Piecewise Regression Using Cubic Splines

DALE J. POIRIER *

Spline theory and piecewise regression theory are integrated to provide a framework in which structural change is viewed as occurring in a smooth fashion. Specifically, structural change occurs at given points through jump discontinuities in the third derivative of a continuous piecewise cubic estimating function. Testing procedures are developed for detecting structural change as well as linear or quadratic segments. Finally, the techniques developed are illustrated empirically in a learning-by-doing model.

> Polynomials are wonderful even after they are cut into pieces, but the cutting must be done with care. One way of doing the cutting leads to so-called spline functions.
>
> I.J. Schoenberg

1. INTRODUCTION

The preceding comment by the father of spline theory seems to be an appropriate prelude to a discussion of the increasingly popular approximation tool—the spline function. Here the topic will focus on cubic splines, i.e., piecewise functions whose "pieces" are polynomials of degrees less than or equal to three, and which are joined together in a suitably smooth fashion.

During the quarter of a century following the pioneering work of Schoenberg [24], cubic splines have drawn a great deal of attention from both theoretical and applied mathematicians. As to be expected, applications have lagged behind theoretical insights, and as a result, many possible areas of application remain to be explored. However, there have been a sufficient number of successful applications, notably in the aerospace and automotive styling industries, to give credibility to the following statement:

> Spline functions, and, more generally, piecewise polynomial functions are the most successful approximating functions in use today. They combine ease of handling in a computer with great flexibility, and are therefore particularly suited for the approximation of experimental data or design curve experiments [7, p. 1].

While continuous piecewise functions have been popular in approximation theory, their use in econometrics as a representation of structural change has not been as widespread. Structural changes at given points has for the most part been handled by analysis of covariance models involving dummy variables. Unfortunately, this has often been at the expense of continuity in the estimating function. Usually this is thought of as a necessary evil. Indeed the whole idea of continuity in economic models is built on shaky grounds since it implies infinitely divisible variables which have no real-world counterparts. However, if continuity is assumed, then it seems that it should hold at the structural shift points just as well as anywhere else. For example, it is one thing to say that the beginning of World War II caused a shift upwards in the functional relationship between government bond sales and time, and quite another thing to say that at the beginning of the war government bond sales took an *instantaneous, abrupt* jump. The first case allows for the more plausible explanation that the beginning of the war ignited a great deal of patriotism which grew rapidly causing bond sales to rise more rapidly than they would have otherwise. The second case is the standard dummy variable interpretation and implies a magical, instantaneous jump in bond sales.

Rather than build a straw man of the dummy variable approach, it should be noted that dummy variables can be used in simple linear regression in a way which permits both the slope and the intercept to change from one interval to the next, but subject to a continuity constraint. However the resulting continuous piecewise linear curve is nothing more than a *linear spline*—a fact that has apparently escaped the literature.[1]

With the exception of Blischke [4], Bellman and Roth [3], Hudson [12], and Robison [23], the basic literature dealing with piecewise regression, beginning with Quandt [21, 22], and more recently Farley and Hinich [9] and McGee and Carleton [14], has dealt with discontinuous curves. Also these authors have concentrated primarily on the extremely difficult problem of unknown points of structural change. Such approaches are hampered by both numerical and statistical problems in deriving estimators for the points of structural change and a corresponding distribution theory.

This study views spline theory and piecewise regression models from a slightly different perspective than either has been viewed previously. Rather than using splines as an approximation tool, i.e., a curve fitting tool, this study uses the cubic spline because its intrinsic piecewise nature is justified from a theoretical stand-

* Dale J. Poirier is assistant professor, Department of Economics, University of Illinois, Urbana, Ill. 61801. An initial version of this study appeared as Social Systems Research Institute Workshop Paper No. 7204. The author is grateful to Professor Dennis J. Aigner for his many helpful comments and suggestions which played a vital role throughout the development of this study. Special thanks are also due Professor Theodore F. Groves and the referees for their helpful comments. The author assumes full responsibility for any remaining errors.

[1] Recognizing this spline theory analog permits the specification of transformed variables which automatically incorporate the continuity constraints. For a recent application of this technique, see [11].

© Journal of the American Statistical Association
September 1973, Volume 68, Number 343
Applications Section

point. Attention is focused on those piecewise regression models in which (1) the piecewise nature reflects the occurrence of structural change, (2) the pieces are defined over adjacent intervals of the independent variable, (3) the overall model is continuous with structural change manifesting itself through jump discontinuities in the third derivative, and (4) the points of structural change are known.[2] Treating the points of structural change as given is not as restrictive as may first appear since often the underlying theory states where a point of structural change may have occurred, and it is the purpose of the model to test this hypothesis. The bond sales example used earlier is a case in point.

2. DERIVATION OF THE CUBIC SPLINE

The development presented in this section is patterned after Ahlberg, Nilson, and Walsh [1, pp. 9–16], Kershaw [13], and Pennington [15, 404–11]. The author feels that this approach is easier to grasp than the more abstract, mathematically sophisticated approaches found in most of the approximation literature, e.g., [25]. The following definition is a precise formulation of the abstract properties of the cubic spline referred to in the Introduction.

Definition: Let the set $\Delta = \{x_0 < x_1 < \cdots < x_k\}$ of abscissa values be referred to as a *mesh* of $[x_0,x_k]$ and the $k+1 \geq 3$ individual points $x_j (j = 0, 1, \cdots, k)$ as *knots*.[3] Let $y = \{y_0, y_1, \cdots, y_k\}$ be an associated set of ordinates. Then a *cubic spline on* Δ *interpolating to* y, denoted $S_\Delta(x)$, as a function satisfying:

(1) $S_\Delta(x) \in C^2[x_0,x_k]$,
(2) $S_\Delta(x)$ coincides with a polynomial of degree at most three on the intervals $[x_{j-1},x_j]$ $(j = 1, 2, \cdots, k)$, and
(3) $S_\Delta(x_j) = y_j$ $(j = 0, 1, \cdots, k)$.

A graphical example of a cubic spline is given in Figure A. Each of the arcs between the points (x_{j-1}, y_{j-1}) and (x_j,y_j) $(j=1, 2, 3)$ is a polynomial of degree at most three, and adjacent arcs are joined together in a continuous fashion according to (1).

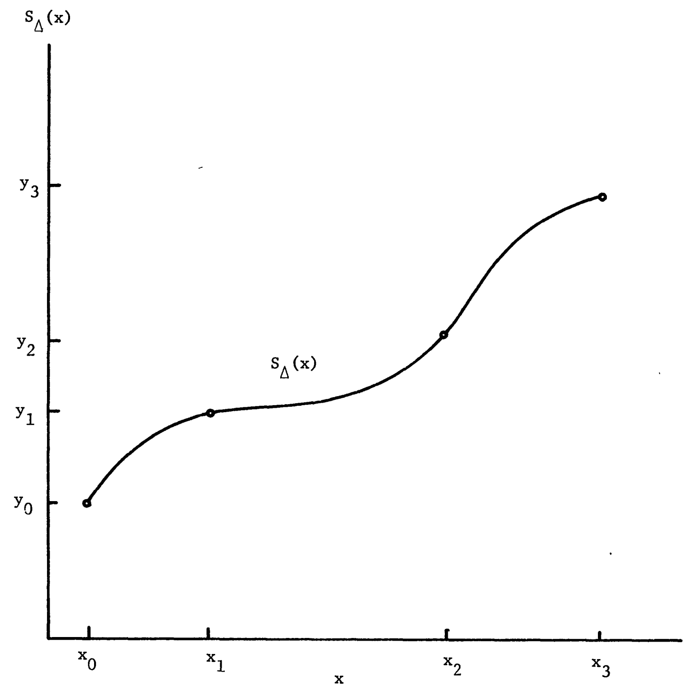
Leaving details to the Appendix, conditions (1)–(3) imply the continuity conditions

$$(1-\lambda_j)M_{j-1} + 2M_j + \lambda_j M_{j+1} = \frac{6y_{j-1}}{h_j(h_j+h_{j+1})} - \frac{6y_j}{h_j h_{j+1}} + \frac{6y_{j+1}}{h_{j+1}(h_j+h_{j+1})} \quad (2.1)$$

$(j = 1, 2, \cdots, k-1)$, where, due to their analog in beam theory, the values of the second derivative at the knots, $M_j = S_\Delta''(x_j)(j = 0, 1, \cdots, k)$, are referred to as *moments* [1, pp. 1, 2, 10]; $h_j = x_j - x_{j-1}$ are the interval lengths; and

$$\lambda_j = \frac{h_{j+1}}{h_j + h_{j+1}} \quad (j = 1, 2, \cdots, k-1). \quad (2.2)$$

A. EXAMPLE OF A CUBIC SPLINE



The continuity conditions (2.1) consist of $k-1$ equations in the $k+1$ unknowns M_j $(j = 0, 1, \cdots, k)$. This underspecification can be corrected by placing constraints, known as *end conditions*, on the spline at its end points x_0 and x_k. While many different types of end conditions have been discussed in the literature, this study focuses attention on only one.[4] By proceeding in this manner the discussion will be simplified and one drawback of the other end conditions will be avoided, namely, having to assert on *a priori* grounds rather specific information, such as the value of a derivative at an end point.

It is assumed that a cubic spline satisfies the following end conditions:

$$M_0 = \pi_0 M_1, \quad |\pi_0| < 2, \quad (2.3)$$

$$M_k = \pi_k M_{k-1}, \quad |\pi_k| < 2, \quad (2.4)$$

where π_0 and π_k are parameters that must be specified. Deciding on values for π_0 and π_k is not as formidable a task as it may at first seem, and in light of the following discussion, it is usually possible to pick values with some degree of *a priori* justification. For simplicity the discussion is in terms of π_0, but completely analogous results hold for π_k.

Choosing $\pi_0 = 1$ implies that $S_\Delta(x)$ reduces to a quadratic over the first interval. If $\pi_0 = 0$, then $M_0 = 0$ regardless of the value of M_1.[5] Furthermore, $M_0 = 0$

[2] Even ignoring distributional questions, obtaining numerical estimates of the points which define the pieces of the cubic spline can be quite difficult. See, e.g., [8].

[3] Hudson [12] refers to the interior knots $x_1, x_2, \cdots, x_{k-1}$ as "join points."

[4] End conditions are usually classified according to the following definition. $S_\Delta(x)$ is said to be periodic of period $x_k - x_0$ if and only if $S_\Delta^{(m)}(x_0^+) = S_\Delta^{(m)}(x_k^-)$ $(m = 0, 1, 2)$. Otherwise, $S_\Delta(x)$ is said to be *nonperiodic*.

[5] When it is also specified that $\pi_k = 0$, $S_\Delta(x)$ is referred to as a *natural spline*.

allows for a linear hookup with the spline at x_0, implies that $S'_\Delta(x)$ has a critical point at x_0, and is compatible with either convexity or concavity of $S_\Delta(x)$ over $[x_0,x_1]$.

If $0 < \pi_0 < 1$, then the results of the preceding paragraph hold at the "pseudo" knot $x_{-1} = (x_0 - \pi_0 x_1)/(1 - \pi_0) < x_0$, i.e., $S''_\Delta(x_{-1}) = 0$.[6] Specifically, $\pi_0 = \frac{1}{2}$ implies that $x_0 - x_{-1} = x_1 - x_0$, and $\pi_0 = x_0/x_1$ implies that $x_{-1} = 0$.

The end conditions (2.3) and (2.4) can be combined with the continuity conditions (2.1) to give a concise matrix formulation in terms of the unknown moments. Define the $(k + 1) \times (k + 1)$ coefficient matrices:

$$\mathbf{\Lambda} = \begin{bmatrix} 2 & -2\pi_0 & 0 & \cdots & 0 & 0 & 0 \\ 1-\lambda_1 & 2 & \lambda_1 & \cdots & 0 & 0 & 0 \\ 0 & 1-\lambda_2 & 2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & \lambda_{k-2} & 0 \\ 0 & 0 & 0 & \cdots & 1-\lambda_{k-1} & 2 & \lambda_{k-1} \\ 0 & 0 & 0 & \cdots & 0 & -2\pi_k & 2 \end{bmatrix} \tag{2.5}$$

$$\boldsymbol{\theta} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \frac{6}{h_1(h_1+h_2)} & \frac{-6}{h_1h_2} & \frac{6}{h_2(h_1+h_2)} & \cdots & 0 & 0 & 0 \\ 0 & \frac{6}{h_2(h_2+h_3)} & \frac{-6}{h_2h_3} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{-6}{h_{k-2}h_{k-1}} & \frac{6}{h_{k-1}(h_{k-2}+h_{k-1})} & 0 \\ 0 & 0 & 0 & \cdots & \frac{6}{h_{k-1}(h_{k-1}+h_k)} & \frac{-6}{h_{k-1}h_k} & \frac{6}{h_k(h_{k-1}+h_k)} \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix} \tag{2.6}$$

and the column vectors

$$\mathbf{M} = [M_0, M_1, \cdots, M_k]', \tag{2.7}$$

$$\mathbf{Y} = [y_0, y_1, \cdots, y_k]'. \tag{2.8}$$

Then the continuity and the end conditions together can be written[7]

$$\mathbf{\Lambda M} = \boldsymbol{\theta}\mathbf{Y}. \tag{2.9}$$

For cases in which $|\pi_0| < 1$ and $|\pi_k| < 1$, Gershgorin's Theorem can be used to show that $\mathbf{\Lambda}$ is nonsingular since the absolute value of each diagonal element is greater than the sum of the absolute values of the off-diagonal elements in its row. With slightly more effort, $\mathbf{\Lambda}$ can be shown to be nonsingular for $1 \leq |\pi_0| < 2$ and $1 \leq |\pi_k| < 2$.[8] Hence, (2.9) can be solved for the unknown moment vector

$$\mathbf{M} = \mathbf{\Lambda}^{-1}\boldsymbol{\theta}\mathbf{Y}. \tag{2.10}$$

In cases where $\mathbf{Y}$ is known (exact approximation problems), (2.9) can be solved for $\mathbf{M}$ efficiently, without actually inverting $\mathbf{\Lambda}$, by using a time-honored method of solving a linear system having a tridiagonal matrix. See, e.g., [1, pp. 14–15]. Unfortunately, in the case of least squares approximation, $\mathbf{Y}$ is unknown and $\mathbf{\Lambda}^{-1}$ is needed in explicit form. Nevertheless, the tridiagonal structure of $\mathbf{\Lambda}$ permits a stable inverting process.

Gathering together (2.10) and (A.3), it is possible for any vector $\boldsymbol{\xi} = [\xi_1, \xi_2, \cdots, \xi_n]'$ of abscissa values to express the corresponding vector

$$S_\Delta(\boldsymbol{\xi}) = [S_\Delta(\xi_1), S_\Delta(\xi_2), \cdots, S_\Delta(\xi_n)]'$$

of spline interpolants as a linear function of the ordinate vector $\mathbf{Y}$. To obtain a matrix formulation of $S_\Delta(\boldsymbol{\xi})$ in terms of $\mathbf{Y}$, it is necessary to define a few coefficient matrices. Considering (A.3), let $\mathbf{P} = [p_{im}]$ and $\mathbf{Q} = [q_{im}]$ be two $(k + 1) \times (k + 1)$ matrices such that for

[6] It is understood that for $x < x_0$ or $x > x_k$, $S_\Delta(x)$ is determined by using the polynomial segment for the first or last interval, respectively.

[7] If $x_0 > 0$ it is possible to force $S_\Delta(x)$ through the origin by replacing left end condition (2.3) by the more general end condition

$$M_0 = \pi_0 M_1 + \mu_1 y_0 + \mu_2 y_1,$$

where

$$\pi_0 = \frac{x_1 - 2x_0}{x_0 - 2x_1} \quad (-2 < \pi_0 < 1),$$

$$\mu_1 = \frac{-6}{x_1^2 - h_1^2}, \quad \mu_2 = \frac{6x_0}{x_1(x_1^2 - h_1^2)}.$$

Substituting this value of M_0 into (A.3) with $j = 1$, it follows that $S_\Delta(0) = 0$. Except for changing the first two elements of the first row of θ from zero to $2\mu_1$ and $2\mu_2$, respectively, the development remains unchanged.

[8] See [1, pp. 61–3]. Requiring π_0 and π_k to be less than two in absolute value is sufficient to insure the invertibility of Λ, however, as will be seen in the example presented in Section 8, it is not necessary.

$x_{j-1} \leq \xi_i \leq x_j$ $(j = 1, 2, \cdots, k)$ $(i = 1, 2, \cdots, n)$,

$$p_{im} = \begin{Bmatrix} \dfrac{x_j - \xi_i}{6h_j}[(x_j - \xi_i)^2 - h_j^2], & \text{for } m = j-1 \\ \dfrac{\xi_i - x_{j-1}}{6h_j}[(\xi_i - x_{j-1})^2 - h_j^2], & \text{for } m = j \\ 0, & \text{otherwise} \end{Bmatrix}, \tag{2.11}$$

$$q_{im} = \begin{Bmatrix} \dfrac{x_j - \xi_i}{h_j}, & \text{for } m = j-1 \\ \dfrac{\xi_i - x_{j-1}}{h_j}, & \text{for } m = j \\ 0, & \text{otherwise} \end{Bmatrix}. \tag{2.12}$$

Then using (2.10)–(2.12) and (A.3) it follows that

$$\begin{aligned} S_\Delta(\xi) &= \mathbf{PM} + \mathbf{QY} \\ &= (\mathbf{P\Lambda^{-1}\theta} + \mathbf{Q})\mathbf{Y} \\ &= \mathbf{WY}, \end{aligned} \tag{2.13}$$

where the transformed $n \times (k+1)$ data matrix $\mathbf{W}$ is

$$\mathbf{W} = \mathbf{P\Lambda^{-1}\theta} + \mathbf{Q}. \tag{2.14}$$

3. LEAST SQUARES CUBIC SPLINES

Attention now turns toward constructing a statistical model in the form of a cubic spline plus white noise, and then using standard least squares procedures to estimate the cubic spline. $S_\Delta(x)$ is well-defined in terms of the mesh Δ, the end condition parameters π_0 and π_k, and the ordinate vector $\mathbf{Y}$. For the purpose of this study, Δ, π_0 and π_k will be treated as given, and $\mathbf{Y}$ will be treated as the unknown population vector to be estimated by the least squares procedure. Fortunately as was seen in (2.13), $S_\Delta(x)$ is linear in $\mathbf{Y}$, and so estimating $\mathbf{Y}$ is easily done.

Model: Let $\Delta = \{x_0 < x_1 < \cdots < x_k\}$ be a given mesh, π_0 and π_k given end condition parameters, and $\mathbf{Y} = [y_0, y_1, \cdots, y_k]'$ an unknown vector of population ordinates corresponding to the knots of Δ. Also let $\boldsymbol{\xi} = [\xi_1, \xi_2, \cdots, \xi_n]'$ and $\boldsymbol{\eta} = [\eta_1, \eta_2, \cdots, \eta_n]'$ be vectors of $n \geq k+1$ data observations on the independent variable ξ and the dependent variable η, and let $\mathbf{W}$ be the matrix defined by (2.14) with rank $(\mathbf{W}) = k+1$. Lastly, let $\boldsymbol{\varepsilon} = [\epsilon_1, \epsilon_2, \cdots, \epsilon_n]'$ be a vector of independent normally distributed disturbance terms such that

$$E(\boldsymbol{\varepsilon}) = \mathbf{0},$$

$$E(\boldsymbol{\varepsilon\varepsilon}') = \sigma^2\mathbf{I},$$

where $\mathbf{0}$ is the null column vector of length n and $\mathbf{I}$ is an $n \times n$ identity matrix.

Then defining the true regression cubic spline as

$$S_\Delta(\xi) = \mathbf{WY},$$

the data observations are said to satisfy a *cubic spline regression model* (CSRM) if and only if

$$\boldsymbol{\eta} = S_\Delta(\xi) + \boldsymbol{\varepsilon} = \mathbf{WY} + \boldsymbol{\varepsilon}.$$

The least squares cubic spline corresponding to the CSRM is the cubic spline

$$\hat{S}_\Delta(\xi) = \mathbf{W\hat{Y}},$$

where $\mathbf{\hat{Y}}$ is the least squares estimator of $\mathbf{Y}$ determined by minimizing

$$[\boldsymbol{\eta} - \hat{S}_\Delta(\xi)]'[\boldsymbol{\eta} - \hat{S}_\Delta(\xi)] = (\boldsymbol{\eta} - \mathbf{W\hat{Y}})'(\boldsymbol{\eta} - \mathbf{W\hat{Y}}) = \mathbf{e}'\mathbf{e}.$$

Using conventional least squares estimating procedures, $\mathbf{\hat{Y}}$ is known to be equal to

$$\begin{aligned} \mathbf{\hat{Y}} &= (\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\boldsymbol{\eta} \\ &= \boldsymbol{\Omega}\mathbf{W}'\boldsymbol{\eta}, \end{aligned} \tag{3.1}$$

where

$$\boldsymbol{\Omega} = (\mathbf{W}'\mathbf{W})^{-1}. \tag{3.2}$$

It is also known that $\mathbf{\hat{Y}}$ is normally distributed, that $\mathbf{\hat{Y}}$ is the best linear unbiased estimator of $\mathbf{Y}$, and that the variance-covariance matrix of $\mathbf{\hat{Y}}$ is

$$\begin{aligned} E(\mathbf{Y} - \mathbf{\hat{Y}})(\mathbf{Y} - \mathbf{\hat{Y}})' &= \boldsymbol{\Omega}\mathbf{W}'E(\boldsymbol{\varepsilon\varepsilon}')\mathbf{W}\boldsymbol{\Omega} \\ &= \sigma^2\boldsymbol{\Omega}. \end{aligned}$$

Finally, an unbiased estimator of σ^2 is

$$s^2 = \frac{\mathbf{e}'\mathbf{e}}{n-k-1},$$

and a $1-\alpha$ confidence interval for $y_j (j = 0, 1, \cdots, k)$ is

$$\hat{y}_j - t_{n-k-1}s_0 \leq y_j \leq \hat{y}_j + t_{n-k-1}s_0, \tag{3.3}$$

where $s_0 = s\sqrt{\omega_{jj}}$, ω_{jj} is the diagonal element in row j of $\boldsymbol{\Omega}$, and t_{n-k-1} is the value of the student t statistic with $n-k-1$ degrees of freedom which cuts off $\alpha/2$ of the area of the distribution in each tail.

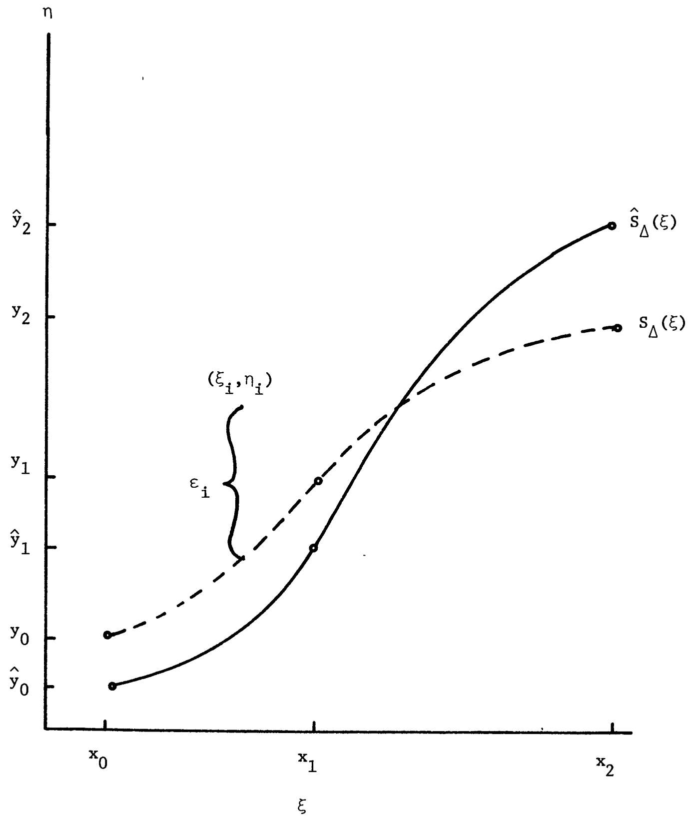
To help visualize this model, Figure B illustrates the sample case in which $k = 2$. It is interesting to note that the parameters being estimated are actual ordinate values (not slopes) of the true regression functions $S_\Delta(\xi)$ for particular abscissa values, namely the knots. In simple linear regression only the constant term has this interpretation, and since it corresponds to an abscissa value of zero, it is often far removed from the data.

4. ORTHOGONALITY OF THE TRANSFORMED VARIABLES

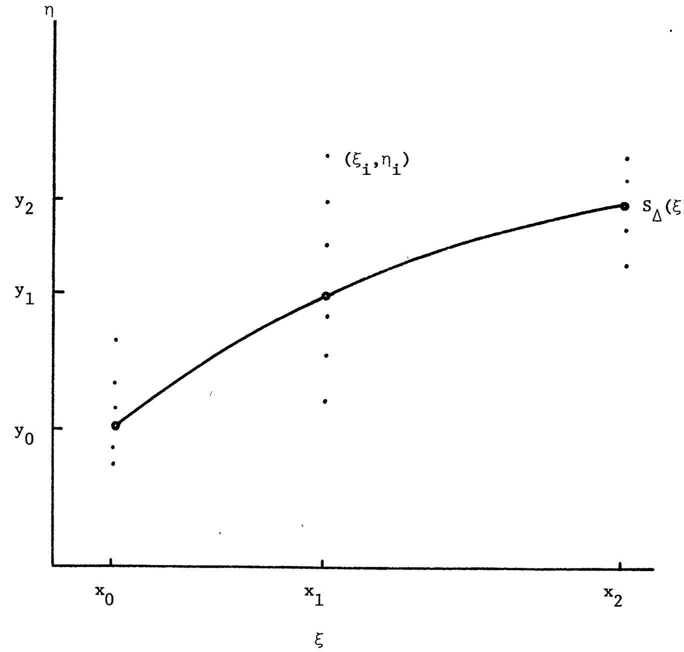
The case of orthogonality in the transformed variables (columns of $\mathbf{W}$) is of particular interest for two reasons: (1) $\mathbf{W}'\mathbf{W}$ is diagonal and hence easily inverted, and (2) the variance-covariance matrix $\boldsymbol{\Omega}$ of $\mathbf{\hat{Y}}$ is diagonal implying that the covariances between the least squares estimators are zero. While orthogonality is clearly the exception rather than the rule, it is still relevant since other cases can be viewed in terms of how much they depart from this special case.

Intuitively, it would seem that $\mathbf{Y}$ should be estimated best when the data observations ξ on the independent variable are very close to the knots. The extreme case in which each ξ_i $(i = 1, 2, \cdots, n)$ equals one of the knots

B. EXAMPLE OF A CUBIC SPLINE REGRESSION MODEL



C. EXAMPLE OF ORTHOGONALITY IN THE TRANSFORMED VARIABLES



x_j $(j = 0, 1, \cdots, k)$, and for each knot there exists at least one ξ_i equal to it, is depicted in Figure C.

In such instances the matrix $\mathbf{P}$ defined by (2.11) is a zero matrix, and the matrix $\mathbf{Q}$ defined by (2.12) takes the form

$$q_{im} = \begin{Bmatrix} 1, & \text{if} \quad m = j \\ 0, & \text{otherwise} \end{Bmatrix}.$$

Thus $\mathbf{W} = \mathbf{Q}$ and

$$\mathbf{W'W} = \begin{bmatrix} n_0 & 0 & \cdots & 0 \\ 0 & n_1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & n_k \end{bmatrix},$$

where $n_j > 0$ $(j = 0, 1, \cdots, k)$ is the number of ξ_i's equal to x_j and $n = n_0 + n_1 + \cdots + n_k$.[9]

Inverting $\mathbf{W'W}$ gives

$$\mathbf{\Omega} = \begin{bmatrix} \dfrac{1}{n_0} & 0 & \cdots & 0 \\ 0 & \dfrac{1}{n_1} & \cdots & 0 \\ \vdots & \vdots & \cdot & \vdots \\ 0 & 0 & \cdots & \dfrac{1}{n_k} \end{bmatrix},$$

and substituting into (3.1) yields the least squares estimator

$$\hat{\mathbf{Y}} = [\bar{\eta}_0, \bar{\eta}_1, \cdots, \bar{\eta}_k]', \tag{4.1}$$

where $\bar{\eta}_j$ $(j = 0, 1, \cdots, k)$ is the arithmetic mean of all the dependent variable observations at x_j.

It is obvious from (4.1) that orthogonality of the transformed variables implies that the least squares estimator $\hat{\mathbf{Y}}$ is *locally dependent* on the dependent variable observation vector $\boldsymbol{\eta}$, i.e., $\hat{y}_j$ is completely determined by the n_j observations at x_j.[10] As the ξ_i's depart from the knots this orthogonality is of course lost. However, for relatively small departures some degree of local dependence should be retained. Being more specific is not in general possible, but for individual cases it is always possible to read off the jth row $(j = 0, 1, \cdots, k)$ of the matrix $\mathbf{\Omega W'}$ to determine whether the η_i's which are assigned the greatest weights correspond to ξ_i's lying close to x_j.[11]

One further interpretation of near local dependence is that $\mathbf{\Omega}$ tends to take on the form of a dominant main diagonal matrix in which the off-diagonal elements decrease rapidly in absolute value as they move away from the main diagonal. In such cases there is little covariance among the least squares estimators $\hat{y}_j$ $(j = 0, 1, \cdots, k)$. Hence, cases in which the abscissa values of the data points tend to be in clusters and in which the knots are chosen to be centered in the cluster, are most likely to approach orthogonality, and hence the desirable estimating properties that characterize it.

[9] Note that the model has reduced to a one-way analysis of variance.

[10] Local dependence is discussed in a slightly different context in [12, p. 1100]. While his Theorem 1 [12, p. 1104] is not applicable to the cubic spline since $S'_\Delta(x)$ is continuous, it is applicable to a linear spline which has its knots "optimally located."

[11] Powell [20] does however obtain some interesting results for *uniform meshes* (equally spaced knots) and the continuous L_2 norm.

5. PREDICTION INTERVALS FOR η_0 AND $E(\eta_0)$

Suppose $x_{j-1} \leq \xi_0 \leq x_j$ and let $\mathbf{w}_0 = \mathbf{p}\Lambda^{-1}\boldsymbol{\theta} + \mathbf{q}$, where $\mathbf{p}$ and $\mathbf{q}$ are row vectors defined analogous to (2.11) and (2.12), respectively. Then, since the CRSM can be equivalently viewed as a standard linear multiple regression model whose independent variables are the transformed columns of $\mathbf{W}$, conventional theory states that at a confidence level of $1 - \alpha$, the prediction interval for an individual value η_0 corresponding to ξ_0 is

$$\hat{S}_\Delta(\xi_0) - t_{n-k-1}s_1 \leq \eta_0 \leq \hat{S}_\Delta(\xi_0) + t_{n-k-1}s_1, \qquad (5.1)$$

and the prediction interval for the mean $E(\eta_0)$ is

$$\hat{S}_\Delta(\xi_0) - t_{n-k-1}s_2 \leq E(\eta_0) \leq \hat{S}_\Delta(\xi_0) + t_{n-k-1}s_2, \qquad (5.2)$$

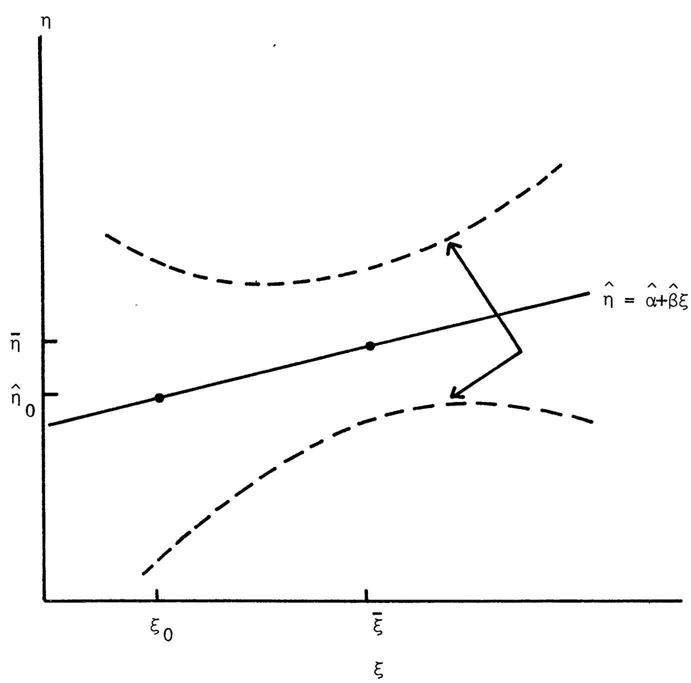
where $s_1 = s(1 + \mathbf{w}_0\boldsymbol{\Omega}\mathbf{w}_0')^{\frac{1}{2}}$ and $s_2 = s(\mathbf{w}_0\boldsymbol{\Omega}\mathbf{w}_0')^{\frac{1}{2}}$.

Of interest is the behavior of these prediction intervals in terms of ξ_0, not the transformed variables. In simple linear regression analysis of η on ξ, prediction intervals take on the form of a hyperbola such as the one shown in Figure D.

The important observation concerning such prediction intervals is that they necessarily become larger as ξ_0 departs from the arithmetic mean $\bar{\xi} = (\xi_1 + \xi_2 + \cdots + \xi_n)/n$. This is not necessarily the case with prediction intervals (5.1) and (5.2).

Since prediction interval (5.2) reduces to confidence interval (3.3) when $\xi_0 = x_j$, and since the bounds of the prediction interval are continuous in ξ_0, the behavior of the prediction interval is largely dependent on the degree of accuracy in which each of the y_j's is estimated. Furthermore, since it was seen in Section 4 that for the orthogonal case this accuracy depends solely on the number of observations at each knot, it is possible for y_0 and y_k to be estimated most accurately. In such a case Figure D would not be accurate at all, and the prediction interval would be the largest for some intermediary $x_0 < \xi_0 < x_k$. Similar remarks hold for prediction interval (5.1).

D. SIMPLE LINEAR REGRESSION PREDICTION INTERVALS



6. HYPOTHESIS TESTING FOR LINEAR AND QUADRATIC SEGMENTS

Since it is common practice in regression analysis to utilize linear (in variables as well as in parameters) functions, a relevant question is whether a particular segment (piece) of the cubic spline reduces to a linear polynomial. While $S_\Delta(x)$ cannot take on the form of a piecewise linear function (except for the trivial case of one linear polynomial over $[x_0,x_k]$), due to the continuity requirements on $S_\Delta''(x)$, the existence of linear segments might be interpreted as justifying standard regression procedures over the relevant intervals.

Fortunately, the test for the linearity of $S_\Delta(x)$ over $[x_{j-1}, x_j]$ $(j = 1, 2, \cdots, k)$ is greatly simplified by the piecewise linear nature of $S_\Delta''(x)$. In particular, $M_{j-1} = M_j = 0$ implies that $S_\Delta''(x) = 0$ for $x_{j-1} \leq x \leq x_j$, and hence that $S_\Delta(x)$ is linear. Thus the relevant hypotheses are:

$$H_0: \quad M_{j-1} = M_j = 0,$$
$$H_1: \quad M_{j-1} \neq 0 \quad \text{or} \quad M_j \neq 0.$$

To test this null hypothesis, define the $2 \times (k + 1)$ matrix $\mathbf{R}$ which consists of the $(j - 1)$th and jth rows of $\Lambda^{-1}\boldsymbol{\theta}$. Then using standard hypothesis testing procedures, H_0 can be rejected if

$$[\hat{M}_{j-1}, \hat{M}_j](\mathbf{R}\boldsymbol{\Omega}\mathbf{R}')^{-1}[\hat{M}_{j-1}, \hat{M}_j]'/2s^2 \geq F_{2,n-k-1},$$

where $[\hat{M}_{j-1}, \hat{M}_j]' = \mathbf{R}\hat{\mathbf{Y}}$ and $F_{2,n-k-1}$ is the value of the F-statistic with 2 and $(n - k - 1)$ degrees of freedom at a confidence level of $1 - \alpha$. Rejecting H_0 implies that at a confidence level of $1 - \alpha$, $S_\Delta(x)$ is not a linear polynomial over $[x_{j-1}, x_j]$.

A completely analogous test can be derived for detecting quadratic segments. $S_\Delta(x)$ reduces to a quadratic over $[x_{j-1}, x_j]$ if and only if $M_{j-1} = M_j$. Hence, the proper hypotheses are:

$$H_0': \quad M_{j-1} = M_j,$$
$$H_1': \quad M_{j-1} \neq M_j.$$

Letting $\mathbf{r}$ be the row vector formed by subtracting row $j - 1$ from row j in $\Lambda^{-1}\boldsymbol{\theta}$, then H_0' can be rejected if

$$(\hat{M}_j - \hat{M}_{j-1})^2/s^2(\mathbf{r}\boldsymbol{\Omega}\mathbf{r}') \geq F_{1,n-k-1}. \qquad (6.1)$$

In words, rejecting H_0' implies that at a level of confidence of $1 - \alpha$, $S_\Delta(x)$ is not a quadratic over $[x_{j-1}, x_j]$. Clearly, if H_0' is rejected, then H_0 must be rejected.

7. HYPOTHESIS TESTING FOR STRUCTURAL CHANGES

The concept of "structural change" has been used in the economic literature in two distinct contexts. The first context, and also the more common, uses it in a historical or temporal sense. In such contexts the "change" that takes place involves altering the environment in which the relationship holds. In such instances time is often an independent variable. The example used in the Introduction involving the effects of World War II on bond sales falls into this category.

The second context is non-temporal in nature and arises from the independent variable reaching a certain prescribed level. For example, considering an individual's income as a function of the number of formal years of education completed, it is reasonable to expect a threshold effect after twelve or sixteen years as a result of receiving a diploma. In other words, completing, say, the twelfth year affects income in two ways: (1) another year of school is completed, and (2) a diploma is received. Additional years of schooling beyond twelve are likely to affect the individual differently because he is then a high school graduate. Similar effects seem likely when age is the independent variable—obvious cases being reaching adulthood or retirement age.

Before constructing a test to determine structural changes in a cubic spline, it is necessary to clear up one small matter. Unless there is no jump discontinuity at x_j ($j = 1, 2, \cdots, k-1$), $S_\Delta'''(x_j)$ is not well-defined since $S_\Delta''(x)$ has a corner at x_j. Thus the following convention will be adopted: define

$$S_\Delta'''(x_j) = S_\Delta'''(x_j^-) \quad (j = 1, 2, \cdots, k)$$

and $S_\Delta'''(x_0) = S_\Delta'''(x_0^+)$. Then in general the graph of $S_\Delta'''(x)$ will look something like Figure E.

As can be seen from this graph, $S_\Delta'''(x)$ is in general a step function, that is, jump discontinuities in $S_\Delta'''(x)$ are permitted at the interior knots. Over any two adjacent intervals a cubic spline is equivalent to a cubic polynomial if and only if $S_\Delta'''(x)$ is continuous over these two intervals. Hence, if there is in fact a structural change in either of the previously mentioned contexts at x_j ($j = 1, 2, \cdots, k-1$), it seems reasonable to expect the jump discontinuity in the third derivative at x_j to be significantly different from zero.

Differentiating $S_\Delta''(x)$ given by (A.1) and using the preceding conventions, the following expression for $S_\Delta'''(x)$ is obtained:

$$S_\Delta'''(x) = (M_j - M_{j-1})/h_j, \quad x_{j-1} < x \leq x_j.$$

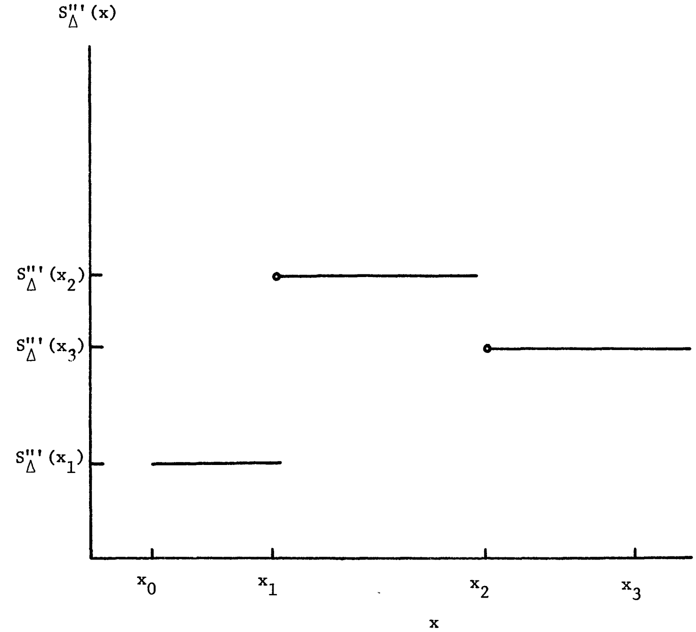
Letting $\mathbf{g}_i$ ($i = j-1, j, j+1$) be the ith row of $\mathbf{\Lambda}^{-1}\boldsymbol{\theta}$, the jump discontinuity τ_j in $S_\Delta'''(x)$ at x_j ($j = 1, 2, \cdots, k-1$) can be written as

$$\begin{aligned}\tau_j &= S_\Delta'''(x_j^+) - S_\Delta'''(x_j^-)\\ &= [(M_{j+1} - M_j)/h_{j+1}] - [(M_j - M_{j-1})/h_j]\\ &= \left[\frac{1}{h_{j+1}}(\mathbf{g}_{j+1} - \mathbf{g}_j) - \frac{1}{h_j}(\mathbf{g}_j - \mathbf{g}_{j-1})\right]\mathbf{Y}\\ &= \mathbf{gY},\end{aligned}$$

where the row vector $\mathbf{g}$ is defined by

$$\mathbf{g} = \left[\frac{1}{h_{j+1}}(\mathbf{g}_{j+1} - \mathbf{g}_j) - \frac{1}{h_j}(\mathbf{g}_j - \mathbf{g}_{j-1})\right].$$

E. EXAMPLE OF $S_\Delta'''(x)$



To test for structural change at x_j (equivalently, whether the jump discontinuity τ_j is significantly different from zero), construct the hypotheses

$$H_0'': \quad \tau_j = 0,$$
$$H_1'': \quad \tau_j \neq 0.$$

Then H_0'' (no structural change) can be rejected at a level of confidence of $1 - \alpha$ if

$$\hat{\tau}_j^2/s^2(\mathbf{g\Omega g}') \geq F_{1,n-k-1}, \tag{7.1}$$

where $\hat{\tau}_j = \mathbf{g\hat{Y}}$ is the estimated jump discontinuity in the third derivative at x_j.

While this test is easily constructed, its interpretation must be performed with care. If H_0'' is rejected, this implies that a level of confidence of $1 - \alpha$ can be attached to the statement that there is structural change at the knot in question. Furthermore, this gives strong support to the choice of a piecewise estimating function at the knot in question. However, this does *not* imply that the knot in question is optimally located in the sense that a small change in its position might not result in a smaller least squares error sum, i.e., give a better fit. Optimally locating knots in terms of minimizing the least squares error sum is a curve fitting problem in which the knots are stripped of any theoretical underpinnings they may have had. Such an issue is not of concern in this study.

If H_0'' cannot be rejected, it is not always clear that this should necessarily imply its acceptance, especially when there exist strong *a priori* reasons for believing there to be structural change at x_j. For example, due to the local dependence of the cubic spline, the inability to reject H_0'' might be due to a lack of data observations near x_j. While not always possible, observations may be added to correct this. Another possible explanation is that while there exists strong *a priori* evidence for structural change at x_j, it may be due to lag effects that such a change has not been uncovered. If the theory involved is compatible with lags, then it seems reasonable to consider an alternative choice for x_j which takes this into account. In such cases where there exists theoretical justification for selecting new knots, the knots will be referred to as *quasi-fixed knots*. This terminology is intended to imply their close resemblance to the fixed knots previously discussed, and to imply their important differences from variable (free) knots which have no theoretical support.

Still another possible explanation for the inability to reject H_0'' might be the inclusion of some other knot x_i $(i \neq j)$ which does not have as strong a theoretical basis. If k is sufficiently large $(k \geq 3)$, then it might be reasonable to treat the knots as quasi-fixed, exclude x_i altogether, and retest H_0''.[12] The added degree of freedom and the resulting change $\hat{r}_j$ might then be sufficient to reject H_0''.

8. AN EMPIRICAL EXAMPLE OF STRUCTURAL CHANGE

This section is based on an article by Barzel [2] dealing with the growth of technical progress in the Indianapolis 500 Race since its conception in 1911. By equating technical advance with progress in the winning speed, and using time as the independent variable, Barzel investigates the effect of the two World Wars on racing performance. In his own words [2, p. 74]: "The two World Wars may have affected racing through war-related technical advance on the one hand and through the interruption of the activity on the other. To the extent that progress is due to 'learning by doing,' the war years were lost to racing."

Although time is not easily thought of as a "causal" variable, Barzel points out that its high correlation with such causal factors as improved cars, experience, etc. justifies its use as a proxy for these variables. Also, the choice of the winning speed as a measure of technical progress seems appropriate and has the added advantage of being measured with great accuracy (usually six digits). However, unlike Barzel, this author feels that these two variables are related in a continuous manner over the entire period *including* the War years. The fact that no racing occurred during these years (1917–18, 1942–45) should not affect continuity any more than the fact that there was no racing in between any of the data observations, yet the conventional dummy variable approach Barzel uses necessitates "magical" shift points in which technical progress *instantaneously* jumps up or down.

Furthermore, it seems that a more flexible functional form than the linear, inverse, double logarithmic, or semilogarithmic forms considered by Barzel should be used. While there exists some theoretical basis for the double logarithmic model [10], Barzel finds it somewhat unsatisfactory. The semilogarithmic form which Barzel prefers implies the proportionate increase in speed is constant over time, and it is not clear whether such an assumption is justifiable. The Great Depression most likely had an impact on many of the causal variables (e.g., research expenditure) and this is reflected in the stagnation of winning times during this period (especially 1932–34), even though learning by doing was going on.

As an alternative to Barzel's models, the CSRM is an attractive choice. Letting ξ_i $(i = 1, 2, \cdots, 55)$ denote the race year minus 1910 and η_i the observed winning speed of that race (see the table), divide the racing period 1911–71 into three subperiods by the mesh $\Delta = \{1, 7.5, 33.5, 61\}$, where the knots are chosen midway through the non-racing war years.[13] Furthermore, let $\mathbf{Y} = [y_0, y_1, y_2, y_3]$ be the vector of theoretical winning speeds (unknown population parameters) corresponding to the knots, and let $\pi_0 = \pi_3 = 2$. Then letting ϵ_i $(i = 1, 2, \cdots, 55)$ denote independent normal disturbance terms with zero mean and constant variance σ^2, the following CSRM may be specified:

$$\eta_i = S_\Delta(\xi_i) + \epsilon_i \quad (i = 1, 2, \cdots, 55). \tag{8.1}$$

Applying ordinary least squares to (8.1) yields the estimated ordinate vector

$$\hat{\mathbf{Y}} = [76.03, 86.71, 116.1, 158.6]',$$

and the corresponding variance-covariance matrix

$$s^2\Omega = \begin{bmatrix} 2.259 & .6668 & -.007446 & -.2784 \\ .6668 & .5093 & -.02960 & .1420 \\ -.007446 & -.02960 & .3681 & -.3268 \\ -.2784 & .1420 & -.3268 & 1.651 \end{bmatrix}.$$

Using (2.10) to compute the estimates of the second derivative at the knots, (A.3) can be rewritten in piecewise cubic form to give the estimated cubic spline

$$\hat{S}_\Delta(x) = \begin{cases} 74.09 + 1.994x - .05058x^2 + .001204x^3, & 1 \leq x \leq 7.5 \\ 74.41 + 1.867x - .03356x^2 + .0004480x^3, & 7.5 \leq x \leq 33.5 \\ 86.03 + .8260x - .002501x^2 + .001389x^3, & 33.5 \leq x \leq 61 \end{cases},$$

which is shown graphically in Figure F. The coefficient of determination for the model (in terms of the percent of raw variance explained) is $R^2 = .9995$, and the standard error of the estimate is 2.793.

In Barzel's model the effects of the wars are determined by testing whether the slopes of the appropriate dummy variables are significantly different from zero. He concludes that they all are and hence that the decline in the trend and rate of increase of winning speeds after the wars is the result of the absence of "learning by doing" during the war years.

As seen in Section 7, in the CSRM the effects of the wars are determined by testing whether any two adjacent

[12] Such a procedure is akin to the popular technique of dropping variables with insignificant coefficients from a regression equation. Any criticisms of the latter technique apply here as well.

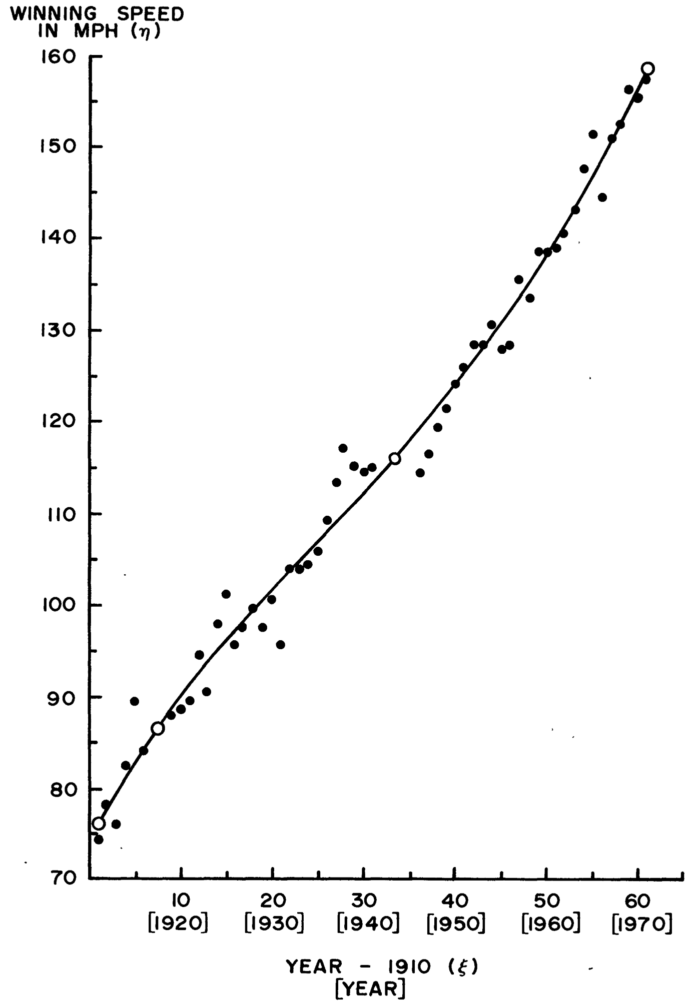
[13] Knots were also tried at the beginning and end of the non-racing war years without significantly changing the results. The data include the years 1970–71, whereas Barzel's results are based on the period 1911–69.

WINNING SPEEDS FOR THE INDIANAPOLIS 500 RACE

Year	Winning Speed	Year	Winning Speed	Year	Winning Speed
1911	74.59	1932	104.114	1955	128.209
1912	78.72	1933	104.162	1956	128.490
1913	75.931	1934	104.863	1957	135.601
1914	82.47	1935	106.240	1958	133.791
1915	89.84	1936	109.069	1959	138.857
1916	84.00	1937	113.580	1960	138.767
1919	88.05	1938	117.200	1961	139.130
1920	88.62	1939	115.035	1962	140.293
1921	89.62	1940	114.277	1963	143.137
1922	94.48	1941	115.117	1964	147.350
1923	90.95	1946	114.820	1965	151.388
1924	98.23	1947	116.338	1966	144.317
1925	101.13	1948	119.814	1967	151.207
1926	95.904	1949	121.327	1968	152.882
1927	97.545	1950	124.002	1969	156.867
1928	99.482	1951	126.244	1970	155.749
1929	97.585	1952	128.922	1971	157.735
1930	100.448	1953	128.740		
1931	96.629	1954	130.840		

Source: For years 1911–60, Bloemker [5, pp. 271-7] and for years 1961–71, [26, p. 901].

F. WINNING SPEEDS, OBSERVED AND PREDICTED, FOR THE INDY 500



polynomial segments can be said to be significantly different in a meaningful sense, i.e., if the jump discontinuities in $S_\Delta'''(x)$ at the interior knots $x_1 = 7.5$ and $x_2 = 33.5$ are significantly different from zero, then it can be concluded that the absence of racing during war years had an effect on technical progress.

The F-ratios corresponding to (7.1) are 7.743 for $x_1 = 7.5$ (the World War I interruption) and 9.721 for $x_2 = 33.5$ (the World War II interruption). Since both of these statistics are significant at the one percent level, and since both of the estimated jump discontinuities $\hat{r}_1 = -.004537$ and $\hat{r}_2 = -.001854$ are negative, the cubic spline regression model gives additional support to Barzel's conclusion that technical progress, as measured by winning speeds, suffered as a result of the non-racing war years.

Last, it is of interest to test if each of cubic segments can be said to be significantly different from a quadratic or linear segment. The F-ratios corresponding to (6.1) are 9.459 for the first interval, 13.49 for the second interval, and 24.62 for the third interval. Since all these statistics are significant at the one percent level, it can be concluded that none of the segments are quadratic.

9. CONCLUSION AND EXTENSIONS

In summary, this shows that the cubic spline can serve as a useful mechanism for viewing structural change in continuous regression models. Linearity in the unknown parameters facilitates both the estimation and hypothesis testing procedures outlined in the text.

While attention has centered on the simple case in which there are no other independent variables in the model, the results extend directly to a multiple regression framework in which the columns of the transformed data matrix $\mathbf{W}$ in (2.14) represent only a subset of the independent variables in the model and also to models in which some transformation of the dependent variable is a cubic spline [16, 17].

Extensions of the piecewise framework to more than one dimension cause no difficulties as long as the different independent variables are not allowed to interact. However, incorporating interaction terms greatly complicates things. Bicubic splines were first investigated by de Boor [6] in an exact approximation setting. However, least squares bicubic splines have not received much attention. Bilinear splines which do involve an interaction term have been investigated in detail in [18, 19].

APPENDIX

Since the second derivative of a cubic polynomial is linear, then using the notation of Section 2, for $x_{j-1} \leq x \leq x_j$ $(j = 1, 2, \cdots, k)$

the two point equation of a straight line yields

$$S_\Delta''(x) = \left[\frac{x_j - x}{h_j}\right]M_{j-1} + \left[\frac{x - x_{j-1}}{h_j}\right]M_j. \quad (A.1)$$

Integrating (A.1) twice and evaluating the constants of integration by imposing the interpolation conditions (3) of Section 2 yields for $x_{j-1} \leq x \leq x_j$ $(j = 1, 2, \cdots, k)$

$$S_\Delta'(x) = \left[\frac{h_j}{6} - \frac{(x_j - x)^2}{2h_j}\right]M_{j-1} + \left[\frac{(x - x_{j-1})^2}{2h_j} - \frac{h_j}{6}\right]M_j + \frac{y_j - y_{j-1}}{h_j}, \quad (A.2)$$

$$S_\Delta(x) = \frac{x_j - x}{6h_j}[(x_j - x)^2 - h_j^2]M_{j-1} + \frac{x - x_{j-1}}{6h_j}[(x - x_{j-1})^2 - h_j^2]M_j + \left[\frac{x_j - x}{h_j}\right]y_{j-1} + \left[\frac{x - x_{j-1}}{h_j}\right]y_j. \quad (A.3)$$

The only quantities in (A.1)–(A.3) which are unknown are the moments M_j $(j = 0, 1, \cdots, k)$, and with a little effort these can be determined. From (A.2) the one-sided limits of the derivative evaluated at the knots x_j $(j = 1, 2, \cdots, k - 1)$ are seen to be

$$S_\Delta'(x_j^-) = h_jM_{j-1}/6 + h_jM_j/3 + (y_j - y_{j-1})/h_j, \quad (A.4)$$

$$S_\Delta'(x_j^+) = - h_{j+1}M_j/3 - h_{j+1}M_{j+1}/6 + (y_{j+1} - y_j)/h_{j+1}. \quad (A.5)$$

Since $S_\Delta'(x)$ is continuous, (A.4) and (A.5) can be equated to give the continuity conditions (2.1) used in calculating the moments.

[Received May 1972. Revised April 1973.]

REFERENCES

[1] Ahlberg, J.H., Nilson, E.N. and Walsh, J.L. *The Theory of Splines and Their Applications*, New York: Academic Press, Inc., 1967.

[2] Barzel, Yoram, "The Rate of Technical Progress: The 'Indianapolis 500'," *Journal of Economic Theory*, 4 (February 1972), 72–81.

[3] Bellman, Richard and Roth, Robert, "Curve Fitting by Segmented Straight Lines," *Journal of the American Statistical Association*, 64 (September 1969), 1079–84.

[4] Blischke, W.R., "Least Squares Estimates of Two Intersecting Lines," Technical Report No. 7, Department of Navy, ONR, Cornell University, 1961.

[5] Bloemker, Al, *500 Miles to Go, The Story of the Indianapolis Speedway*, New York: Coward-McCann, Inc., 1961.

[6] de Boor, Carl E., "Bicubic Spline Interpolation," *Journal of Mathematics and Physics*, 41 (1962), 212–18.

[7] ——— and Rice, John R., "Least Squares Cubic Spline Approximation I—Fixed Knots," Computer Science Department, Division of Mathematical Sciences, Purdue University, CSD TR 20, 1968.

[8] ———, "Least Squares Cubic Spline Approximation II—Variable Knots," Computer Science Department, Division of Mathematical Sciences, Purdue University, CSD TR 21, 1968.

[9] Farley, John U. and Hinich, Melvin J., "Testing for a Shifting Slope Coefficient in a Linear Model," *Journal of the American Statistical Association*, 65 (September 1970), 1320-9.

[10] Fellner, William, "Specific Interpretation of Learning by Doing," *Journal of Economic Theory*, 1 (August 1969), 119–40.

[11] Garber, Steven G. and Poirier, Dale J., "Determinants of Aerospace Profit Rates 1951–71," Social Systems Research Institute Workshop Paper No. 7308, University of Wisconsin, Madison, Wisconsin, 1973.

[12] Hudson, Derek J., "Fitting Segmented Curves Whose Join Points Have To Be Estimated," *Journal of the American Statistical Association*, 61 (December 1966), 1097–129.

[13] Kershaw, D., "A Note on the Convergence of Interpolating Cubic Splines," *SIAM Journal of Numerical Analysis*, 8 (March 1971), 67–74.

[14] McGee, Victor E. and Carleton, Willard T., "Piecewise Regression," *Journal of the American Statistical Association*, 65 (September 1970), 1109–24.

[15] Pennington, Ralph H., *Introductory Computer Methods and Numerical Analysis*, New York: The Macmillan Company, 1965.

[16] Poirier, Dale J., "Structural Change in Continuous Multiple Regression Models," Unpublished manuscript, 1972.

[17] ———, "Structural Change in Growth Models," Unpublished manuscript, 1972.

[18] ———, "Multiple Regression Using Bilinear Splines," Social Systems Research Institute Workshop Paper No. 7230, University of Wisconsin, Madison, Wisconsin, 1972.

[19] ———, "Cobb-Douglas Spline Production Functions," Social Systems Research Institute Workshop Paper No. 7231, University of Wisconsin, Madison, Wisconsin, 1972.

[20] Powell, M.J.D., "The Local Dependence of Least Squares Cubic Splines," *SIAM Journal of Numerical Analysis*, 6 (September 1969), 398–413.

[21] Quandt, Richard E., "The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes," *Journal of the American Statistical Association*, 53 (December 1958), 873–80.

[22] ———, "Tests of the Hypothesis that a Linear Regression System Obeys Two Separate Regimes," *Journal of the American Statistical Association*, 55 (June 1960), 324–30.

[23] Robison, D.E., "Estimates for the Points or Intersection of Two Polynomial Regressions," *Journal of the American Statistical Association*, 59 (March 1964), 214–24.

[24] Schoenberg, I.J., "Contributions to the Problem of Approximation of Equidistant Data by Analytic Functions. Part A—On the Problem of Smoothing or Graduation. A First Class of Analytic Approximation Formulae," *Quarterly Journal of Applied Mathematics*, 4 (April 1946), 45–99, "Part B—On the Problem of Osculatory Interpolation. A Second Class of Analytical Approximation Formulae," *Quarterly Journal of Applied Mathematics*, 4 (July 1946), 112–41.

[25] ———, "On Spline Functions," in Oved Shisha ed., *Inequalities*, New York: Academic Press, Inc., 1967, 255–91.

[26] *The 1973 World Almanac and Book of Facts*, New York: Newspaper Enterprise Association, 1972.