated efficiently by minimizing

$$||(I - S)(z - \phi_\tau \beta)||^2 \tag{2}$$

over β (see Eubank and Speckman, 1991), and the resulting estimate is

$$\hat{\beta}(\tau) = \frac{\phi_\tau'(I - S)^2 z}{\phi_\tau'(I - S)^2 \phi_\tau}.$$

Setting

$$P_\tau = \frac{(I-S)\phi_\tau\phi'_\tau(I-S)}{\phi'_\tau(I-S)^2\phi_\tau},$$

a natural estimate of g is provided by

$$\hat{g}_\tau = S(z - \phi_\tau\hat{\beta}(\tau)) + \phi_\tau\hat{\beta}(\tau) = Sz + P_\tau(I-S)z,$$

and the residual sum of squares can be written as

$$\|z - \hat{g}_\tau\|^2 = \|z - Sz - P_\tau(I-S)z\|^2 = \|(I-P_\tau)(I-S)z\|^2. \tag{3}$$

We propose estimating τ_0 by minimizing (3) as a function of τ. Note that this is equivalent to minimizing (2) jointly in (τ, β).

This problem simplifies because β is a scalar and P_τ is a projection. Since

$$\|(I-P_\tau)(I-S)z\|^2 = \|(I-S)z\|^2 - \|P_\tau(I-S)z\|^2,$$

minimizing (3) is equivalent to maximizing $\|P_\tau(I-S)z\|^2 = \hat{\beta}(\tau)^2\phi'_\tau(I-S)^2\phi_\tau$ with respect to τ. We will show in Lemma 4 below that $\phi'_\tau(I-S)^2\phi_\tau$ is essentially independent of τ and $\hat{\beta}(\tau)$ is consistent. Hence, it is asymptotically equivalent to maximize (or minimize) $\hat{\beta}(\tau)$ in τ.

Several authors including Müller (1992), Qiu (1992) and Wu and Chu (1992) have used the difference of two kernel estimates to estimate the change in a function value or derivative value at a point. For example, Müller's estimate of β at τ is based on the difference of two one-sided kernel estimates of the form

$$\hat{g}'_\pm(\tau) = \frac{1}{nh^2}\sum_{j=1}^n K'_\pm\left(\frac{\tau - t_j}{h}\right)z_j,$$

where K_+ is a smooth kernel with support $[-1,0]$, K_- is a smooth kernel with support $[0,1]$, and $K'_\pm$ consequently are both suitable kernels for estimating the first derivative. Typically, one can take $K'_+(x) = -K'_-(x)$. Then $\tilde{\beta}(\tau) = \hat{g}'_+(\tau) - \hat{g}'_-(\tau)$ is an estimate of the change (if any) in g' at τ analogous to $\hat{\beta}(\tau)$.

In an asymptotic sense, the semiparametric method used here is closely related. One consequence of Lemma 3 below is that there is a symmetric function $\tilde{\psi}$ with support [-2,2] such that to a good approximation

$$\phi'_\tau(I-S)^2z \approx h\sum_{j=1}^n \tilde{\psi}\left(\frac{t_j - \tau}{h}\right)z_j.$$

It follows that

$$\hat{\beta}(\tau) \approx \frac{1}{Bnh^3} \sum_{j=1}^{n} \tilde{\psi}\left(\frac{t_j - \tau}{h}\right) z_j$$

for some constant B. As in Lemma 1 below, it can be shown that $\tilde{\psi}$ is symmetric and satisfies the integral conditions $\int_0^2 \tilde{\psi}(u)du = 0$, $\int_0^2 u\tilde{\psi}(u)du \neq 0$. Thus, suitably rescaled, $\tilde{\psi}$ behaves like K'_+ on $[-2, 0]$ and like $-K'_-$ on $[0, 2]$. Because the kernels $K_\pm$ of Müller (1992) have stronger smoothness assumptions than we obtain for $\tilde{\psi}$, Müller's proofs do not carry over directly. However, his methods are extended below to obtain the asymptotic behavior of $\hat{\tau}$ and $\hat{\beta}$.

3. The Main Results. The method of proof is modeled on Müller (1992), and the approach is based on the following weak convergence result. Assume $\tau_0 \in (0, 1)$, let

$$\tau = \tau_0 + y\frac{h}{\sqrt{nh^3}}, \quad y \in [-A, A],$$

where $A > 0$, and define a process in $C[-A, A]$

$$C_n(y) = nh^3[\hat{\beta}(\tau) - \hat{\beta}(\tau_0)].$$

In the following, the notation $h \sim n^{-\alpha}$ means $n^\alpha h \to c$, where c is a constant satisfying $0 < c < \infty$, and "$\Rightarrow$" denotes weak convergence.

THEOREM 1. *If $h \sim n^{-\alpha}$ for $\alpha \geq 1/5$ and $n \to \infty$,*

$$C_n(\cdot) \Rightarrow X(\cdot)$$

on $C[-A, A]$, where

$$X(y) = -\beta_0 C_1 y^2 + (C_2 \Delta_1 L + C_3 \sigma U)y,$$

$$U \sim N(0, 1),$$

$$\Delta_1 = f''(\tau_0^+) - f''(\tau_0^-),$$

and

$$L = \lim_{n\to\infty} nh^4/\sqrt{nh^3}.$$

The constants are given by $C_1 = B_4/B_1$, $C_2 = B_3/B_1$ and $C_3 = \sqrt{B_5}/B_1$, where B_1, B_3, B_4 and B_5 are defined in Lemmas 4, 7, 8 and 9 below respectively.

The proof is postponed to Section 4.

THEOREM 2. *Under the conditions of Theorem 1,*

$$\frac{\sqrt{nh^3}}{h}(\hat{\tau} - \tau_0) \xrightarrow{\mathcal{D}} N\left(\frac{C_2\Delta_1 L}{2\beta_0 C_1}, \left(\frac{C_3\sigma}{2\beta_0 C_1}\right)^2\right).$$

REMARK 1. If $h \sim n^{-1/5}$ (the "optimal" rate for second order smoothing), Theorem 2 gives

$$\hat{\tau} - \tau_0 = O_p(n^{-2/5}).$$

On the other hand, if $h \sim n^{-\alpha}$ for $\alpha > 1/5$ implying that

$$L = \lim_{n\to\infty} \frac{nh^4}{\sqrt{nh^3}} = 0,$$

or if f'' exists and is continuous at τ_0 so that $\Delta_1 = 0$, then

$$\frac{\sqrt{nh^3}}{h}(\hat{\tau} - \tau_0) \Rightarrow N\left(0, \left(\frac{C_3\sigma}{\beta_0 C_1}\right)^2\right).$$

PROOF OF THEOREM 2. Under Whitt's (1970) metric, the technique of Eddy (1980) can be used to extend convergence of C_n from $C[-A, A]$ to $C(-\infty, \infty)$ (cf. Müller, 1992). The process $X(\cdot)$ has a unique maximum if $\beta_0 > 0$ (or minimum if $\beta_0 < 0$) at

$$\begin{aligned} y^* &= \frac{C_2\Delta_1 L + C_3\sigma U}{2\beta_0 C_1} \\ &\sim N\left(\frac{C_2\Delta_1 L}{2\beta_0 C_1}, \left(\frac{C_3\sigma}{2\beta_0 C_1}\right)^2\right). \end{aligned} \tag{4}$$

The result now follows from the functional mapping theorem (cf. Billingsley, 1968). ∎

THEOREM 3. *Under the conditions of Theorems 1 and 2,*

$$\sqrt{nh^3}[\hat{\beta}(\tau_0) - \beta_0] \xrightarrow{\mathcal{D}} N(C_4\Delta_2 L, C_5), \tag{5a}$$

and

$$\sqrt{nh^3}[\hat{\beta}(\hat{\tau}) - \beta_0] = \sqrt{nh^3}[\hat{\beta}(\tau_0) - \beta_0] + o_p(1), \tag{5b}$$

where L is defined in Theorem 1 and

$$\Delta_2 = [f''(\tau_0^+) + f''(\tau_0^-)]/2.$$

The constants in equation (5a) are $C_4 = B_2/B_1$ and $C_5 = \tilde{B}_1/B_1^2$, where B_1, $\tilde{B}_1$ and B_2 are defined in Lemmas 4, 5 and 6 respectively.

The proof will be given in Section 4.

REMARK 2. It is possible to show also that $\hat{\tau}$ and $\hat{\beta}(\hat{\tau})$ are asymptotically independent. For brevity, the proof will not be given here.

REMARK 3. If $h \sim n^{-1/5}$, then Theorem 3 shows that $\hat{\beta}(\hat{\tau}) - \beta_0 = O_p(n^{-1/5})$. This is exactly the classic "optimal" nonparametric rate for estimating $g'(\tau_0)$ when $g \in C^2[0,1]$. Note that it is possible to estimate τ_0 better than β_0.

4. Proofs of Main Results. We begin with several preliminary definitions and results. Let

$$\psi(t) = t_+ - \int_{-2}^{2} K(t-v)v_+ dv, \tag{6}$$

where $t_+ = \max\{0, t\}$. Note that $\psi(t) = 0$ for $|t| > 1$ by (A3) and $\psi(-t) = \psi(t)$ from the symmetry of K. We also tacitly assume that h is small enough so that $(\tau - 4h, \tau + 4h) \in (0,1)$ for all $y \in (-A, A)$. Expression (6) can be simplified somewhat by defining

$$(IK)_j(t) = \int_{-1}^{t} u^j K(u) du.$$

It follows that

$$\psi(t) = t\,[1 - (IK)_0(t)] + (IK)_1(t)$$

for $t > 0$.

We will also need a related kernel defined as $\tilde{K} = 2K - K * K$. This new kernel is the result of "twicing" (Stützle and Mittal, 1979) and is a *fourth order* symmetric kernel with support $[-2, 2]$, i.e.,

$$\int_{-2}^{2} u^j \tilde{K}(u) du = 0, j = 0, 1, 2, 3, \int_{-2}^{2} u^4 \tilde{K}(u) du \neq 0.$$

Let

$$\tilde{\psi}(t) = t_+ - \int_{-4}^{4} \tilde{K}(t-v)v_+ dv, \tag{7}$$

and define

$$(I\tilde{K})_j(t) = \int_{-2}^{t} u^j \tilde{K}(u) du.$$

It can be shown that $\tilde{\psi}$ is symmetric with support $[-2, 2]$, and

$$\tilde{\psi}(t) = t\big[1 - (I\tilde{K})_0(t)\big] + (I\tilde{K})_1(t) \tag{8}$$

for $t > 0$. We further assume that K is chosen so that

$$\int_{0}^{2} u\tilde{K}(u) du \neq 0.$$

Several facts about $\tilde{\psi}$ are collected in the following lemma. The proof is routine and is omitted.

LEMMA 1. *Under the conditions above, $\tilde{\psi}$ is piecewise twice differentiable,*

$$\tilde{\psi}''(t) = \tilde{K}(t), \quad t \neq 0, \tag{9a}$$

$$\int_{-2}^{2} \tilde{\psi}(u)du = \int_{-2}^{2} u\tilde{\psi}(u)du = \int_{-2}^{2} u\tilde{\psi}'(u)du = 0, \tag{9b}$$

$$\int_{-2}^{2} u^2\tilde{\psi}(u)du = -\frac{1}{12}\int_{-2}^{2} u^4\tilde{K}(u)du \neq 0, \tag{9c}$$

and there exists a constant $M < \infty$ such that

$$|\tilde{\psi}(u-w) - \tilde{\psi}(u)| \leq M|w| \tag{9d}$$

for all u, w.

LEMMA 2.

$$(I-S)\phi_\tau = h(0,\ldots,\psi(q_i),\ldots,\psi(q_{i+r}),\ldots,0)' + O\left(\frac{1}{nh}\right),$$

where $q_j = \frac{t_j-\tau}{h}$, i is the first index such that $q_i > -1$ and $r = [2nh]$.

PROOF. Up to quadrature error, $\psi(q_j) = 0$ for all $|t_j - \tau| > h$ by the second order assumptions (A2) and (A3). The result follows from standard arguments and a substitution. ∎

LEMMA 3.

$$(I-S)^2\phi_\tau = h(0,\ldots,\tilde{\psi}(q_i),\ldots,\tilde{\psi}(q_{i+r}),\ldots,0)' + O\left(\frac{1}{nh}\right),$$

where q_j is defined as in Lemma 2, i is the first index such that $q_i > -2$ and $r = [4nh]$.

PROOF. It can be shown that $2S - S^2$ has typical element

$$\frac{1}{nh}\tilde{K}\left(\frac{t_i-t_j}{h}\right) + O\left(\frac{1}{(nh)^2}\right).$$

Because $\tilde{K}$ is also a kernel, the result follows as in the last lemma. ∎

LEMMA 4. *$\phi_\tau'(I-S)^2\phi_\tau \sim B_1 nh^3$ for*

$$B_1 = 2\int_0^1 \psi(t)^2 dt.$$

PROOF. Using Lemma 2,

$$\begin{aligned}\phi_\tau'(I-S)^2\phi_\tau &\sim h^2 \sum_{q_j\in[-1,1]} \psi^2(q_j)\\ &= nh^3\int_{-1}^{1}\psi^2(u)du + O(h^3) + O\left(\frac{1}{nh^2}\right).\end{aligned}$$

∎

The following result is similar, and the proof is omitted.

LEMMA 5. *$\phi_\tau'(I-S)^4\phi_\tau \sim \tilde{B}_1 nh^3$ for*

$$\tilde{B}_1 = 2\int_0^2 \tilde{\psi}(t)^2 dt.$$

In the proof of the next lemma, we will use the fact that the assumptions on f permit the decomposition $f(t) = f_0(t) + f_1(t)$ with

$$f_1(t) = \Delta_1(t-\tau_0)_+^2/2$$

for $\Delta_1 = f''(\tau_0^+) - f''(\tau_0^-)$. Here $(t-\tau_0)_+^2 = (t-\tau_0)^2$ for $t \geq \tau_0$ and zero otherwise. It follows that $f_0(t) \in C^2[0,1]$ and $f_0''(\tau_0^+) = f''(\tau_0^-)$.

LEMMA 6.

$$\phi_\tau'(I-S)^2 f \sim \begin{cases} B_2 nh^4 f''(\tau), & \tau \neq \tau_0\\ B_2 nh^4 \Delta_2, & \tau = \tau_0\end{cases}$$

where

$$B_2 = -\frac{1}{24}\int_{-2}^{2} u^4 \tilde{K}(u)du$$

and $\Delta_2 = (f''(\tau_0^+) + f''(\tau_0^-))/2$.

PROOF. Assume first that f has two continuous derivatives at τ. By Lemma 3 with $q_i = (t_i - \tau)/h$ and a Taylor series expansion,

$$\begin{aligned}\phi_\tau'(I-S)^2 f &\sim h \sum_{t_i\in[\tau-2h,\tau+2h]} \tilde{\psi}(q_i) f(t_i)\\ &\sim nh\int_{\tau-2h}^{\tau+2h} \tilde{\psi}\left(\frac{t-\tau}{h}\right) f(t)dt\\ &\sim nh^2\int_{-2}^{2}\tilde{\psi}(u)\big(f(\tau) + huf'(\tau) + \frac{(hu)^2}{2}f''(\tau)\big)du\\ &= B_2 f''(\tau) nh^4,\end{aligned}$$

using the moment conditions (9b) and (9c) and finally (9a). This proves the Lemma for $\tau \neq \tau_0$.

Consider now the case $\tau = \tau_0$. Arguing as above,

$$\phi'_{\tau_0}(I-S)^2 f_1 \sim \frac{\Delta_1 h}{2} \sum_{t_i \in [\tau_0, \tau_0+2h]} \tilde{\psi}(q_i)(t_i - \tau_0)^2$$
$$\sim \frac{\Delta_1}{2} nh^4 \int_0^2 u^2 \tilde{\psi}(u) du.$$

But by the symmetry of $\tilde{\psi}$, this last expression is equal to

$$\frac{\Delta_1}{4} nh^4 \int_{-2}^2 u^2 \tilde{\psi}(u) du = B_2 \Delta_1 nh^4/2.$$

Thus for $\tau = \tau_0$,

$$\begin{aligned} \phi'_{\tau_0}(I-S)^2 f &= \phi'_{\tau_0}(I-S)^2(f_0 + f_1) \\ &\sim B_2 nh^4 (f''(\tau_0^-) + \Delta_1/2) \\ &= B_2 nh^4 \Delta_2. \end{aligned}$$

∎

LEMMA 7.

$$(\phi_\tau - \phi_{\tau_0})'(I-S)^2 f \sim y B_3 \Delta_1 \frac{nh^4}{\sqrt{nh^3}}.$$

with

$$B_3 = -\frac{1}{6} \int_0^2 u^3 \tilde{K}(u) du.$$

PROOF. We again use the decomposition $f = f_0 + f_1$. Consider f_1 first and recall that $\tau = \tau_0 + hy/\sqrt{nh^3}$. Without loss of generality, take $y > 0$ and apply Lemma 3 to obtain

$$(\phi_\tau - \phi_{\tau_0})'(I-S)^2 f_1 \sim \frac{h\Delta_1}{2} \sum_{t_i \in [\tau_0, \tau+2h]} \left(\tilde{\psi}\left(\frac{t_i - \tau}{h}\right) - \tilde{\psi}\left(\frac{t - \tau_0}{h}\right) \right) (t_i - \tau_0)^2$$
$$\sim \frac{nh\Delta_1}{2} \int_{\tau_0}^{\tau+2h} \left(\tilde{\psi}\left(\frac{t-\tau}{h}\right) - \tilde{\psi}\left(\frac{t-\tau_0}{h}\right) \right) (t-\tau_0)^2 dt$$
$$= \frac{nh^4 \Delta_1}{2} \int_0^{2+y/\sqrt{nh^3}} \left(\tilde{\psi}(u - y/\sqrt{nh^3}) - \tilde{\psi}(u) \right) u^2 du.$$

Now let

$$H_1(w) = \int_0^{2+w} \left(\tilde{\psi}(u-w) - \tilde{\psi}(u) \right) u^2 du.$$

Using the fact that $\tilde{\psi}(2) = 0$ and $\tilde{\psi}$ is continuous and piecewise continuously differentiable by (7), $H_1(0) = 0$ and $H_1'(0) = -\int_0^2 u^2 \tilde{\psi}'(u) du$. Integration by parts using (9a) shows that $H_1'(0) = 2B_3$. Thus $H_1(w) \sim 2wB_3$, and with $w = y/\sqrt{nh^3}$ we obtain

$$(\phi_\tau - \phi_{\tau_0})'(I-S)^2 f_1 \sim yB_3\Delta_1 nh^4/\sqrt{nh^3}.$$

The other term is handled similarly. Using a second order Taylor expansion for f_0 at τ_0 and (9b), it can be shown that $(\phi_\tau - \phi_{\tau_0})'(I-S)^2 f_0 = o(ynh^4/\sqrt{nh^3})$, hence the contribution from f_0 is negligible. ∎

LEMMA 8.

$$(\phi_\tau - \phi_{\tau_0})'(I-S)^2\phi_{\tau_0} \sim -B_4 y^2$$

for

$$B_4 = \frac{1}{2}\int_0^2 u\tilde{K}(u)du.$$

PROOF. Arguing as above,

$$(\phi_\tau - \phi_{\tau_0})'(I-S)^2\phi_{\tau_0} \sim nh^3 \int_0^{2+y/\sqrt{nh^3}} [\tilde{\psi}(u - y/\sqrt{nh^3}) - \tilde{\psi}(u)]\,udu.$$

Letting

$$H_3(w) = \int_0^{2+w} [\tilde{\psi}(u-w) - \tilde{\psi}(u)]\,udu,$$

Lemma 1 implies that $H_3(0) = H_3'(0) = -\int_0^2 u\tilde{\psi}'(u)du = 0$. It can be shown that H_3' is differentiable at zero with $H_3''(0) = \int_0^2 u\tilde{\psi}''(u)du = -\int_0^2 u\tilde{K}(u)du$, so $H_3(y/\sqrt{nh^3}) \sim -B_4 y^2/(nh^3)$. ∎

LEMMA 9. *Let* $\eta = \tau_0 + hz/\sqrt{nh^3}$. *Then*

$$(\phi_\tau - \phi_{\tau_0})'(I-S)^4(\phi_\eta - \phi_{\tau_0}) \sim yzB_5$$

with

$$\begin{aligned} B_5 &= \int_0^2 \left[\int_s^2 \tilde{K}(y)dy\right]^2 ds \\ &= 2\int_0^2\int_0^2 s(y-s)_+\tilde{K}(s)\tilde{K}(y)dsdy. \end{aligned}$$

PROOF. Once again, by Lemma 3

$$\begin{aligned} &(\phi_\tau - \phi_{\tau_0})'(I-S)^4(\phi_\eta - \phi_{\tau_0}) \\ &\sim nh^2 \sum \left(\tilde{\psi}\left(\frac{t_i-\tau}{h}\right) - \tilde{\psi}\left(\frac{t_i-\tau_0}{h}\right)\right)\left(\tilde{\psi}\left(\frac{t_i-\eta}{h}\right) - \tilde{\psi}\left(\frac{t_i-\tau_0}{h}\right)\right) \\ &\sim nh^3 \int \left(\tilde{\psi}(u-y/\sqrt{nh^3}) - \tilde{\psi}(u)\right)\left(\tilde{\psi}(u-z/\sqrt{nh^3}) - \tilde{\psi}(u)\right)du \\ &= G(y/\sqrt{nh^3}, z/\sqrt{nh^3}), \end{aligned}$$

where

$$G(v,w) = \int \left(\tilde{\psi}(u-v) - \tilde{\psi}(u)\right)\left(\tilde{\psi}(u-w) - \tilde{\psi}(u)\right) du.$$

It can be shown that

$$\frac{\partial G}{\partial v}(0,0) = \frac{\partial^2 G}{\partial v^2}(0,0) = \frac{\partial G}{\partial w}(0,0) = \frac{\partial^2 G}{\partial w^2}(0,0) = 0,$$

$$\frac{\partial^2 G}{\partial v \partial w}(0,0) = \int_{-2}^{2} \tilde{\psi}'(u)^2 du = B_5 \neq 0.$$

Thus $G(y/\sqrt{nh^3}, z/\sqrt{nh^3}) \sim B_5 yz/(nh^3)$. ∎

LEMMA 10. *Let $\ell = (\ell_1, \ldots, \ell_p)'$ be an arbitrary vector of length p, and let $y_j, j = 1, \ldots, p$, be arbitrary elements of $(-A, A)$. Then, $V_n = \sum_{j=1}^{p} \ell_j(\phi_{\tau_j} - \phi_{\tau_0})'(I-S)^2\varepsilon \xrightarrow{\mathcal{D}} N(0, \ell'\Sigma\ell)$ with $\Sigma = \{y_i y_j B_5\}_{i,j=1,p}$, for B_5 in Lemma 9 and $\tau_j = \tau_0 + y_j h/\sqrt{nh^3}$.*

PROOF. We have $E(V_n) = 0$ and

$$\begin{aligned}\mathrm{Var}(V_n) &= \sigma^2 \sum_j \sum_k \ell_j \ell_k (\phi_{\tau_j} - \phi_{\tau_0})'(I-S)^4(\phi_{\tau_k} - \phi_{\tau_0}) \\ &\sim \sigma^2 \ell'\Sigma\ell.\end{aligned}$$

From Lemma 3

$$\begin{aligned}V_n = \sum_{i=1}^{n} \varepsilon_i c_i &= \sum_{i=1}^{n} \varepsilon_i \sum_{j=1}^{p} \ell_j \left[(\phi_{\tau_j} - \phi_{\tau_0})'(I-S)^2\right]_j \\ &= h \sum_{t_i \in T} \varepsilon_i \sum_{j=1}^{p} \ell_j \left[\tilde{\psi}\left(\frac{t_i - \tau_j}{h}\right) - \tilde{\psi}\left(\frac{t_i - \tau_0}{h}\right)\right] + \sum_{t_i \notin T} \varepsilon_i c_i,\end{aligned}$$

with summation in the first term above over the range

$$T = \left[\tau_0 - h(2 + A/\sqrt{nh^3}), \tau_0 + h(2 + A/\sqrt{nh^3})\right]$$

and c_i uniformly $O((nh)^{-1})$ in the last sum.

Let $q_i = (t_i - \tau_0)/h$. Then, for $\tau = \tau_0 + hy/\sqrt{nh^3}$, Lemma 1 yields

$$\left|\tilde{\psi}(\frac{t_i - \tau}{h}) - \tilde{\psi}(\frac{t_i - \tau_0}{h})\right| = |\tilde{\psi}(q_i + y/\sqrt{nh^3}) - \tilde{\psi}(q_i)| \leq M|y/\sqrt{nh^3}|.$$

It follows that the terms in the sum for $t_i \in T$ have coefficients c_i of order $h/\sqrt{nh^3} = 1/\sqrt{nh} = O(n^{-2/5})$ if $h \sim n^{-\alpha}$ for $\alpha \geq 1/5$. Consequently, for any

$\theta > 0$, if the ε_i have common distribution function F_ε,

$$\sum_i \int_{|c_i x|>\theta} (c_i x)^2 dF_\varepsilon(x) \lesssim \ell' \Sigma \ell \int_{\max|c_i||x_i|>\theta} x^2 dF_\varepsilon(x)$$
$$= \ell' \Sigma \ell \int_{|x|>c\theta\sqrt{nh^3}/h} x^2 dF_\varepsilon(x) \to 0,$$

and the Lindeberg condition holds. ∎

PROOF OF THEOREM 1. The mean, covariance and joint asymptotic normality conditions have been established in Lemmas 7–10. Thus to prove weak convergence, we need only establish tightness. For this it suffices to show that

$$E(C_n(y_1) - C_n(y_2))^2 \le c(y_1 - y_2)^2$$

for some constant c and all n sufficiently large (Billingsley, 1968). But

$$\begin{aligned} E(C_n(y_1) - C_n(y_2))^2 &= \mathrm{Var}(C_n(y_1)) - 2\mathrm{Cov}(C_n(y_1), C_n(y_2)) + \mathrm{Var}(C_n(y_2)) \\ &\sim B_5(y_1 - y_2)^2. \end{aligned}$$

The approximation is uniform for $y \in [-A, A]$. ∎

PROOF OF THEOREM 3. To prove (5a), begin by writing

$$\begin{aligned} \sqrt{nh^3}\left(\hat\beta(\tau_0) - \beta_0\right) &= \sqrt{nh^3}\left(\frac{\phi'_{\tau_0}(I-S)^2 z}{\phi'_{\tau_0}(I-S)^2\phi_{\tau_0}} - \beta_0\right) \\ &= \sqrt{nh^3}\left(\frac{\phi'_{\tau_0}(I-S)^2 f}{\phi'_{\tau_0}(I-S)^2\phi_{\tau_0}} + \frac{\phi'_{\tau_0}(I-S)^2 \varepsilon}{\phi'_{\tau_0}(I-S)^2\phi_{\tau_0}}\right). \end{aligned}$$

The first term is asymptotic to

$$\frac{nh^4 B_2 \Delta_2}{\sqrt{nh^3} B_1}$$

by Lemmas 4 and 6, establishing the asymptotic mean. The second term is asymptotically equivalent to $\phi'_{\tau_0}(I-S)^2\varepsilon/(B_1\sqrt{nh^3})$ by Slutsky's theorem and Lemma 2 again, so it has asymptotic variance $\tilde B_1/B_1^2$. For normality, we have

$$\frac{1}{\sqrt{nh^3}}\phi'_{\tau_0}(I-S)^2\varepsilon = \frac{1}{\sqrt{nh^3}} \sum_{t_i\in[\tau_0-h,\tau_0+h]} c_i\varepsilon_i + \frac{1}{\sqrt{nh^3}} \sum_{t_i\notin[\tau_0-h,\tau_0+h]} c_i\varepsilon_i.$$

For $t_i \in [\tau_0 - h, \tau_0 + h]$, $c_i \doteq h\psi(q_i)$ with the other $c_i = O((nh)^{-1})$. An application of the Lindeberg condition gives the result.

It remains to establish (5b). Theorem 1 and the functional mapping theorem imply that

$$nh^3\left(\hat\beta(\hat\tau) - \hat\beta(\tau_0)\right) \xrightarrow{\mathcal{D}} y^*$$

with y^* defined by (4). But this shows that $\sqrt{nh^3}(\hat{\beta}(\hat{\tau}) - \hat{\beta}(\tau_0)) = o_p(1)$, and the proof is complete. ∎

Acknowledgement. The authors are indebted to Professor Hans-Georg Müller for sharing his work prior to publication and to two anonymous referees for their helpful comments and suggestions. Eubank's research was supported under NSF Grant No. DMS-9024879.

REFERENCES

BILLINGSLEY, P. (1968), *Convergence of Probability Measures*, Wiley: New York.

EDDY, W. (1980), "Optimum kernel estimators of the mode", *Annals of Statistics* **8**, 870–882.

ENGLE, R., GRANGER, C., RICE, J., and WEISS, A. (1986), "Nonparametric estimates of the relation between weather and electricity sales," *J. Amer. Statist. Assoc.* **81**, 310–320.

EUBANK, R. (1988), *Spline Smoothing and Nonparametric Regression*, Marcel Dekker: New York.

EUBANK, R. and SPECKMAN, P. (1991), "A bias reduction theorem with applications in nonparametric regression," *Scand. J. Statist.* **18**, 211–222.

MÜLLER, H.-G. (1992), "Change-points in nonparametric regression analysis," *Ann. Statist.* **20**, 737–761.

QIU, P. (1992), "Nonparametric estimation of the jump surfaces." Unpublished manuscript.

STÜTZLE, W. and MITTAL, Y. (1979), "Some comments on the asymptotic behavior of robust smoothers," in *Smoothing Techniques for Curve Estimation*, Th. Gasser and M. Rosenblatt, eds, 191–195. Lecture Notes in Mathematics, No. 757. Heidelberg: Springer-Verlag.

WAHBA, G. (1984), "Partial spline models for the semiparametric estimation of functions of several variables," in *Analyses for Time Series, Japan–US Joint Seminar*, pp. 319–329. Tokyo: Institute of Statistical Mathematics.

WHITT, W. (1970), "Weak convergence of probability measures on the function space $C([0, \infty])$," *Ann. Math. Statist.* **41**, 939–944.

WU, J. S. and CHU, C. K. (1992), "Kernel type estimators of jump points and values of a regression function," to appear in *Ann. Statist.*

DEPARTMENT OF STATISTICS
TEXAS A&M UNIVERSITY
COLLEGE STATION, TX77843-3143

DEPARTMENT OF STATISTICS
UNIVERSITY OF MISSOURI
COLUMBIA, MO 65211