Asymptotic Theory for Bent-Cable Regression — the Basic Case

Grace Chiu,[1,2] Richard Lockhart[2], and Richard Routledge[2]
[1] Pacific Institute for the Mathematical Sciences
[2] Department of Statistics and Actuarial Science
Simon Fraser University, Burnaby, British Columbia, V5A 1S6, Canada.

Article also available from *Journal of Statistical Planning and Inference* Volume **127**, Issues 1–2, 1 January 2005, Pages 143–156. doi:10.1016/j.jspi.2003.09.007

Abstract

We use what we call the bent-cable model to describe potential change-point phenomena. The class of bent cables includes the commonly used broken stick (a bent cable without a bend segment). Theory for least-squares (LS) estimation is developed for the basic bent cable, whose incoming and outgoing linear phases have slopes 0 and 1, respectively, and are joined smoothly by a quadratic bend. Conditions on the design are given to ensure regularity of the estimation problem, despite non-differentiability of the model's first partial derivatives (with respect to the covariate and model parameters). Under such conditions, we show that the LS estimators (i) are consistent, regardless of a zero or positive true bend width; and (ii) asymptotically follow a bivariate normal distribution, if the underlying cable has all three segments. In the latter case, we show that the deviance statistic has

[1]Corresponding author.

E-mail addresses: *grace@pims.math.ca* (G. Chiu), *lockhart@stat.sfu.ca* (R. Lockhart), *routledg@stat.sfu.ca* (R. Routledge).

This research has been funded by the Natural Sciences and Engineering Research Council of Canada through a Postgraduate Scholarship to G. Chiu during her graduate studies at SFU, and Research Grants to R. Lockhart and R. Routledge.

an asymptotic chi-squared distribution with two degrees of freedom.

Mathematical subject codes: 62F12 (62J02)

Keywords: Change points; Segmented regression; Least-squares Estimation; Asymptotic Theory

---

1 Introduction

Given known design points $x_1, \ldots, x_n$, we consider random responses $Y_1, \ldots, Y_n$ generated by the regression model

$$Y_i = q(x_i; \tau_0, \gamma_0) + \varepsilon_i\,, \qquad i = 1, \ldots, n\,, \qquad \text{where} \tag{1}$$

$$q(x; \tau, \gamma) = \frac{(x - \tau + \gamma)^2}{4\gamma}\,\mathbf{1}\big\{|x - \tau| \le \gamma\big\} + (x - \tau)\,\mathbf{1}\big\{x - \tau > \gamma\big\} \tag{2}$$

is referred to as the *basic bent cable* (Figure 1), and ε_i's are i.i.d. random errors with mean 0 and known, constant standard deviation σ. [2] We write $\boldsymbol{\theta}_0 = (\tau_0, \gamma_0)$.

Least-squares (LS) estimation of $\boldsymbol{\theta}_0 \in \Omega = (-\infty, M] \times [0, \infty)$ on the open regression domain $\mathcal{X} = \mathbb{R}$ is considered. Here, M is some finite positive upper bound (large) for the candidate τ-values. Zero is the natural lower bound for candidate γ-values. Any basic bent cable $q(x; \boldsymbol{\theta})$ for $\boldsymbol{\theta} \in \Omega$ is a candidate model. In this article, we prove, given a set of conditions on the location of the design points $x_1, \ldots, x_n$ (Section 3), that the least-squares estimators (LSE's) for τ_0 and γ_0 are consistent when $\gamma_0 \ge 0$, and asymptotically follow a bivariate normal distribution when $\gamma_0 > 0$. Asymptotic distributional properties for the case of $\gamma_0 = 0$ appear in Chiu *et al.* (2002a). A bent cable with free slope parameters is required

[2] In practice, estimation of σ may be required. Chiu (2002) shows that the results of this article extend to LS estimation assuming unknown σ, and that the LSE of σ is consistent.

in practice. The full bent-cable model can be written as $f(x;\beta_0,\beta_1,\beta_2,\tau,\gamma) = \beta_0 + \beta_1\ x + \beta_2\ q(x;\tau,\gamma)$. This article is intended to provide a framework for the complex estimation theory associated with the full model.

Seber and Wild (1989, Chapter 9) have suggested employing the class of bent-cable models — which includes the piecewise-linear "broken-stick" model when γ=0 — in situations where both smooth and sharp transitions are plausible. However, modeling change phenomena by the broken stick remains common (Barrowman and Myers, 2000; Naylor and Su, 1998; Neuman *et al.*, 2001). Numerical instability due to the non-differentiability of this model prompted Tishler and Zang (1981) to develop (2). Their introduction of a "phoney" bend of fixed, non-trivial width γ to replace the kink at τ was a computational tactic. Upon numerical convergence, γ would be ignored.

However, when no law of nature or auxiliary knowledge is available to support an abruptness notion, a broken-stick fit would encourage the investigator to look for sources of change associated with the sole value of $\widehat{\tau}$. In contrast, the bent cable incorporates γ as part of the parametric model. It generalizes the broken stick by removing the *a priori* assumption of a sharp threshold, allowing for a possibly gradual transition. A bent-cable fit would point to one or more sources of change whose influence took hold gradually over a certain covariate range. Thus, it helps to avoid data misinterpretation due to possible over-simplication of the nature of change. We call it the "bent cable" due to the smooth bend as opposed to a sharp break in a snapped stick. The performance of bent-cable regression for assessing the abruptness of change is discussed in Chiu *et al.* (2002b).

2 Theory for Segmented Models

Whether γ is zero or positive, the second partial derivatives of (2) fail to exist where the segments meet. This irregularity prevents the direct application of asymptotics readily available for LS problems that are regular. Feder (1975) and Ivanov (1997), among others, have discussed asymptotics for non-linear regression involving segmented models of general forms with unknown join points. Similar regression problems include the consideration of broken-stick models (Bhattacharya, 1990; Hušková, 1998) and general two-phase linear-nonlinear or other multiphase non-linear models (Gallant, 1974 & 1975; Hušková and Steinebach, 2000; Jarušková, 1998a,b & 2001; Rukhin and Vajda, 1997).

Feder's principal assumption is continuity of the underlying function without extra smoothness constraints. If the underlying model has an odd order of smoothness (number of continuous derivatives plus one), then the asymptotics are radically different compared to those for an even order. Gallant considers once-differentiable candidate and underlying models of common functional form. In the case of a once-differentiable quadratic-quadratic model, consistency and asymptotic normality of the LSE for the unknown joint is established by, among other requirements, (i) taking the design points to be "near" replicates of a basic design with five distinct covariate values (see Gallant, 1975, p. 26; Gallant, 1974, pp. 6–7), and (ii) restricting the search for candidate knots to within a compact subset trapped between two consecutive covariate values from (i) (see Gallant, 1975, p. 26). Gallant, Ivanov, and Rukhin and Vajda all establish consistency by assuming a bounded or compact parameter space. Ivanov adds a somewhat unintuitive condition which relates the candidate model to the response error variability $\big($see Ivanov, 1997, p. 30, expression (3.13)$\big)$. Rukhin and Vajda simply exclude the LSE from their M-estimators $\big($see condition (C3) in their article$\big)$. To establish asymptotic normality, they and Ivanov assume twice-differentiability of the regression function, although Rukhin and Vajda cite a tactic for relaxing such an assumption (see Vaart and Wellner, 1996, Chapter 3).

Hušková, Hušková and Steinebach, Jarušková, and Rukhin and Vajda all consider evenly-spaced regressors, while Bhattacharya considers a more general design. These authors have all shown standard LS asymptotics despite a lack of higher order derivatives of the regression function. However, their results are not directly applicable to our specialized problem. We consider underlying and candidate models within the class of bent cables. General and structurally simple design conditions are provided and proved to suffice for establishing regularity.

3 Conditions for the Basic Bent Cable

Besides an unspecified upper bound on the τ-space, our only regularity conditions are placed on the design. Given $\delta > 0$ and sequence $\xi_n \downarrow 0$, first define

$$c_0(\delta)=\liminf_{n\longrightarrow\infty} \frac{1}{n}\sum \mathbf{1}\Big\{|x_i-\tau_0| \le \gamma_0+\delta\Big\},\ c_-(\delta)=\liminf_{n\longrightarrow\infty} \frac{1}{n}\sum \mathbf{1}\Big\{|x_i-\tau_0| \le \gamma_0-\delta\Big\},$$

$$c_+(\delta)=\liminf_{n\longrightarrow\infty} \frac{1}{n}\sum \mathbf{1}\Big\{x_i \ge \tau_0+\gamma_0+\delta\Big\},\ \zeta_n(\xi_n)=\frac{1}{n}\sum \mathbf{1}\left\{\Big||x_i-\tau_0|-\gamma_0\Big| \le \xi_n\right\}.$$

The regularity conditions on $x_1,\ldots,x_n$ are

[A_1] If $\gamma_0 = 0$, then $\forall\ \delta > 0$, $c_0(\delta) > 0$.

[A_2] If $\gamma_0 > 0$, then $\exists\ \delta_1 > 0$ such that $c_-^* \equiv c_-(\delta_1) > 0$.

[B] For $\gamma_0 \ge 0$, $\exists\ \delta_{11} > 0$ such that $c_+^* \equiv c_+(\delta_{11}) > 0$.

[C] If $\gamma_0 > 0$, then $\forall\ \xi_n \downarrow 0$, $\zeta_n(\xi_n) \longrightarrow 0$.

[D] If $\gamma_0 > 0$, then $x_i \neq \tau_0 \pm \gamma_0 \; \forall \; i = 1, \ldots, n$.

The practical value in these conditions is that they indicate a general design (not necessarily equidistant) which ensures that data are collected at appropriate x-locations for reliable LS estimation of $\boldsymbol{\theta}_0$. While $\boldsymbol{\theta}_0$ is unknown, the investigator's expertise in the subject matter should suggest a range of design points that easily satisfies conditions [A] and [B]. Satisfying condition [C] in practice is not difficult as the precise locations of $\tau_0 \pm \gamma_0$ are unknown. For a continuous covariate, it is reasonable to rule out exact equality that would violate Condition [D]. At the cost of some notational complexity, [D] could be entirely eliminated.

4 Method and Results

For normally distributed ε_i's in (1), the log-likelihood as a function of $\boldsymbol{\theta}$ is

$$\ell_n(\boldsymbol{\theta}) = -\frac{1}{2\sigma^2} S_n(\boldsymbol{\theta}) + \text{constant}\,,$$

where $S_n(\boldsymbol{\theta}) = \sum_{i=1}^n \big[Y_i - q(x_i;\boldsymbol{\theta})\big]^2$. In this case, maximum likelihood (ML) and LS theories coincide. In the absence of normality, we maximize ℓ_n but do not refer to it as the ML function. Maximization of ℓ_n is equivalent to minimization of S_n (i.e. LS estimation). The results in this section imply that ML and LS asymptotics coincide even without the normality of the ε_i's in the latter case. When a sample yields multiple maximizers of ℓ_n, we take the LSE $\widehat{\boldsymbol{\theta}}_n$ to be the one selected sequentially as follows: (i) pick out the one(s) with the least vector norm; (ii) keep the one(s) with the least γ; (iii) if necessary, select that with the least $|\tau|$.

The following notation is used in stating the results:

$$U_{n\tau}(\boldsymbol{\theta}) = \frac{\partial \ell_n}{\partial \tau}(\boldsymbol{\theta}) \ , \quad U_{n\gamma}(\boldsymbol{\theta}) = \frac{\partial \ell_n}{\partial \gamma}(\boldsymbol{\theta}) \quad , \quad \boldsymbol{U}_n(\boldsymbol{\theta}) = \big(U_{n\tau}(\boldsymbol{\theta}), U_{n\gamma}(\boldsymbol{\theta})\big)^T \ ,$$

$$\mathbb{V}_n(\boldsymbol{\theta}) = \nabla \boldsymbol{U}_n(\boldsymbol{\theta}) \quad \text{(wherever defined)} \quad , \quad \mathbb{I}_n(\boldsymbol{\theta}) = \mathrm{Cov}_{\boldsymbol{\theta}_0}\big[\boldsymbol{U}_n(\boldsymbol{\theta})\big] \ .$$

Note that $\mathbb{I}_n$ and $\boldsymbol{U}_n$ are analogous to the Fisher Information and the score function in ML estimation. With conditions [A] to [D], our only irregularity lies in a Hessian, $\mathbb{V}_n$, that is not well-defined everywhere. However, [D] preserves regularity at $\boldsymbol{\theta}_0$, and is used to establish the uniform convergence (in probability) of $-\mathbb{V}_n$ and $\mathbb{I}_n$ in a shrinking neighborhood of $\boldsymbol{\theta}_0$ (Lemma 3). Regular LS asymptotics follow.

Condensed proofs for the following key results appear in Section 6. Motivational and mathematical details are in Chiu (2002).

Lemma 1 (Identifiability) *Take δ_1 and δ_{11} from [A] and [B]. Suppose $w_1 \in [\tau_0 + \gamma_0 + \delta_{11}, \infty)$. Furthermore, (1) if $\gamma_0 > 0$ and $w_0 \in [\tau_0 - \gamma_0 + \delta_1, \tau_0 + \gamma_0 - \delta_1]$, then $q(w_i\,; \boldsymbol{\theta}_0) = q(w_i\,; \boldsymbol{\theta})$ for $i = 1, 2$ implies $\boldsymbol{\theta} = \boldsymbol{\theta}_0$; (2) if $\gamma_0 = 0$ and $w_0 \in [\tau_0 - \delta/8, \tau_0 + \delta/8]$, where $\delta \in (0, \delta_{11}/2)$, then $q(w_i\,; \boldsymbol{\theta}_0) = q(w_i\,; \boldsymbol{\theta})$ for $i = 1, 2$ implies $|\boldsymbol{\theta} - \boldsymbol{\theta}_0| \le \delta$.*

Remark 1. This lemma is fundamental to establishing consistency.

Theorem 1 (Consistency) *Consider LS estimation for model (1) as defined in Section 1. Under conditions [A] and [B], $\widehat{\boldsymbol{\theta}}_n \xrightarrow{\mathrm{P}} \boldsymbol{\theta}_0$ as $n \to \infty$.*

Before stating our next theorem, we provide a fact about the "square-root" matrix to avoid ambiguity. For a symmetric positive definite matrix $\mathbb{A}$, there is a unique symmetric positive definite matrix $\mathbb{B}$ such that $\mathbb{A} = \mathbb{B}^2 = \mathbb{B}^T\mathbb{B}$ (Golub and Van Loan, 1996, Exercise P11.2.4.). We write $\mathbb{B} = \mathbb{A}^{\frac{1}{2}}$, the "square root" of $\mathbb{A}$.

Theorem 2 (Asymptotic Normality) *Consider LS estimation for model (1) with as defined in Section 1. Under conditions [A] to [D] and* $\gamma_0 > 0$,

1. $\frac{1}{n}\mathbb{I}_n(\boldsymbol{\theta}_0)$ *is positive definite for all sufficiently large* n;
2. $\sqrt{n}\left[\frac{1}{n}\mathbb{I}_n(\boldsymbol{\theta}_0)\right]^{\frac{1}{2}}\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\right)$ *converges in distribution to a standard bivariate normal random variable;*
3. $P_{\boldsymbol{\theta}_0}\left\{\frac{1}{n}\mathbb{I}_n\left(\widehat{\boldsymbol{\theta}}_n\right) \text{is positive definite}\right\} \longrightarrow 1$;
4. $\sqrt{n}\left[\frac{1}{n}\mathbb{I}_n\left(\widehat{\boldsymbol{\theta}}_n\right)\right]^{\frac{1}{2}}\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\right)$ *converges in distribution to a standard bivariate normal random variable;*
5. $G_n \equiv 2\left[\ell_n\left(\widehat{\boldsymbol{\theta}}_n\right) - \ell_n\left(\boldsymbol{\theta}_0\right)\right]$, *the deviance statistic, converges in distribution to a* χ^2 *random variable with 2 degrees of freedom.*

5 Auxiliary Lemmas for Proving Theorem 2

First, note that the ith summands of $U_{n\tau}(\boldsymbol{\theta})$ and $U_{n\gamma}(\boldsymbol{\theta})$ are, respectively,

$$U_{\tau,i}(\boldsymbol{\theta}) = -\frac{1}{\sigma^2}\left[\varepsilon_i + d_{\boldsymbol{\theta}_0,\boldsymbol{\theta}}(x_i)\right]\frac{\partial q_{\boldsymbol{\theta}}(x_i)}{\partial \tau} \quad , \quad U_{\gamma,i}(\boldsymbol{\theta}) = -\frac{1}{\sigma^2}\left[\varepsilon_i + d_{\boldsymbol{\theta}_0,\boldsymbol{\theta}}(x_i)\right]\frac{\partial q_{\boldsymbol{\theta}}(x_i)}{\partial \gamma}$$

$$\begin{aligned} \text{where} \quad d_{\boldsymbol{\theta}_1,\boldsymbol{\theta}_2}(x) &= q(x;\boldsymbol{\theta}_1) - q(x;\boldsymbol{\theta}_2), \\ \frac{\partial q_{\boldsymbol{\theta}}(x_i)}{\partial \tau} &= -\frac{x_i - (\tau - \gamma)}{2\gamma}\mathbf{1}\big\{|x_i - \tau| \le \gamma\big\} - \mathbf{1}\{x_i > \tau + \gamma\}, \qquad (3) \\ \frac{\partial q_{\boldsymbol{\theta}}(x_i)}{\partial \gamma} &= \frac{1}{4}\left[1 - \left|\frac{x_i - \tau}{\gamma}\right|^2\right]\mathbf{1}\big\{|x_i - \tau| \le \gamma\big\}. \qquad (4) \end{aligned}$$

Thus, for each i, both $U_{\tau,i}$ and $U_{\gamma,i}$ are continuous surfaces over the (τ,γ)-plane, but they have folds along the rays defined by

$$R_{i+} = \big\{\boldsymbol{\theta} \in \Omega : \gamma = \tau - x_i\big\}, \quad R_{i-} = \big\{\boldsymbol{\theta} \in \Omega : \gamma = x_i - \tau\big\}.$$

Summing these surfaces over i produces continuous $U_{n\tau}$ and $U_{n\gamma}$ surfaces, each with n pairs of folds indexed by the data $x_1, \ldots, x_n$ (Chiu, 2002, Figure 4.1). As a result, the matrix $\mathbb{V}_n$ is well-defined everywhere on Ω except along the $R_{i\pm}$'s. To avoid technical difficulties we define a "directional" Hessian:

$$V^+_{\tau j,i}(\boldsymbol{\theta}) = \lim_{h \downarrow 0} \frac{\partial}{\partial \tau} U_{j,i}(\tau + h, \gamma) \quad , \quad V^+_{\gamma j,i}(\boldsymbol{\theta}) = \lim_{h \downarrow 0} \frac{\partial}{\partial \gamma} U_{j,i}(\tau, \gamma + h)$$

for all $i = 1, \ldots, n$ and $j = \tau, \gamma$. Of course, these V_i^+'s are merely regular derivatives when evaluated at some $\boldsymbol{\theta} \notin R_{k\pm}$, $k = 1, \ldots, n$. Now, we can replace $\mathbb{V}_n$ by

$$\mathbb{V}^+_n(\boldsymbol{\theta}) = \sum_{i=1}^n \begin{bmatrix} V^+_{\tau\tau,i}(\boldsymbol{\theta}) & V^+_{\tau\gamma,i}(\boldsymbol{\theta}) \\ V^+_{\gamma\tau,i}(\boldsymbol{\theta}) & V^+_{\gamma\gamma,i}(\boldsymbol{\theta}) \end{bmatrix}$$

which is well-defined on Ω, and coincides with $\mathbb{V}_n(\boldsymbol{\theta})$ except along the $R_{k\pm}$'s.

Lemma 2 *For all $\boldsymbol{\theta} \in \Omega$, we have*

$$\boldsymbol{U}_n(\boldsymbol{\theta}) = \boldsymbol{U}_n(\boldsymbol{\theta}_0) + \left[\int_0^1 \mathbb{V}^*_n(\boldsymbol{\theta}, t)\, dt\right]^T (\boldsymbol{\theta} - \boldsymbol{\theta}_0) \tag{5}$$

where, for all $t \in [0, 1]$ and $\boldsymbol{\theta} \in \Omega$,

$$\mathbb{V}^*_n(\boldsymbol{\theta}, t) = \sum_{i=1}^n \begin{bmatrix} V^+_{\tau\tau,i}\Big(\tau_0 + t(\tau - \tau_0), \quad \gamma_0 \quad\Big) & V^+_{\tau\gamma,i}\Big(\tau_0 + t(\tau - \tau_0), \quad \gamma_0 \quad\Big) \\ V^+_{\gamma\tau,i}\Big(\quad \tau \quad, \gamma_0 + t(\gamma - \gamma_0)\Big) & V^+_{\gamma\gamma,i}\Big(\quad \tau \quad, \gamma_0 + t(\gamma - \gamma_0)\Big) \end{bmatrix}.$$

Remark 2. Had the Hessian been continuous, all its components in the integrand would have been simply $\big(\tau_0 + t(\tau - \tau_0), \gamma_0 + t(\gamma - \gamma_0)\big) = \boldsymbol{\theta}_0 + t(\boldsymbol{\theta} - \boldsymbol{\theta}_0)$. As $\mathbb{V}^+_n$ is discontinuous, slightly different arguments are used, as given by $\mathbb{V}^*_n$.

Lemma 3 *Given are conditions [A] to [D], and a sequence $\delta_n \downarrow 0$. Then,*

$$\forall\ j, k = \tau, \gamma \qquad \sup_{\boldsymbol{\theta}\, \in\, \Theta_{\delta_n}} \left| 1 + \frac{V^+_{n,jk}(\boldsymbol{\theta})}{I_{n,jk}(\boldsymbol{\theta}_0)} \right| \xrightarrow{\text{P}} 0 \qquad \text{as } n \longrightarrow \infty \tag{6}$$

where $\Theta_r = \{\boldsymbol{\theta} : |\boldsymbol{\theta} - \boldsymbol{\theta}_0| \le r\}$ and $V^+_{n,jk}$ denotes the (j,k)-th component of $\mathbb{V}^+_n$, and similarly for $I_{n,jk}$.

Lemma 4 (Corollary to Theorem 1) *Under conditions [A] and [B], there exists a decreasing sequence* $\xi_n \downarrow 0$ *such that* $P_{\boldsymbol{\theta}_0}\{|\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0| \le \xi_n\} \longrightarrow 1$ *as* $n \longrightarrow \infty$.

Remark 3. In proving Theorem 2, we concentrate on the behavior of the $\boldsymbol{U}_n$ surface over Θ_{ξ_n}.

6 Proofs

In addition to the notation from Sections 4 and 5, write

$$T_n(\boldsymbol{\theta}) = \frac{1}{n}\sum_1^n \left| d_{\boldsymbol{\theta}_0,\boldsymbol{\theta}}(x_i) \right|^2 ,$$

$$S_n(\boldsymbol{\theta}) = \sum_1^n \left| \varepsilon_i + d_{\boldsymbol{\theta}_0,\boldsymbol{\theta}}(x_i) \right|^2 , \quad \overline{S}_n(\boldsymbol{\theta}) = \frac{1}{n} S_n(\boldsymbol{\theta}) \quad , \quad H_n(\boldsymbol{\theta}) = E_{\boldsymbol{\theta}_0}\left[\overline{S}_n(\boldsymbol{\theta}) \right] .$$

Proof of Lemma 1: See Chiu (2002). ■

Proof of Theorem 1

This proof of consistency is mostly standard, with the exceptions that the parameter space is unbounded, and that the boundary value of $\gamma_0 = 0$ is handled properly. Below, we highlight the crucial elements of the proof.

Claim 1: There exists $M^* > 0$ such that $P_{\boldsymbol{\theta}_0}\{|\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0| > M^*\} \longrightarrow 0$.

Proof: By conditions [A] and [B], there are $M_1, M_2, N > 0$ such that

$$n > N \implies \inf_{\substack{\gamma > M_2 \\ \tau \in [-M_1, M]}} \inf_{\substack{\tau < -M_1 \\ \gamma \ge 0}} \frac{1}{n} T_n(\boldsymbol{\theta}) \ \ge \ 6\sigma^2 .$$

This lower bound and the Strong Law of Large Numbers can then be applied to show that $\inf_{\boldsymbol{\theta} \notin \Theta_{M^*}} \overline{S}_n(\boldsymbol{\theta}) \overset{\text{a.s.}}{>} \overline{S}_n(\widehat{\boldsymbol{\theta}}_n)$ for all $n > N$, where $M^* = \max\{M_1, M_2\}$.

Claim 2: For each $\delta \in (0, \delta_{11}/2)$, there is an $\eta > 0$ such that $\liminf_{n\to\infty} \inf_{\boldsymbol{\theta} \in D_\delta}\{H_n(\boldsymbol{\theta}) - H_n(\boldsymbol{\theta}_0)\} > \eta$, where $D_\delta = \{\boldsymbol{\theta} : \delta \le |\boldsymbol{\theta} - \boldsymbol{\theta}_0| \le M^*\}$.

Proof: Define $T^*(w_0, w_1, \boldsymbol{\theta}) = \left|d_{\boldsymbol{\theta}_0,\boldsymbol{\theta}}(w_0)\right|^2 + \left|d_{\boldsymbol{\theta}_0,\boldsymbol{\theta}}(w_1)\right|^2$ and

$$C_{\delta,R} = \left\{ (w_0, w_1) : \begin{array}{lcl} w_0 & \in & \left\{ \begin{array}{ll} [\tau_0 - \gamma_0 + \delta_1, \tau_0 + \gamma_0 - \delta_1] & \text{if } \ \gamma_0 > 0 \\ [\tau_0 - \delta/8, \tau_0 + \delta/8] & \text{if } \ \gamma_0 = 0 \end{array} \right. \\ w_1 & \in & [\tau_0 + \gamma_0 + \delta_{11}, R] \end{array} \right\},$$

for some $R \in (\tau_0 + \gamma_0 + 2M^*, \infty)$. By Lemma 1, a compactness argument can show that $\eta_{\delta,R} \equiv \inf\{T^*(w_0, w_1, \boldsymbol{\theta}) : \boldsymbol{\theta} \in D_\delta, \ (w_0, w_1) \in C_{\delta,R}\} > 0$. Next, pick $\epsilon^* \in (0, \delta/4)$ and define $c^* = \min\{ c^*_-, c_0(\epsilon^*), c^*_+ \} > 0$. Then, choose from the dataset $\{x_1, \ldots, x_n\}$ pairs of design points $(x_{0j}, x_{1j}) \in C_\delta \equiv \lim_{K\to\infty} C_{\delta,K}$. Conditions [A] and [B] imply that there are at least nc^* such pairs for all $n > N_1$ for some N_1. Moreover, it can be shown that $\eta_{\delta,K} = \eta_{\delta,R}$ for all $K \geq R$. Take $\eta = \eta_{\delta,R}\, c^*$. Then,

$$T_n(\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^{n}\left|d_{\boldsymbol{\theta}_0,\boldsymbol{\theta}}(x_i)\right|^2 \geq \frac{1}{n}\sum_{j=1}^{nc^*} T^*(x_{0_j}, x_{1_j}, \boldsymbol{\theta}) \geq \frac{1}{n}\sum_{j=1}^{nc^*}\left\{\inf_{C_\delta} T^*\right\} \geq \eta_{\delta,R}\, c^* = \eta\,.$$

This lower bound is positive and $\boldsymbol{\theta}$-free. Then, $\inf_{\boldsymbol{\theta} \in D_\delta} T_n(\boldsymbol{\theta}) \geq \eta$ for all $n > N_1$ for the given D_δ. Since $T_n(\boldsymbol{\theta}) = H_n(\boldsymbol{\theta}) - H_n(\boldsymbol{\theta}_0)$, the claim follows.

Claim 3: For all $\epsilon > 0$, $P_{\boldsymbol{\theta}_0}\left\{\sup_{\boldsymbol{\theta} \in \Theta_{M^*}} \left|\overline{S}_n(\boldsymbol{\theta}) - H_n(\boldsymbol{\theta})\right| \leq \epsilon\right\} \longrightarrow 1$.

Proof: We employ a standard chaining argument, using the fact that the basic bent cable $q(x_i;\, \boldsymbol{\theta})$ satisfies the Lipschitz condition

$$\left|q(x;\, \boldsymbol{\theta}_1) - q(x;\, \boldsymbol{\theta}_2)\right| = \left|d_{\boldsymbol{\theta}_1,\boldsymbol{\theta}_2}(x)\right| \leq B_r\, |\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2| \ \ \forall\, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \Theta_r \ \ \forall\, x \in \mathbb{R}\,, \qquad (7)$$

where $r \leq M^*$, and B_r is a positive constant that can be derived using (3) and (4). The remainder of the proof is straightforward. ■

Proof of Lemma 2

Each summand of $\boldsymbol{U}_n$ is piecewise continuously differentiable as a function of either τ or γ. Thus, for example, integrating $V^+_{n,\tau\tau}$ over τ gives

$$\begin{aligned} U_{n\tau}(\tau,\gamma) - U_{n\tau}(\tau_0,\gamma) &= \int_{\tau_0}^{\tau} \sum_{i=1}^{n} V^+_{\tau\tau,i}(s,\gamma)\, ds \\ &= (\tau-\tau_0) \int_0^1 \sum_{i=1}^{n} V^+_{\tau\tau,i}\Big(\tau_0 + t(\tau-\tau_0), \gamma\Big)\, dt\,. \end{aligned}$$

The algebra is similar for integrating $V^+_{n,\gamma\tau}$ over γ and for other $V^+_{n,jk}$'s. Then, since

$$U_{nj}(\boldsymbol{\theta}) = U_{nj}(\tau_0,\gamma_0) + \big[U_{nj}(\tau,\gamma_0) - U_{nj}(\tau_0,\gamma_0)\big] + \big[U_{nj}(\tau,\gamma) - U_{nj}(\tau,\gamma_0)\big]$$

for all $j = \tau, \gamma$, the lemma follows by routine algebra. ■

Proof of Lemma 3

We examine the components of $\mathbb{V}^+_n$ and $\mathbb{I}_n$. To simplify the algebra, write

$$\alpha_{1i} = \mathbf{1}\big\{x_i > \tau + \gamma\big\} \quad , \quad \alpha_{2i} = \frac{x_i - (\tau-\gamma)}{2\gamma}\mathbf{1}\big\{|x_i - \tau| \le \gamma\big\}$$

$$\alpha_{3i} = \frac{1}{4}\left[1 - \left|\frac{x_i-\tau}{\gamma}\right|^2\right]\mathbf{1}\big\{|x_i-\tau| \le \gamma\big\} \quad , \quad \alpha_{4i} = x_i - \tau$$

all of which are functions of $\boldsymbol{\theta}$ (but the argument is suppressed in the notation).

Now, $\partial q_{\boldsymbol{\theta}}(x_i)/\partial\tau = \alpha_{1i} + \alpha_{2i}$ and $\partial q_{\boldsymbol{\theta}}(x_i)/\partial\gamma = -\alpha_{3i}$, and the directional derivatives for the summands of $\boldsymbol{U}_n$ are

$$\begin{aligned} V^+_{\tau\tau,i}(\boldsymbol{\theta}) &= -I_{\tau\tau,i} + \frac{1}{2\gamma\sigma^2}\Big[\varepsilon_i + d_{\boldsymbol{\theta}_0,\boldsymbol{\theta}}(x_i)\Big]\mathbf{1}\big\{\tau-\gamma < x_i \le \tau+\gamma\big\} \\ V^+_{\tau\gamma,i}(\boldsymbol{\theta}) &= -I_{\tau\gamma,i} + \frac{\alpha_{4i}}{2\gamma^2\sigma^2}\Big[\varepsilon_i + d_{\boldsymbol{\theta}_0,\boldsymbol{\theta}}(x_i)\Big]\mathbf{1}\big\{\tau-\gamma < x_i \le \tau+\gamma\big\} \\ V^+_{\gamma\tau,i}(\boldsymbol{\theta}) &= -I_{\tau\gamma,i} + \frac{\alpha_{4i}}{2\gamma^2\sigma^2}\Big[\varepsilon_i + d_{\boldsymbol{\theta}_0,\boldsymbol{\theta}}(x_i)\Big]\mathbf{1}\big\{|x_i-\tau| \le \gamma\big\} \\ V^+_{\gamma\gamma,i}(\boldsymbol{\theta}) &= -I_{\gamma\gamma,i} + \frac{\alpha_{4i}^2}{2\gamma^3\sigma^2}\Big[\varepsilon_i + d_{\boldsymbol{\theta}_0,\boldsymbol{\theta}}(x_i)\Big]\mathbf{1}\big\{|x_i-\tau| \le \gamma\big\} \end{aligned}$$

where the summands of the components of $\mathbb{I}_n$ are

$$I_{\tau\tau,i}(\boldsymbol{\theta}) = \frac{\alpha_{1i} + \alpha_{2i}^2}{\sigma^2}\,,\ I_{\tau\gamma,i}(\boldsymbol{\theta}) = -\frac{\alpha_{2i}\,\alpha_{3i}}{\sigma^2} = I_{\gamma\tau,i}(\boldsymbol{\theta})\,,\ I_{\gamma\gamma,i}(\boldsymbol{\theta}) = \frac{\alpha_{3i}^2}{\sigma^2}\,. \tag{8}$$

By conditions [A] and [B], one can show that there are N and $M_{jk} > m_{jk} > 0$, $j,k = \tau,\gamma$, such that $n > N$ implies

$$m_{jk} \le \frac{1}{n} I_{n,jk}(\boldsymbol{\theta}_0) \le M_{jk} \quad \forall\ j,k = \tau,\gamma\,. \tag{9}$$

That is, for all $j,k = \tau,\gamma$, $I_{n,jk}(\boldsymbol{\theta}_0)$ is bounded between two non-trivial multiples of n for all sufficiently large n. Thus, to prove the lemma, it suffices to show

$$\sup_{\boldsymbol{\theta}\,\in\,\Theta_{\delta_n}} \left|\eta_{n,jk}(\boldsymbol{\theta})\right| \;=\; o(n) \tag{10}$$

$$\sup_{\boldsymbol{\theta}\,\in\,\Theta_{\delta_n}} \left|\Delta_{n,jk}(\boldsymbol{\theta})\right| \;=\; o_p(n) \tag{11}$$

for each pair of (j,k), where $\eta_{n,jk}(\boldsymbol{\theta}) = I_{n,jk}(\boldsymbol{\theta}) - I_{n,jk}(\boldsymbol{\theta}_0)$ and $\Delta_{n,jk}(\boldsymbol{\theta}) = V^{+}_{n,jk}(\boldsymbol{\theta}) + I_{n,jk}(\boldsymbol{\theta})$. There are three cases: (a) $j = k = \tau$, (b) $j = k = \gamma$, and (c) $j \ne k$. We do only Case (a).

First, consider (10). Given any $\boldsymbol{\theta} \in \Theta_{\delta_n}$, apply (8) to get

$$\sigma^2\left|\eta_{n,\tau\tau}(\boldsymbol{\theta})\right| \le \sum\left|\alpha_{1i}^2 - \alpha_{1i,0}^2\right| + \sum\left|\alpha_{2i}^2 - \alpha_{2i,0}^2\right| \tag{12}$$

where $\alpha_{ki,0}$ $(k = 1,2,3)$ is the value of α_{ki} with $\boldsymbol{\theta}$ replaced by $\boldsymbol{\theta}_0$.

For the second sum in (12), partition the index set $\{1,\ldots,n\}$ into the four sets $\mathcal{I}_1 \equiv \big\{i : |x_i-\tau| > \gamma,\, |x_i-\tau_0| > \gamma_0\big\}$, $\mathcal{I}_2 \equiv \big\{i : |x_i-\tau| \le \gamma,\, |x_i-\tau_0| \le \gamma_0\big\}$, $\mathcal{I}_3 \equiv \big\{i : |x_i-\tau| \le \gamma,\, |x_i-\tau_0| > \gamma_0\big\}$, and $\mathcal{I}_4 \equiv \big\{i : |x_i-\tau| > \gamma,\, |x_i-\tau_0| \le \gamma_0\big\}$. Then, use (7) to show that $\sup_{\boldsymbol{\theta}\,\in\,\Theta_{\delta_n}}\left|\alpha_{2i}^2 - \alpha_{2i,0}^2\right| = O(\delta_n)$ over $\mathcal{I}_1$ and $\mathcal{I}_2$. For $i \in \mathcal{I}_3 \cup \mathcal{I}_4$, note that $\left|\alpha_{2i}^2 - \alpha_{2i,0}^2\right| \le 1$. Now, define $K_n^{+} = [\tau_0+\gamma_0-\delta_n, \tau_0+\gamma_0+\delta_n]$

and $K_n^- = [\tau_0 - \gamma_0 - \delta_n, \tau_0 - \gamma_0 + \delta_n]$, which shrink to the join points $\tau_0 \pm \gamma_0$ as $n \longrightarrow \infty$. Then, apply condition [C] to see that for all sufficiently large n,

$$\begin{aligned} \sum_{i=1}^{n} \left| \alpha_{2i}^2 - \alpha_{2i,0}^2 \right| &= n\,O(\delta_n) + \sum_{i \in \mathcal{I}_3 \cup \mathcal{I}_4} \left| \alpha_{2i}^2 - \alpha_{2i,0}^2 \right| \\ &\leq n\,o(1) + \sum_{i:\, x_i \in K_n^{\pm}} 1 \;\leq\; o(n) + n\,\zeta_n(\delta_n) \;=\; o(n)\,. \end{aligned}$$

This bound is $\boldsymbol{\theta}$-free. A similar argument applies to the first sum in (12), with a suitable partition of the index set. Therefore, $\left|\eta_{n,\tau\tau}(\boldsymbol{\theta})\right|$ is uniformly bounded by $o(n)$ over Θ_{δ_n}. (For Cases (b) and (c), not having "accumulation" at the joints $\tau_0 \pm \gamma_0$ is crucial in keeping the matrices $\mathbb{I}_n(\boldsymbol{\theta})$ and $\mathbb{I}_n(\boldsymbol{\theta}_0)$ close everywhere on Θ_{δ_n}.)

To show (11), Case (a), first, define $a_{n,\tau\tau}(\boldsymbol{\theta}) = (2\gamma)^{-1} \sum d_{\boldsymbol{\theta}_0,\boldsymbol{\theta}}(x_i)\, \mathbf{1}\{\tau - \gamma < x_i \leq \tau + \gamma\}$ and $B_{n,\tau\tau}(\boldsymbol{\theta}) = (2\gamma)^{-1} \sum \varepsilon_i\, \mathbf{1}\{\tau - \gamma < x_i \leq \tau + \gamma\}$ so that $\sigma^2\, \Delta_{n,\tau\tau}(\boldsymbol{\theta}) = a_{n,\tau\tau}(\boldsymbol{\theta}) + B_{n,\tau\tau}(\boldsymbol{\theta})$. By (7), one can show that $\sup_{\boldsymbol{\theta} \in \Theta_{\delta_n}} \left|a_{n,\tau\tau}(\boldsymbol{\theta})\right| = o(n)$. Now, relabel the data so that $x_1 \leq x_2 \leq \ldots \leq x_n$, and define the martingale $\mathcal{M}_m = \sum_{i=1}^{m} \varepsilon_i$, $1 \leq m \leq n$. For each $\boldsymbol{\theta} \in \Theta_{\delta_n}$, there are integers s and t, $1 \leq s \leq t \leq n$, such that $\sum_{i=1}^{n} \varepsilon_i\, \mathbf{1}\big\{\tau - \gamma < x_i \leq \tau + \gamma\big\} = \mathcal{M}_t - \mathcal{M}_s$. Hence,

$$\left|B_{n,\tau\tau}(\boldsymbol{\theta})\right| = \frac{\left|\mathcal{M}_t - \mathcal{M}_s\right|}{2\gamma} \leq \frac{\left|\mathcal{M}_t\right| + \left|\mathcal{M}_s\right|}{2\gamma} \leq \frac{1}{\gamma_0 - \delta_n} \left| \max_{1 \leq m \leq n} \mathcal{M}_m \right|.$$

This $\boldsymbol{\theta}$-free bound is $O_p(\sqrt{n})$ by the Doob-Kolmogorov inequality (see Breiman, 1968, p. 89, Problem 2). Thus, $\sup_{\boldsymbol{\theta} \in \Theta_{\delta_n}} \left|\Delta_{n,\tau\tau}(\boldsymbol{\theta})\right| \leq o(n) + O_p(\sqrt{n}) = o_p(n)$.

Note that for Cases (b) and (c), each summand of $B_{n,\tau\gamma}(\boldsymbol{\theta})$ or $B_{n,\gamma\gamma}(\boldsymbol{\theta})$ has the form $\varepsilon_i\, \omega_i(\boldsymbol{\theta})$, and the martingale tactic above does not apply. Instead, the reader can verify that a routine chaining argument suffices in these cases. ■

Proof of Theorem 2

Part 1: Denote the eigenvalues of the covariance matrix $n^{-1}\, \mathbb{I}_n(\boldsymbol{\theta}_0)$ by λ_{n1} and

λ_{n2}, where $0 \leq \lambda_{n1} \leq \lambda_{n2}$. With the upper bounds of (9), it remains to show

$$\exists \text{ integer } N \text{ and } \epsilon > 0 \quad \text{s.t.} \qquad n > N \implies \lambda_{n1} \geq \epsilon\,. \tag{13}$$

One can prove (13) by relating λ_{n1} and λ_{n2} to the trace and determinant of $n^{-1}\,\mathbb{I}_n(\boldsymbol{\theta}_0)$, then applying conditions [A] and [B]. We omit the details.

Part 2: By eigenvalue properties, Lemma 3, and Part 1 of the theorem,

$$P_{\boldsymbol{\theta}_0}\left\{\frac{1}{n}\mathbb{V}_n^{+}(\boldsymbol{\theta}) \text{ is negative definite for all } \boldsymbol{\theta} \in \Theta_{\xi_n}\right\} \longrightarrow 1\,. \tag{14}$$

Now, we introduce a lemma about the concavity of a once-differentiable function. Its first assertion is due to Theorem 4.4.10 in Stoer and Witzgall (1970), and the second, the definition of concavity. We omit the proof.

Lemma 5 *(1) Let $\mathcal{N}$ be a subset of [0,1] consisting of isolated points. Suppose that a differentiable function $f : [0,1] \longrightarrow \mathbb{R}$ has a continuous second derivative, f'', in $[0,1] \setminus \mathcal{N}$. Then, f is strictly concave on $[0,1]$ if $f''(t) < 0$ for all but those isolated values of $t \in \mathcal{N}$. (2) A differentiable function $g : \Theta_{\xi_n} \longrightarrow \mathbb{R}$ is strictly concave if and only if $h(t) \equiv g\big(\boldsymbol{\theta}_1 + t(\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1)\big)$ is concave on $[0,1]$ for each $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \Theta_{\xi_n}$, where $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_2$.*

For our problem, we study $h_n(t) \equiv \ell_n\big(\boldsymbol{\theta}_1 + t(\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1)\big)$ for $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \Theta_{\xi_n}\,, \boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_2\,, t \in [0,1]$. In what follows, we restrict our attention to the event defined in (14).

First, consider $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ which do not both lie on $R_{k\pm}$ for any k. They define a line segment along which ℓ_n is piecewise twice differentiable. By the chain rule,

$$\frac{1}{n}h_n''(t) = (\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1)^T \left[\frac{1}{n}\mathbb{V}_n\big(\boldsymbol{\theta}_1 + t(\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1)\big)\right] (\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1) \tag{15}$$

$$= (\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1)^T \left[\frac{1}{n}\mathbb{V}_n^{+}\big(\boldsymbol{\theta}_1 + t(\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1)\big)\right] (\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1) \;<\; 0 \tag{16}$$

for all but isolated values of $t \in [0,1]$. By Lemma 5, it follows that $n^{-1}\,\ell_n$, hence, ℓ_n, is concave along this line segment.

We are left to examine the case where $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ both lie on a ray. By symmetry, we need only to consider, say, the 45-degree R_{k+}'s. Here, the chain rule cannot be applied to yield (15), since $\mathbb{V}_n$ is not defined at $\boldsymbol{\theta}_1 + t(\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1)$ for any $t \in [0,1]$. Instead, one can expand $h_n(t)$, differentiate it twice, and verify, after some lengthy algebra, that $n^{-1}h_n(t)$ indeed equals the expression in (16) for all but isolated values of $t \in [0,1]$. These isolated values correspond to the intersections of the line segment (joining $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ on R_{k+}) with the 135-degree R_{i-}'s.

Thus, we have shown that every cross-section, and hence, the entire surface, of ℓ_n over Θ_{ξ_n} is strictly concave on the event defined in (14). By consistency,

$$P_{\boldsymbol{\theta}_0}\Big\{\widehat{\boldsymbol{\theta}}_n \text{ is the unique maximizer of } \ell_n \text{ over } \Theta_{\xi_n}\Big\} \longrightarrow 1\,.$$

Now, this unique $\widehat{\boldsymbol{\theta}}_n$ can be substituted into the Taylor-type expansion of Lemma 2. Finally, asymptotic normality of $\widehat{\boldsymbol{\theta}}_n$ results from standard matrix manipulation by noting that the summands of $\boldsymbol{U}_n(\boldsymbol{\theta}_0)$ satisfy the Lindeberg Condition, and thus Part 1 and the Lindeberg-Feller Central Limit Theorem ensure that $\big[\mathbb{I}_n(\boldsymbol{\theta}_0)\big]^{-1/2} \times \boldsymbol{U}_n(\boldsymbol{\theta}_0)$ is asymptotically standard normal.

Parts 3 and 4: The results follow respectively from Parts 1 and 2 due to consistency and (10).

Part 5: Since ℓ_n is once-differentiable, we can write

$$\ell_n(\boldsymbol{\theta}) = \ell_n(\boldsymbol{\theta}_0) + (\boldsymbol{\theta} - \boldsymbol{\theta}_0)^T \int_0^1 \boldsymbol{U}_n\big(\boldsymbol{\theta}_0 + t(\boldsymbol{\theta} - \boldsymbol{\theta}_0)\big)\, dt\,. \tag{17}$$

By Lemma 2, we have, for all $t \in [0,1]$,

$$\boldsymbol{U}_n\big(\boldsymbol{\theta}_0 + t(\boldsymbol{\theta} - \boldsymbol{\theta}_0)\big) = \boldsymbol{U}_n(\boldsymbol{\theta}_0) + \left[\int_0^1 \mathbb{V}_n^*\big(\boldsymbol{\theta}_0 + t(\boldsymbol{\theta} - \boldsymbol{\theta}_0), s\big)\, ds\right]^T t(\boldsymbol{\theta} - \boldsymbol{\theta}_0)\,. \tag{18}$$

Substitute (18) into (17) and apply Lemma 3 and Part 2 to complete the proof. ■

7 Concluding Remarks

We have given some structurally simple conditions on the design. These provide a practical guideline for data collection when considering basic bent-cable regression. We have shown that these few regularity conditions suffice to compensate for the intrinsic irregularity of the problem due to non-differentiability of the model's first partial derivatives. In particular, they ensure that the $\mathbb{I}_n(\boldsymbol{\theta}_0)$ matrix is asymptotically well-behaved, in the sense that the amount of information about the unknown parameter $\boldsymbol{\theta}_0$ contained in all sufficiently large datasets is no less than a non-trivial multiple of the sample size. Furthermore, as $-\mathbb{V}_n^+(\boldsymbol{\theta})$ is the negative (directional) Hessian of ℓ_n, its uniform closeness to $\mathbb{I}_n(\boldsymbol{\theta}_0)$ on Θ_{ξ_n} ensures that its discontinuities are asymptotically negligible, leading to an ℓ_n-surface that is uniformly well-behaved on this ever-decreasing neighborhood of $\boldsymbol{\theta}_0$. This is the basis of the "regularity" of our problem, despite $\boldsymbol{U}_n$ surfaces that are folded. Thus, standard asymptotic results apply.

With slightly modified conditions, our results here can be extended to the full bent-cable model with non-normal errors and unknown constant variance. The details are to appear in a future article and are currently available in Chiu (2002).

Acknowledgments

We thank the Co-ordinating Editor and referees for providing valuable suggestions and extra references.

References

Barrowman, N. J. and R. A. Myers (2000). Still more spawner-recruitment curves: the hockey stick and its generalizations. *Can. J. Fish. Aquat. Sci.* 57, 665–676.

Bhattacharya, P. K. (1990). Weak convergence of the log likelihood process in the two-phase linear regression problem. Proceedings of the R. C. Bose Symposium on Probability, Statistics and Design of Experiments, 145–156. Wiley Eastern, New Delhi.

Breiman, L. (1968). *Probability.* Addison-Wesley, Reading.

Chiu, G. S. (2002). *Bent-Cable Regression for Assessing Abruptness of Change.* Ph. D. Thesis. Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, BC.

Chiu, G., R. Lockhart, and R. Routledge (2002a). Bent-cable asymptotics when the bend is missing. *Statist. Prob. Letters* 56, 9–16.

Chiu, G., R. Routledge, and R. Lockhart (2002b). Using the bent cable to assess an abrupt change in species abundance and other phenomena. Winning student paper. WNAR/IMS 2002 Conference, Los Angeles, CA.

Feder, P. I. (1975). On asymptotic distribution theory in segmented regression problems — identified case. *Ann. Statist.* 3, 49–83.

Gallant, A. R. (1974). The Theory of nonlinear regression as it relates to segmented polynomial regressions with estimated join points. Mimeograph Series No. 925. Institute of Statistics, North Carolina State University, Rayleigh, NC.

Gallant, A. R. (1975). Inference for nonlinear models. Mimeograph Series No. 875. Institute of Statistics , North Carolina State University, Rayleigh, NC.

Golub, G. H. and C. F. Van Loan (1996). *Matrix Computations.* Johns Hopkins University Press, Baltimore.

Hušková, M. (1998). Estimators in the location model with gradual changes. *Comment. Math. Univ. Carolinae* 39, 147–157.

Hušková, M. and J. Steinebach (2000). Limit theorems for a class of tests of gradual changes. *J. Statist. Plann. Inference* 89, 57–77.

Ivanov, A. V. (1997). *Asymptotic Theory of Nonlinear Regression.* Kluwer Academic Publishers, Dordrecht.

Jarušková, D. (1998a). Testing appearance of linear trend. *J. Statist. Plann. Inference* 70, 263–276.

Jarušková, D. (1998b). Change-point estimator in gradually changing sequences. *Comment. Math. Univ. Carolinae* 39, 551–561.

Jarušková, D. (2001). Change-point estimator in continuous quadratic regression. *Comment. Math. Univ. Carolinae* 42, 741–752.

Naylor, R. E. L. and J. Su (1998). Plant development of triticale cv. Lasko at different sowing dates. *J. Agric. Sci.* 130, 297–306.

Neuman, M. J., D. A. Witting and K. W. Able (2001). Relationships between otolith microstructure, otolith growth, somatic growth and ontogenetic transitions in two cohorts of windowpane. *J. Fish Biol.* 58, 967–984.

Rukhin, A. L. and I. Vajda (1997). Change-point estimation as a nonlinear regression problem. *Statistics* 30, 181–200.

Seber, G. A. F. and C. J. Wild (1989). *Nonlinear Regression.* Wiley, New York.

Stoer, J. and C. Witzgall (1970). *Convexity and Optimization in Finite Dimensions*, I. Springer-Verlag, Berlin.

Tishler, A. and I. Zang (1981). A new maximum likelihood algorithm for piecewise regression. *J. Amer. Statist. Assoc.* 76, 980–987.

Vaart, A. W. van der and J. A. Wellner (1996). *Weak Convergence and Empirical Processes.* Springer-Verlag, NY.

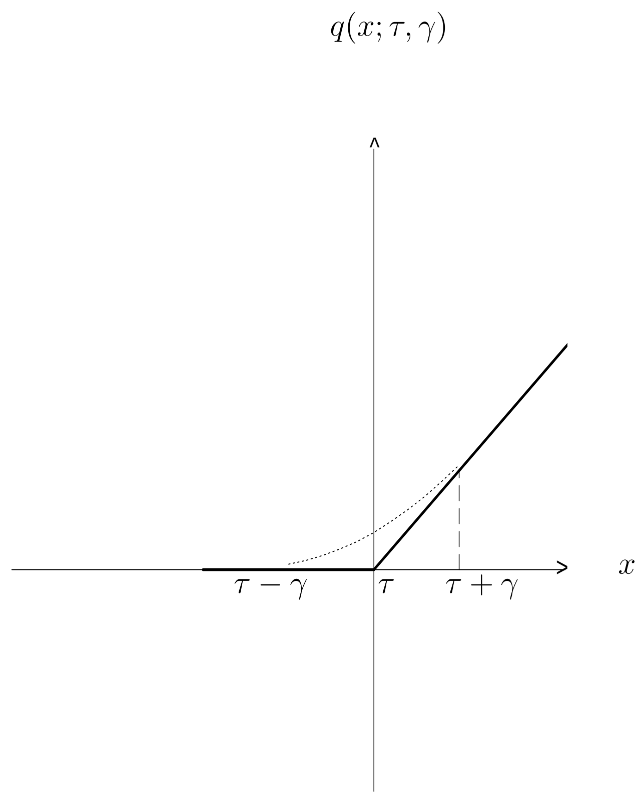

Figure 1. The basic bent cable $q(x;\tau,\gamma)$, with the dotted quadratic bend. (The solid line is the broken stick.)