Neighbourhood 'Correlation Ratio' Curves

By

Kjell Doksum*
Department of Statistics
367 Evans Hall
University of California at Berkeley
Berkeley, California 94720

and

Sorana Froda**
Département de mathématiques
Université du Québec à Montréal
C.P. 8888, Succ. "A"
Montréal, Québec, Canada, H3C 3P8

Technical Report No. 412
April 1994

*Research partially supported by NSF grant DMS-9106752
**Research supported by a grant from NSERC, Canada

Department of Statistics
University of California
Berkeley, California 94720

NEIGHBOURHOOD 'CORRELATION RATIO' CURVES

Kjell Doksum *

Department of Statistics
367 Evans Hall
University of California at Berkeley
Berkeley, California 94720

and

Sorana Froda **

Département de mathématiques
Université du Québec à Montréal
C. P. 8888, Succ. "A"
Montréal, Québec, Canada, H3C 3P8

* Research partially supported by NSF grant DMS-9106752.
** Research supported by a grant from NSERC, Canada.

NEIGHBOURHOOD 'CORRELATION RATIO' CURVES

Kjell Doksum and Sorana Froda

Abstract

Pearson's nonparametric R-squared, also called the *correlation ratio*, or *eta-squared*, is the ratio of the variance explained by nonparametric regression to the total variance of the response Y. We obtain a local version $\eta_h^2(x)$ of this η^2 by calibrating the conditional eta-squared obtained by restricting the explanatory variable X to an interval, $[x-h, x+h]$. $\eta_h^2(x)$ is a local measure of the explanatory power of X. Nonparametric estimators of $\eta_h^2(x)$ are introduced and their root-n distributional convergence to normal distributions is established. We propose a local bandwith selection procedure for choosing the bandwith b in the nonparametric regression function $\mu(x) = \mathbb{E}(Y|X=x)$. The procedure consists in choosing the b which maximizes the local explanatory power of X. Monte Carlo comparisons of kernel and locally quadratic approaches are presented. The locally quadratic methods only do better when $\mu(x)$ has sharp turns. Finally we illustrate our local procedures by doing a local ANOVA on a real data set.

Key words: local R-squared, nonparametric correlation, kernel regression, local polynomial regression, local ANOVA.

1. Introduction.

For experiments where the relationship between a response variable Y and a covariate X is not necessarily linear, a very useful measure of the strength of the relationship between X and Y is Pearson's *correlation ratio*

$$\eta^2 = \frac{\operatorname{Var}\mu(X)}{\operatorname{Var}Y} = 1 - \frac{\mathbb{E}\,\sigma^2(X)}{\operatorname{Var}Y}, \tag{1.1}$$

where $\mu(x) = \mathbb{E}[Y|X=x]$ and $\sigma^2(x) = \operatorname{Var}(Y|X=x)$. The coefficient η^2 is based on the ANOVA decomposition $\operatorname{Var}Y = \operatorname{Var}\mu(X) + \mathbb{E}\,\sigma^2(X)$ and thus gives the fraction of the variability of Y that can be explained by the regression $\mu(X)$; in linear models η^2 reduces to the usual (Galton-Pearson) squared correlation

$$\rho^2 = \frac{\operatorname{Cov}(X,Y)}{\operatorname{Var}X \operatorname{Var}Y}.$$

In nonlinear models, which, without loss of generality, can be written as

$$Y = \mu(X) + \sigma(X)\epsilon\ ,\ X \text{ and } \epsilon \text{ uncorrelated}\ , \tag{1.2}$$

η^2 is a better measure of the strength of the relationship between X and Y than ρ^2. In particular, there are many models with X and Y strongly related, where $\rho^2 = 0$ while η^2 gives a good measure of the relationship (see Rényi, 1959; more recently, Doksum and Samarov, 1994, discuss the properties of η^2).

As discussed by Bjerve and Doksum, 1993, and Doksum et al., 1994, there are many studies where the strength of the relationship between X and Y is different for different values x of the covariate, and in these cases it is useful to have a local mesure of the strength of the relation. They introduced as local version of ρ^2

$$\rho^2(x) = \frac{\beta^2(x)}{\beta^2(x) + \sigma^2(x)/\sigma_1^2}\ , \tag{1.3}$$

with $\sigma_1^2 = \operatorname{Var}X$, and $\beta(x) = \mu'(x) = d\mu(x)/dx$.

QTEX

However, just as ρ^2 can be zero when X and Y are strongly related, so can $\rho^2(x)$ be zero when X and Y are strongly related locally. For instance, if $Y = 1 + (X-2)^2 + \epsilon$, then $\rho^2(2) = 0$, provided $\mathrm{Var}\,\epsilon > 0$. This happens because $\rho(x)$ measures the strength of the *locally linear* relationship between X and Y. In this paper we consider local versions of the *correlation ratio* η^2 which pick up nonlinear local dependence between X and Y.

The local squared correlation $\rho^2(x)$ was obtained from the formula

$$\rho^2 = \frac{\beta^2}{\beta^2 + \sigma_\epsilon^2/\sigma_1^2},$$

by replacing the fixed linear model slope β with the local counterpart $\beta(x) = \mu'(x)$, and the fixed linear model residual variance σ_ϵ^2 with the local version $\sigma^2(x)$. An alternative approach to local correlation, which corresponds to the usual approach to local regression, is to use a measure based on the conditional distribution of Y given X. In this paper we introduce a local correlation measure based on the conditional version $\eta_{CO}^2(x)$ of η^2 given X in a neighbourhood of a given covariate value x of interest. This $\eta_{CO}^2(x)$ can be interpreted as the *correlation ratio* for a biased sampling plan which drives the value of $\eta_{CO}^2(x)$ towards zero. Therefore we propose to calibrate $\eta_{CO}^2(x)$ and thus correct for the biased sampling by requiring that our local measure coincide with the usual *correlation ratio* in the linear models. The details of the derivation of the calibrated measure $\eta_{CA}^2(x)$, which we also refer to as a *local (or neighbourhood) R-squared*, are given in Section 2.

Besides providing a local measure of the strength of the relationship between X and Y, estimators $\widehat{\eta}_{CA}^2(x)$ of $\eta_{CA}^2(x)$ can be used to select the bandwith $b(x)$ for the kernel estimator $\widehat{\mu}_b(x)$ of the regression curve $\mu(x)$. The idea is similar to cross-validation. That is, we consider the estimate $\widehat{\eta}_{CA}^2(x)$ as function of the bandwith b, and we select the value b which maximizes $\widehat{\eta}_{CA}^2(x)$. In other words, we view $\widehat{\eta}_{CA}^2(x)$ as a measure of the explanatory power of the covariate X in a neighbourhood of a given x, and we choose b to maximize this explanatory power. This idea is closely related to minimizing mean squared error, e.g. the supersmoother (Friedman, 1984), the bootstrap smoother (Härdle and Bowman, 1988), and LOWESS (Cleveland, 1979, Cleveland and Devlin, 1988). The advantage of our approach is that it links the local regression estimate with an estimate of the local explanatory power of the covariate.

We study also the asymptotic properties of $\widehat{\eta}_{CA}^2(x)$. Here we adapt the results of Doksum and Samarov, 1994, who investigated the asymptotic properties of estimates $\widehat{\eta}^2$ of the global measure of explanatory power η^2. We give the asymptotic mean squared error of $\widehat{\eta}_{CA}^2(x)$ and we find the asymptotic normal distribution of $\sqrt{n}\left[\widehat{\eta}_{CA}^2(x) - \eta_{CA}^2(x)\right]$, where n is the sample size. We use Monte-Carlo methods to investigate and compare various estimators based on Nadaraya-Watson kernel estimators and locally quadratic estimators of $\mu(x)$.

2. Nonparametric Correlation through Calibration of the Conditional Correlation Ratio.

Our aim is to develop a measure of the strength of the relationship between X and Y near a particular value x of the covariate. We start by considering the 'conditional' *correlation ratio* given $X \in N_h(x) = [x-h, x+h]$, where h is a number which determines the length of the interval on which the strength of the relationship is to be measured. The law of $\{X | X \in [x-h, x+h]\}$ is given by (we assume throughout that X has a density $f(x)$):

$$f_h(z) = \begin{cases} f(z)/P_h & \text{if } z \in [x-h, x+h]; \\ 0 & \text{otherwise,} \end{cases} \tag{2.1}$$

where $P_h = Pr(X \in [x-h, x+h])$. We consider the conditional η^2

$$\eta_{CO,h}^2(x) = \frac{\sigma_{\mu,h}^2(x)}{\sigma_{Y,h}^2(x)},$$

where $\mu(X) = \mathbb{E}[Y|X]$, $\sigma^2_{\mu,h}(x) = \mathrm{Var}(\mu(X)|X \in N_h(x))$, and $\sigma^2_{Y,h}(x) = \mathrm{Var}(Y|X \in N_h(x))$. Note that $\eta^2_{CO,h}(x)$ can be interpreted as the *correlation ratio* for a biased sampling plan, that is, a sampling plan where X measurements can only be obtained for X in the interval.

It is easy to see that $\eta^2_{CO,h}(x)$ is much smaller than η^2 computed for X unrestricted; in fact $\mathrm{Var}(\mu(X)|X \in N_h(x))$ tends to zero as $h \to 0$, while $\mathrm{Var}(Y|X \in N_h(x))$ does not (except in trivial cases). The proposed measure is obtained from $\eta^2_{CO,h}(x)$ via a *calibration* procedure.

Before we derive the calibrated measure, we state the following result, which follows from (2.1) by straightforward computing.

Proposition 2.1. *The conditional expectation $\mu(x)$, and the conditional variance, $\sigma^2(x)$, are preserved under restriction of X to an interval. That is: let x_0 be fixed and let (X_h, Y_h) be of law $\{(X,Y)|X \in [x_0-h, x_0+h]\}$; then, for $x \in [x_0-h, x_0+h]$:*

$$\begin{aligned}\mu_h(x) &= \mathbb{E}[Y_h|X_h = x] = \mu(x)\\ \sigma^2_h(x) &= \mathrm{Var}[Y_h|X_h = x] = \sigma^2(x).\end{aligned}$$

As a first step in the derivation of the calibrated measure consider the linear model,

$$Y = \alpha + \beta X + \epsilon\ , \tag{2.2}$$

with X and ϵ uncorrelated. For this linear case Gulliksen, 1951, ch.11, studied how the regression and correlation are affected by restricting the values of X. Gulliksen used the term *selection on the basis of X* and pointed out that the regression line of Y on X, given that X is restricted, will not be affected by a *selection* based on X. Therefore the regression lines with X restricted and unrestricted can be assumed to be the same and the two slopes to be equal; that is, in our notation

$$\rho^2\left(\frac{\sigma^2_Y}{\sigma^2_X}\right) = \eta^2_{CO,h}(x)\left(\frac{\sigma^2_{Y,h}(x)}{\sigma^2_{X,h}(x)}\right)\ , \tag{2.3}$$

where $\sigma^2_{X,h}(x) = \mathrm{Var}(X|X \in N_h(x))$. At the same time Gulliksen noted that the residual variances are the same, i.e., in our notation

$$\sigma^2_Y\left(1-\rho^2\right) = \sigma^2_{Y,h}(x)\left(1-\eta^2_{CO,h}(x)\right)\ . \tag{2.4}$$

Let $\tau^2_h(x) = \sigma^2_{X,h}(x)/\sigma^2_X$ and write $\eta^2_{CO,h}(x)$ as function of $\tau^2_h(x)$ and ρ^2. If we isolate the ratio $\sigma^2_{Y,h}(x)/\sigma^2_Y$ in both (2.3) and (2.4)

$$\frac{\sigma^2_{Y,h}(x)}{\sigma^2_Y} = \frac{1-\rho^2}{1-\eta^2_{CO,h}(x)}\ ,\quad \frac{\sigma^2_{Y,h}(x)}{\sigma^2_Y} = \frac{\rho^2\tau^2_h(x)}{\eta^2_{CO,h}(x)}\ ,$$

we obtain:

$$\eta^2_{CO,h}(x) = \frac{\rho^2\tau^2_h(x)}{\rho^2\tau^2_h(x) + (1-\rho^2)}\ . \tag{2.5}$$

In order to define the local measure we make the following point: since in the linear model (2.2) the local measure should be the correlation coefficient ρ^2 we define the local measure by solving for ρ^2 in (2.5) and letting the calibrated conditional *correlation ratio* $\eta^2_{CA,h}(x)$ be equal to the result. We obtain:

$$\begin{aligned}\eta^2_{CA,h}(x) &= \frac{\eta^2_{CO,h}(x)}{\eta^2_{CO,h}(x) + \tau^2_h(x)(1-\eta^2_{CO,h}(x))} = \frac{\eta^2_{CO,h}(x)}{\eta^2_{CO,h}(x)(1-\tau^2_h(x)) + \tau^2_h(x)}\\ &= \frac{\sigma^2_{\mu,h}(x)}{\sigma^2_{\mu,h}(x)(1-\tau^2_h(x)) + \tau^2_h(x)\sigma^2_{Y,h}(x)}\ .\end{aligned} \tag{2.6}$$

The preceding formula (2.6) is our local *correlation ratio* measure. It is the conditional *correlation ratio* given $X \in N_h(x)$, $\eta^2_{CO,h}$, calibrated to coincide with the *correlation ratio* $\eta^2 = \rho^2$ in linear models.

As a final remark: in many cases formula (2.6) can be computed explicitly, as function of $\mu(x)$ and h (see examples in Section 5 and *Appendix*). In all cases it can be easily estimated by considering only data $\{(X_i, Y_i)\}_i$, with X_i in a proper neighbourhood of x (see Section 6).

3. A Local ANOVA Decomposition

Let $\sigma_\mu^2 = \operatorname{Var}\mu(X)$ and $\sigma_{Y|X}^2 = \mathbb{E}\sigma^2(X)$. The classical ANOVA decomposition is: $\operatorname{Var} Y = \sigma_\mu^2 + \sigma_{Y|X}^2$, which leads to the global correlation ratio

$$\eta^2 = \frac{\sigma_\mu^2}{\sigma_\mu^2 + \sigma_{Y|X}^2}.$$

We next show that the local *correlation ratio* $\eta_{CA,h}^2(x)$ is based on a similar but *local* ANOVA decomposition.

Consider the decomposition $\sigma_{Y,h}^2(x) = \sigma_{\mu,h}^2(x) + \sigma_{Y|X,h}^2(x)$, where $\sigma_{Y|X,h}^2(x)$ is the conditional expected residual variance of Y given X, when $X \in N_h(x)$, that is $\sigma_{Y|X,h}^2(x) = \mathbb{E}[\operatorname{Var}(Y_h|X_h)]$, where (X_h, Y_h) is distributed as $\{(X,Y)|X \in N_h(x)\}$. Substituting this decomposition into (2.6), we obtain

$$\eta_{CA,h}^2(x) = \frac{\sigma_{\mu,h}^2(x)}{\sigma_{\mu,h}^2(x) + \tau_h^2(x)\sigma_{Y|X,h}^2(x)}. \tag{3.1}$$

The sum in the denominator of (3.1) gives a *local ANOVA decomposition* of $\{(X,Y)|X \in N_h(x)\}$; we call the denominator

$$D_h(x) = \sigma_{\mu,h}^2(x) + \tau_h^2(x)\sigma_{Y|X,h}^2(x) \tag{3.2}$$

the *local variability of* Y. Under general regularity conditions, $\tau_h^2(x) \to 1$, $\sigma_{\mu,h}^2(x) \to \sigma_\mu^2$, and $\sigma_{Y|X,h}^2(x) \to \sigma_{Y|X}^2$, when $h \to \infty$. Thus the *local ANOVA decomposition* tends to the classical ANOVA decomposition.

4. Properties of the Neighbourhood Correlation Ratio.

Assume that X has a density $f(x)$; let $\eta_h^2(x) = \eta_{CA,h}^2(x)$ and consider its properties.

1) $\eta_h^2(x) \le 1$;
2) $\eta_h^2(x)$ is invariant under linear transformations $X \mapsto a + bX$, $Y \mapsto c + dY$; this property follows from the invariance of $\eta_{CO,h}^2(x)$ and of $\tau_h^2(x)$ with respect to scale changes;
3) $\eta_h^2(x) \equiv \rho^2$ in the normal bivariate case;
4) If $\mu'(x)$ is continuous and $\mathbb{E}(Y^k|X \in N_h(x)) \to \mathbb{E}(Y^k|X = x)$, $k = 1, 2$, then, as $h \to 0$, $\eta_h^2(x) \to \rho^2(x)$, with $\rho^2(x)$ given in (1.3).
5) $\eta_h^2(x) = 0$ for all x, if X and Y are independent; indeed, $\eta_{CO,h}^2(x) = 0$ in this case, while $\sigma_{Y|X,h}^2(x) > 0$, except in trivial cases;
6) $\eta_h^2(x) = 1$ when Y is a function of X, because $\eta_{CO,h}^2(x) = 1$ and $\sigma_{Y|X,h}^2(x) = 0$;
7) If the density $f(x) > 0$, then the equality $\eta_h^2(x) \equiv 1$ for all x implies that Y is a function of X;
8) **Interchangeability**: write $\eta_{X,Y}^2(x)$ for $\eta_{CA,h}^2(x)$ as defined in Section 2; then we can define

$$\eta^2(x,y) = \sqrt{\eta_{X,Y}^2(x)}\sqrt{\eta_{Y,X}^2(y)},$$

   and obtain a measure where X and Y can be interchanged; we now need to assume that also Y has a density.
9) If $f(x)$ has infinite support, then, for $h \to \infty$, $\tau_h^2(x) \to 1$, and $\eta_h^2(x) \to \eta^2$, with η^2 given in (1.1). If $f(x)$ has as support an interval of length $2L$, then, for $h \to \max(x, L - x)$, $\tau_h^2(x) \to 1$, and $\eta_h^2(x) \to \eta^2$.
10) **Converting to a signed correlation ratio**: while $\eta_h^2(x)$ gives the strength of the relationship or the variance explained locally, the quantity $\operatorname{sign}\{\mathbb{E}(\beta(X)|X \in [x - h, x + h])\}$ indicates whether this relationship is positive or negative.
11) **Conditioning on a probability interval.** Instead of conditioning on the fixed width interval $N_h(x) = [x - h, x + h]$, we could condition on the fixed probability interval $I_\delta(x) =$

$[x_{p-\delta}, x_{p+\delta}]$, where $p = F(x)$ and x_q denotes the quantile $F^{-1}(q)$. Since $X \in I_\delta(x)$ is equivalent to $F(X) \in [p-\delta, p+\delta]$, and since $U = F(X)$ has a uniform $\mathcal{U}[0,1]$ distribution, we are thus conditioning on a uniform variable. Properties 1)–10) hold as before. In addition, note that $\eta^2_{CO,\delta}(x)$ is invariant under one-to-one transformations of X.

Remark 4.1. Tarter and Lock, 1991, Chapter 8, have proposed $\eta_T^2(x) = 1 - \sigma^2(x)/\sigma_Y^2$ as local version of $\eta^2 = 1 - \mathbb{E}\,\sigma^2(X)/\sigma_Y^2$. Note, however, that this measure is not necessarily positive in heteroscedastic models. For instance, take $Y = \alpha + \beta X + X\epsilon$, where X and ϵ are independent, $\mathbb{E}\,\epsilon = 0$, and assume $\beta = 0$. Then, since $\sigma_Y^2 = \beta^2\sigma_X^2 + \mathbb{E}\,[X^2]\sigma_\epsilon^2$, $\eta_T^2(x)$ is negative for x such that $x^2 > \mathbb{E}\,[X^2]$. In homoscedastic models $Y = \mu(X) + \epsilon$ with ϵ and X independent, $\eta_T^2(x) = 1 - \sigma_\epsilon^2/\left[\operatorname{Var}\mu(X) + \sigma_\epsilon^2\right]$, and it is no longer local.

Remark 4.2. $\eta_h^2(x)$ differs from $\rho^2(x)$ in that it involves the local relative variance $\tau_h^2(x)$ of X instead of the global variance σ_X^2. We would expect $\tau_h^2(x) \le 1$. Here are conditions for this to be the case. Consider the sequences of inequalities:

$$\begin{aligned}\left(\int_{x_0-h}^{x_0+h} |x - x_h|\, f_h(x)dx\right)^2 &\le \int_{x_0-h}^{x_0+h} (x - x_h)^2 f_h(x)dx \\ &\le \int_{x_0-h}^{x_0+h} (x - x_0)^2 f_h(x)dx \le h^2\ ,\end{aligned} \tag{4.1}$$

and

$$\begin{aligned}\int_{x_0-h}^{x_0+h} |x - x_h|\, f_h(x)dx &\ge \int_{x_0-h}^{x_0+h} \min\left(h/3, |x - x_h|\right) f_h(x)dx \\ &\ge \frac{h}{3}\int_{|x-x_h|\ge h/3} f_h(x)dx = hI_h(x_0)\ ,\end{aligned} \tag{4.2}$$

where $x_h = \mathbb{E}\,[X_h]$. Let

$$I_h(x_0) = \int_{|x-x_h|\ge h/3} f_h(x)dx\ . \tag{4.3}$$

From (4.1) and (4.2) we obtain the following result:

Proposition 4.1. *Let X, X_h be as in Proposition 2.1, let $I_h(x_0)$ be given by (4.3), and let $\mathbb{E}\,X_h = x_h$. For X with infinite support, let $h < \infty$, and for X supported in a finite interval of length $2L$, let $h < 2L$. Then, for any x_0, $I_h(x_0) > 0$, and the following conditions hold:*

a) sufficiency:

$$h \le \sigma_X \Longrightarrow \tau_h^2(x_0) \le 1\ ;$$

b) necessity:

$$\tau_h^2(x_0) \le 1 \Longrightarrow h \le 3\sigma_X/I_h(x_0)\ .$$

In the uniform $\mathcal{U}[0,1]$ case, $\tau_h^2(x_0) \le 4h^2$ for all x_0, and $\tau_h^2(x_0) \le 1$ if $h \le 1/2$. (for the formula of $\tau_h^2(x_0)$ see (6.6).) Actually, in this case, at each x_0 one can compute $\tau_h^2(x_0)$ with h in one of the three intervals $[0, \min(x_0, 1-x_0)]$, $[\min(x_0, 1-x_0), \max(x_0, 1-x_0)]$, and $[\max(x_0, 1-x_0), 1]$, and show that the sufficient condition in Proposition 4.1 is satisfied for all h and x_0; thus $\tau_h^2(x_0) \le 1$ for all h and x_0. When X is any other variable, Proposition 4.1 suggests a rule of thumb for choosing h in practice: one should always consider neighbourhood measures $\eta^2_{CA,h}(x)$ with $h \le \widehat{\sigma}_X$, where $\widehat{\sigma}_X$ is the estimate of σ_X.

We now give an example where $\tau_h^2(x) > 1$. Suppose that X has a density which equals ε on [0,1], and $(1/\varepsilon - 1)$ on $[1, 1+\varepsilon]$. Then $\operatorname{Var}(X|X \in [0,1]) = 1/12$, but $\sigma_X^2 = \operatorname{Var} X \to 0$ as $\varepsilon \to 0$. It follows that there exists $\varepsilon > 0$ such that $\tau_{.5}^2(.5) > 1$. In this example $x_{.5} = x_0 = .5$ and $I_{.5}(.5) = 2(1 - (1 - 1/6)) = 1/3$; the 'pathological' behaviour is possible because $h^2 = (.5)^2 > 81\sigma_X^2$ for some $\varepsilon > 0$, since $\sigma_X \to 0$ as $\varepsilon \to 0$.

5. Examples of Neighbourhood Correlation Ratios.

In this section we consider the behaviour of $\eta_h^2(x) = \eta_{CA,h}^2(x)$ for some models which have been studied in the literature, and we compare $\eta_h^2(x)$ with ρ^2, η^2, and $\rho^2(x)$.

Example 5.1 First consider a simple quadratic model (Hall and Wehrly, 1991),

$$Y = (X - 1/2)^2 + \tau\epsilon\,,\ \ X \sim \mathcal{U}[0,1]\,,\ \ \epsilon \sim \mathcal{N}(0,1)\,, X \text{ and } \epsilon \text{ independent}\,.$$

It is easy to see that

$$\rho^2(x) = \frac{(x-1/2)^2}{(x-1/2)^2 + 3\tau^2}\,,$$

while $\rho^2 = 0$ but η^2 is not; more precisely $\eta^2 = 1/(1 + 180\tau^2)$. In the appendix we derive $\eta_h^2(x)$ for this model, and in Figure (5.1) we plot $\eta_h^2(x)$, with $h = .3$, ρ^2, η^2, and $\rho^2(x)$. It appears that $\eta_h^2(x)$ is larger than $\rho^2(x)$ in the center and smaller in the tails, which makes sense, because $\rho^2(x)$ behaves more like $[\mu'(x)]^2$ and thus is drawn to 0 at $x = 1/2$, while tending to one for $x \to 0, 1$.

Example 5.2 Next consider the 'bump' model (Härdle, 1991),

$$Y = 2 - 5X + 5e^{-100(X-1/2)^2} + \tau\epsilon\,,\ \ X \sim \mathcal{U}[0,1]\,,\ \ \epsilon \sim \mathcal{N}(0,1)\,, X \text{ and } \epsilon \text{ independent}\,.$$

In this case the correlation curve is given by

$$\rho^2(x) = \frac{25\left[1 + 200(x-1/2)e^{-100(x-1/2)^2}\right]^2}{25\left[1 + 200(x-1/2)e^{-100(x-1/2)^2}\right]^2 + 12\tau^2}\,.$$

The formula of $\eta_h^2(x)$ is derived in the appendix. Figure (5.2) gives plots of $\eta_{.3}^2(x)$, ρ^2, η^2, and $\rho^2(x)$ when $\tau^2 = 1/4,\ 1,\ 4,\ 16$. This example is interesting because the difference between $\eta_h^2(x)$ and $\rho^2(x)$ is very striking. Since the conditional variance τ^2 is constant, like in the previous example the behaviour of $\rho^2(x)$ follows quite closely the behaviour of $[\mu'(x)]^2$, and therefore is drawn to 0 at the points $x_1 = .309$ and $x_2 = .495$ where the derivative $\mu'(x) = 0$. On the other hand, at $x_i = 1/2 \pm 1/\sqrt{200}$, $i = 3,4$, $\mu''(x) = 0$, and $\mu'(x)$ has an extremum at x_i, $i = 3,4$. This explains why $\rho^2(x_i)$, $i = 3,4$ is largest at these values.

The new measure $\eta_h^2(x)$ shows its strength in this case, because it clearly smoothes out the wild behaviour of $\rho^2(x)$ by proposing an averaging over intervals of length $2h$.

Example 5.3 Finally consider the 'twisted pear' model, of non-constant conditional variance, as well as non-linear conditional mean (Doksum et al., 1994):

$$Y = aXe^{(b-cX)} + (\gamma + \lambda X)\sigma\tau\epsilon\,,\ X \sim \mathcal{N}(\mu,\sigma^2)\,,\ \epsilon \sim \mathcal{N}(0,1)\,; a,\, b,\, c,\, \tau,\, \sigma > 0\,; \mu,\, \gamma,\, \lambda \in R\,,$$

with X and ϵ independent. This model represents a situation where the relationship betwen X and Y is strong for small x, but then tapers off. The correlation curve is

$$\rho^2(x) = \frac{a^2e^{2b-2cx}\,(1-cx)^2}{a^2e^{2b-2cx}\,(1-cx)^2 + \tau^2\,(\gamma + \lambda x)^2}\,.$$

Figure (5.4) plots $\eta_{.3}^2(x)$ ρ^2, η^2, and $\rho^2(x)$ for $a = .1$, $b = 5$, $c = .5$, $\gamma = 1$, $\lambda = .5$, $\mu = 1.2$, $\sigma = 1/3$ and $\tau^2 = 1/4,\ 1,\ 4,\ 16$. In this example the difference betwen $\rho^2(x)$ and $\eta_h^2(x)$ is less important. This can be explained as follows: unlike the bump model, this is a non-constant conditional variance model, where the square of $\mu'(x)$ and the conditional variance are polynomials of same degree, and therefore $\rho^2(x)$ behaves smoothly. On the other hand, the interpretation of $\eta_h^2(x)$ in terms of calibrated local correlation permits to use estimators which are easy to compute and converge rapidly, as can be seen in the next section.

6. Estimation. Asymptotic Results.

An estimator of $\eta^2_{CA,h}(x)$ can be defined in a natural way as the sampling equivalent of formula (2.6). That is: compute first the respective estimates of $\eta^2_{CO,h}(x)$ and of $\tau^2_h(x)$, and then insert the resulting sampling values in the first ratio of (2.6). Doksum and Samarov, 1994, have proposed three consistent estimators of the *correlation ratio*; in this paper we are using the conditional version of the estimator that performed best in their Monte Carlo study. That is, we take the sample squared correlation $\widehat{\rho}^2(\widehat{\mu}(X), Y)$, where $\widehat{\mu}(x)$ is the estimated regression curve. Another advantage of this estimator is that it takes values in the interval$[0,1]$. Let $(X_{j,h}, Y_{j,h})$ be data (X_j, Y_j) with X_j belonging to $N_h(x) = [x-h, x+h]$, let n_h be the number of X_j in N_h, while $\overline{\mu}_h$ and $\overline{Y}_h$ are the respective means of $\widehat{\mu}(X_{j,h})$ and $Y_{j,h}$. Then the estimator of $\eta^2_{CO,h}(x)$ is given by:

$$\begin{aligned}\widehat{\eta}^2_{CO,h}(x) &= \left\{ \frac{\sum_j \left[\widehat{\mu}(X_{j,h}) - \overline{\mu}_h\right]\left[Y_{j,h} - \overline{Y}_h\right]/n_h}{\sqrt{\sum_j \left[\widehat{\mu}(X_{j,h}) - \overline{\mu}_h\right]^2/n_h}\sqrt{\sum_j \left[Y_{j,h} - \overline{Y}_h\right]^2/n_h}} \right\}^2 \\ &= \left\{ \frac{\sum_{i=1}^n \left[\widehat{\mu}(X_i) - \sum_{k=1}^n w_k \widehat{\mu}(X_k)\right]\left(Y_i - \overline{Y}\right) W_i}{\sqrt{\sum_{i=1}^n \left[\widehat{\mu}(X_i) - \sum_{k=1}^n w_k \widehat{\mu}(X_k)\right]^2 W_i}\sqrt{\sum_{i=1}^n \left[Y_i - \sum_{k=1}^n w_k Y_k\right]^2 W_i}} \right\}^2,\end{aligned} \tag{6.1}$$

where:

$$W_i = \begin{cases} 1 & \text{if } X_i \in [x-h, x+h] \\ 0 & \text{otherwise,} \end{cases} \tag{6.2}$$

and $w_k = W_k / \sum_{i=1}^n W_i$.

In the same way we estimate $\tau^2_h(x)$ by:

$$\widehat{\tau}^2_h(x) = \frac{\sum_j \left(X_{j,h} - \overline{X}_h\right)^2/n_h}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2/n} = \frac{\sum_{i=1}^n \left(X_i - \sum_{k=1}^n w_k X_k\right)^2 w_i}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2/n}. \tag{6.3}$$

The proposed estimator of $\eta^2_{CA,h}(x)$ is:

$$\widehat{\eta}^2_{CA,h}(x) = \frac{\widehat{\eta}^2_{CO,h}(x)}{\widehat{\eta}^2_{CO,h}(x) + \widehat{\tau}^2_h(x)(1 - \widehat{\eta}^2_{CO,h}(x))}, \tag{6.4}$$

and it is consistent by the consistency of $\widehat{\eta}^2_{CO,h}(x)$ and of $\widehat{\tau}^2_h(x)$. In Section 7 we present the results of a simulation study of this estimator.

In situations where the design density of X is chosen to be uniform $\mathcal{U}(0,1)$, let $h \le 1/2$, and consider the simpler estimator:

$$\widetilde{\eta}^2_{CA,h}(x) = \frac{\widehat{\eta}^2_{CO,h}(x)}{\widehat{\eta}^2_{CO,h}(x) + \widetilde{\tau}^2_h(x)(1 - \widehat{\eta}^2_{CO,h}(x))}, \tag{6.5}$$

where

$$\widetilde{\tau}^2_h(x) = \tau^2_h(x) = \begin{cases} (x+h)^2 & \text{if } x \in [0,h] \\ 4h^2 & \text{if } x \in [h, 1-h] \\ (1-x+h)^2 & \text{if } x \in [1-h, 1]. \end{cases} \tag{6.6}$$

In other words, the estimate of $\tau^2_h(x)$ is replaced by its known population value (see *Appendix*). Note that $\widetilde{\tau}^2_h(x)$ is non-random.

The asymptotic results of this section can be derived using the approach of Samarov, 1993, and Doksum and Samarov, 1994. Since x is kept fixed, we shall omit indicating the dependence on x for the remainder of this section. We start with a common list of assumptions.

Assumptions.

1) the expectations $\mathbb{E}\{|X|^4\}$ and $\mathbb{E}\{|Y|^4\}$ are finite;
2) the conditional variance $\sigma^2(x)$ is bounded;
3) for fixed i_0 the estimate $\widehat{\mu}(X_{i_0})$ is a kernel estimate which does not depend on the data pair (X_{i_0}, Y_{i_0}) (that is, it is a 'leave-one-out' kernel estimate);
4) the kernel is a nonnegative, symmetric, bounded function with compact support, bounded away from 0 on some neighbourhood of the origin;
5) the bandwith b satisfies $b = O(n^{-1/3})$;
6) $\mu(x)$ is first order Lipschitz;
7) the density $f(x)$ is positive on $(x-\varepsilon, x+\varepsilon)$ for some $\varepsilon > 0$.

The first lemma can be proved by modifying the proofs of Doksum and Samarov, 1994.

Lemma 6.1. *Let $y_h = \mathbb{E}[Y|X \in [x-h, x+h]]$, and for $i = 1, \ldots, n$, let*

$$e_i = (Y_i - y_h)/\sigma_{Y,h}\ ,\ u_i = (Y_i - \mu(X_i))/\sigma_{Y,h}\sqrt{(1-\eta^2_{CO,h})}\ .$$

The estimator $\widehat{\eta}^2_{CO,h}$ introduced at (6.1) admits the following asymptotic expansion, as $n \to \infty$:

$$\sqrt{n}\,[\widehat{\eta}^2_{CO,h} - \eta^2_{CO,h}] = n^{-1/2}(1-\eta^2_{CO,h})\sum_{i=1}^{n}(e_i^2 - u_i^2)\,W_i + o_P(1)\ . \tag{6.7}$$

The next result follows from Lemma 6.1 and the delta method.

Proposition 6.2. *Assume that X is uniform $\mathcal{U}(0,1)$, and let $\widehat{\eta}^2_{CA,h}$ be the estimator introduced at (6.5). Then*

$$\sqrt{n}\,[\widehat{\eta}^2_{CA,h} - \eta^2_{CA,h}] \xrightarrow{\mathcal{L}} \mathcal{N}(0, M_{CA})\,,$$

with

$$M_{CA} = \left\{\tau_h^2 D_h^{-2}\right\}^2 \times M_{CO}\,,$$

where M_{CO} is the asymptotic variance of $\widehat{\eta}^2_{CO,h}$, and

$$D_h^2 = \left[\eta^2_{CO,h} + \tau_h^2(1-\eta^2_{CO,h})\right]^2\ ,$$

is the squared local variability of Y defined by (3.2).

The following lemma gives an asymptotic expansion of $\widehat{\tau}_h^2$ similar to the one for $\widehat{\eta}_h^2$.

Lemma 6.3. *Let $x_h = \mathbb{E}X_h$, and, for $i = 1, \ldots, n$, let W_i, be as in (6.2), and let*

$$d_i = (X_i - x_h)/\sigma_{X,h}\ ,\ f_i = (X_i - \mathbb{E}X)/\sigma_X\ .$$

The estimator $\widehat{\tau}_h^2$ admits the following asymptotic expansion, as $n \to \infty$:

$$\sqrt{n}\,[\widehat{\tau}_h^2 - \tau_h^2] = n^{-1/2}\tau_h^2\sum_{i=1}^{n}\left(d_i^2 W_i - f_i^2 - (W_i - P_h)/P_h\right) + o_P(1)\ . \tag{6.8}$$

From Lemmas 6.1 and 6.3 we obtain the main result on asymptotic normality.

Proposition 6.4. *The estimator defined at (6.4) is asymptotically normal,*

$$\sqrt{n}\,[\widehat{\eta}^2_{CA,h} - \eta^2_{CA,h}] \xrightarrow{\mathcal{L}} \mathcal{N}(0, M_{CA})\,,$$

with

$$M_{CA} = \left\{\frac{\tau_h^2}{D_h^2}\right\}^2 M_{CO} + \left\{\frac{\eta^2_{CO,h}(1-\eta^2_{CO,h})}{D_h^2}\right\}^2 M_T$$
$$-2\left\{\frac{\tau_h^2}{D_h^2}\right\}\left\{\frac{\eta^2_{CO,h}(1-\eta^2_{CO,h})}{D_h^2}\right\} M_{T,CO}\ ,$$

where D_h^2 is given by (3.2), M_{CO} is the asymptotic variance of $\widehat{\eta}^2_{CO,h}$, M_T is the asymptotic variance of $\widehat{\tau}_h^2$, and $M_{T,CO}$ is the asymptotic covariance of $\widehat{\tau}_h^2$ and $\widehat{\eta}^2_{CO,h}$.

Standard errors.

From propositions 6.2 and 6.4 we can obtain expressions for the asymptotic standard deviations of $\widehat{\eta}^2_{CO,h}$ and $\widehat{\eta}^2_{CO,h}$. If we replace the unknown parameters in these expressions with estimates we get (approximate) standard errors. For $i = 1, \ldots, n$, let

$$A_i = \widehat{d}_i^2 W_i - \widehat{f}_i^2 - (W_i - \widehat{P}_h)/\widehat{P}_h\,, \quad B_i = \left(\widehat{e}_i^2 - \widehat{u}_i^2\right) W_i\,, \text{ and } C_i = B_i - \widehat{\eta}^2_{CO,h} A_i\,,$$

where:

$$\widehat{e}_i = \left(Y_i - \overline{Y}_h\right)/s_{Y,h}\,,\ \widehat{u}_i = (Y_i - \widehat{\mu}(X_i))/s_{Y,h}\sqrt{\left(1 - \widehat{\eta}^2_{CO,h}\right)}\,,$$

$$\widehat{d}_i = \left(X_i - \overline{X}_h\right)/s_{X,h}\,,\ \widehat{f}_i = \left(X_i - \overline{X}\right)/s_X\,,\ \widehat{P}_h = n^{-1}\sum_{i=1}^{n} I\left[X_i \in [x-h, x+h]\right]\,;$$

$s^2_{X,h}$ and $s^2_{Y,h}$ are the sample variances of $X_{j,h}$ and $Y_{j,h}$ with j such that $X_j \in N_h$. Then the (approximate) standard errors of $\widehat{\eta}^2_{CA,h}$ and $\widehat{\eta}^2_{CA,h}$ are, respectively:

$$\tau_h^2 \widehat{D}_h^{-2} s_B\,, \text{ and } \tau_h^2\left(1 - \eta^2_{CO,h}\right)\widehat{D}_h^{-2} s_C\,,$$

where s_B^2 and s_C^2 are the respective sample variances of B_i and C_i, $i = 1, \ldots, n$.

7. A Simulation Study. Bandwidth Selection.

In this section we present the results of a Monte-Carlo study. Our purpose is to illustrate the finite sample size behaviour of the estimator given at (6.4), as well as to propose a simple bandwith selection procedure, based on the maximization of $\widehat{\eta}_h^2(x) = \widehat{\eta}^2_{CA,h}(x)$.

We consider the following models (presented in Section 3): quadratic, bump and twisted-pear, and for the quadratic and bump regression models we consider both fixed and random designs. For each of these 5 models we simulated 200 samples of 200 data each; in each data set let the pairs be (x_i, y_i), $i = 1, \ldots, 200$. For each sample we computed the kernel (Nadaraya-Watson) and the locally quadratic estimator as well as their leave-one-out counterparts, which differ from the usual ones in that they do not use the data (x_{i_0}, y_{i_0}), i_0 fixed, in the estimation of $\mu(x_{i_0})$. In the locally quadratic estimates, and in the kernel leave-one out estimate, at sample points x_i where the regular formula failed to work because $N_h(x)$ did not contain enough points, we replaced $\widehat{\mu}(x_i)$ by the average $[y_{i-1} + y_{i+1}]/2$ (assuming the x's have been ordered.) Such cases ocurred less than 1% of the time for sample size 200 and larger. We used the tricube kernel function, suggested by Cleveland, 1979, i.e. $K(t) = \left[1 - |t|^3\right]^3 I(|t| \leq 1)$.

A first set of results are on the comparison of six estimators of $\eta_h^2(x)$: three are based on kernel type regression-smoothers, and three are based on locally quadratic type regression-smoothers. Here is a brief description on how the estimates are obtained: for each type of smoother we compute $\widehat{\eta}^2_{CA,h}(x)$ in three ways: first we insert in (6.1) $\widehat{\mu}(x)$ as given by the usual smoother ('all in' estimates), second we insert $\widehat{\mu}(x)$ as given by the leave-one-out smoother ('one out' estimates), and a last estimate $\widehat{\eta}^2_{CA,h}(x)$ is obtained as the average of the first two $\widehat{\eta}^2_{CA,h}(x)$ estimates ('average' estimates, as suggested by Doksum and Samarov, 1994). The estimates are computed at points x with $x = x_{gj}$, $j = 1, \ldots, m$, where x_{gj}, $j = 1, \ldots, m$ are grid points ($m = 50$ in our examples). The grid points are of the form $F^{-1}(i/m)$, $i = 1, \ldots, m$, where $F(x)$ is the distribution function of X; thus the grid points are equally spaced when X is uniform. In particular, in the fixed design case with $n = 200$ data points and $m = 50$ grid-points, the latter are $x_{gj} = j/n$, $j = 4k+1$, $k = 0, \ldots, K = [(n-1)/4]$.

The first set of tables contain summary statistics concerning the estimated MISE of the six estimators of $\eta^2_{CA}(x)$, where, for each sample, the estimated MISE is given by:

$$\widehat{MISE} = \sum_{i=1}^{m}\left[\widehat{\eta}^2_{CA,h}(x_i) - \eta^2_{CA,h}(x_i)\right]/m\,,$$

where m is the number of grid points. The length of the interval is taken as $2h$, where $h = .3$, which, in all three cases, is the approximate value of the standard deviation σ_1 of X; the bandwith is $b = .22 \approx .7\sigma_1$. The results are presented in Table 1. Before any comment on the results we would like to note that in this context a sample size of only 200 is quite modest. In spite of this relatively small sample size, we found that all estimators perform very well in all models, with the exception of the twisted pear model of large τ, (i.e. $\tau = 2$ and $\tau = 4$), where the performance is slightly less good. This has to be expected, since in this model the conditional variance is non-constant, and a much larger sample size might be needed. (For example, when $\tau = 2$, at $x = 1.5$ the conditional variance is $\sigma^2(1.5) \approx 7.1$.) In most cases the 'all-in' versions perform best (i.e. have lowest mean and/or median $\widehat{MISE}$) in their class (kernel or locally quadratic). For the bump model with $\tau = .5$ and $\tau = 1$ the locally quadratic estimators perform much better than their kernel counterparts. In view of these results we decided to plot the median, 5% and 95% quantile estimated curves as given by kernel 'all in' and locally quadratic 'all in' estimates. Each curve is obtained by computing, respectively, the median, 5%, and 95% quantile of the 200 estimated values at each of the 50 grid points. The 5% and 95% curves are called envelope curves by Hall and Wehrly, 1991. In all three models, the data were generated for the random design. We took only two values of τ in the bump and twisted pear models: one small, $\tau = .5$, and one moderately large, $\tau = 2$. The results are given in Figures (7.1), (7.2), and (7.3). Figure (7.2) shows a dramatic improvement by the locally quadratic over the kernel estimate in the bump model. In the other two models the difference is not pronounced.

Next we propose and study bandwith selection by maximizing $\widehat{\eta}_h^2(x)$ at selected values of x. For $h = .3$ we studied bandwith selection at typical quantiles x_q, with q: .25, .5, and .75. The idea is to compute $\widehat{\eta}_h^2(x)$ as function of the bandwith b, where b is the bandwith used in $\widehat{\mu}(x)$, and to choose $\tilde{b}$ which maximizes $\widehat{\eta}_h^2(x)$. That is, we choose the b which locally maximizes the empirical explanatory power of the explanatory variable. The global version of this strategy has been proposed by Doksum and Samarov, 1994. In our study we simulated $ms = 200$ samples of size 200 each, and we chose the bandwith $\tilde{b}$, which maximized $\widehat{\eta}_{.3}^2(x_q)$ in each sample, over 25 equally spaced values b in [.06, .3]. Because of overfitting, only leave-one-out regression smoothers make sense in bandwith selection, and thus our selection was done by maximizing the corresponding $\widehat{\eta}_{.3}^2(x_q)$. In order to assess the quality of the bandwith selection procedure we computed summary statistics across 200 samples of the sample values $\widehat{\mu}_{\tilde{b}}(x_q)$ (both 'all-in' and 'leave-one-out' regression-smoothers), and of the sample values $\widehat{\eta}_{.3,\tilde{b}}^2(x_q)$ ('all-in', 'leave-one-out', and 'average' estimators).

Table 2 gives the median and quartiles of the selected bandwith for each regression model and both types of regression smoothers. Note that our procedure selects a larger bandwith for the locally quadratic estimate than for the kernel estimate. This makes sense, because the kernel estimate fits a constant locally, and for larger bandwidth the resulting estimate is less correlated with data which are generated from a curved regression. For the twisted pear model, where the curvature is small, b is most often chosen as the maximum possible value (i.e. $b = h = 0.3$), while for the bump model, where the curvature is high, a much smaller bandwidth is chosen. For the kernel estimate the smallest possible value 0.06 is chosen in most cases.

For $\widehat{\eta}_{.3}^2(x_q)$ ($q = .25, .5$, and .75) based on our bandwidth selection rule we give median, quartiles and mean squared error (MSE), as well as the true $\eta_{.3}^2(x)$ value, which can be compared with the median of $\widehat{\eta}_{.3}^2(x_q)$. These statistics are listed in Table 3; for economy we chose to report only the results for two values of τ, .5 and 2, in the bump and twisted pear model. All estimators perform extremely well; among locally quadratic estimators the 'average' version has the smallest MSE in most cases, while among kernel estimators the 'all in' version is the best. Surprisingly, overall the kernel estimators perform slightly better.

Further, in Table 4, we compare the median value of the four regression-smoothers $\widehat{\mu}_{\tilde{b}}(x_q)$ with the true value $\mu(x_q)$, and we give the Monte-Carlo bias and mean squared error of $\widehat{\mu}_{\tilde{b}}(x_q)$, $q = .25$, .5, and .75, for our three models under random design. (Again we retain only two values of τ, .5 and 2.) From Table 4 it appears that the estimators don't perform well in the bump model at $x_{.5} = .5$, with the kernel estimate turning in the poorest performance. This can be explained by the 'wild' behaviour of $\mu(x)$ around $x = .5$: at $x = .495$ the regression function has a maximum (the 'bump'), while at $x = .5707$ its second derivative is 0, and the function changes concavity. This suggests that both estimation and choice of an optimal bandwith at $x = .5$ might require more data. Therefore, we decided to repeat the procedure with $ms = 200$ samples of increased sample size, $n = 400$. With this new sample size, both bias and MSE were reduced in an important way (from half to a third) but we decided to report here only the results based on the same sample size for all three models.

8. A Data Example.

Finally we illustrated the behaviour of the neighbourhood *correlation ratio* estimates on a real data set. The data are from the Family Expenditure Survey, 1968-1983; scatter plots of this data set can be found in Härdle, 1991, Figure 2.1 and Figure 2.2. The data are: (X, Y), where Y is the expenditure for food, and X is the net income of $n = 7125$ households. By inspecting the scatter plots one can infer that this data set is a typical example of a 'twisted pear' model data set. In our estimation we used the same kernel as in the simulation study, and we computed the six estimators presented in Section 7, at $m = 100$ grid points $X_{(l_j)}, j = 1, \ldots, 100$ with $l_j = [(j/101)n]$. ($x_{(k)}$ denotes the ordered k-th observation). The results for the 'all-in', 'one-out', and 'average' vesions were extremely close (at least two decimals, except for the last few grid points in the right tail). Therefore, in Figure 8.1, we plotted the 'all-in' version only, kernel and locally quadratic. The curves exhibit the expected behaviour, as they decrease steadily from 85% to very low values. In Table 5 we give the local ANOVA decomposition proposed in Section 3, at six selected quantiles of this data set, i.e. at $x_{.1}$, $x_{.25}$, $x_{.5}$, $x_{.75}$, and $x_{.9}$. We also give the corresponding values of $\eta_h^2(x)$, here labeled as the *local R-squared.*

9. Appendix.

In the first half of this section we present some details of the derivation of $\eta_h^2(x) = \eta_{CA,h}^2(x)$ for the three models considered in Section 5.

1. Quadratic model.

Let $h \leq 1/2$, and take $h \leq x \leq 1 - h$; then: $\sigma_{\mu,h}^2(x) = (4/45)h^4 + (4/3)(x - 1/2)^2 h^2$, $\tau_h^2(x) = 12(h^2/3)$, while the expected conditional variance is constant, $\sigma_{Y|X,h}^2(x) = \sigma^2$. Thus we obtain:

$$\eta_h^2(x) = \frac{15(x - 1/2)^2 + h^2}{15(x - 1/2)^2 + h^2 + 45\sigma^2} = \frac{60x^2 - 60x + 4h^2 + 15}{60x^2 - 60x + 4h^2 + 15 + 180\sigma^2} .$$

For $0 < x < h$, $\tau_h^2(x) = (x + h)^2$ and the neighbourhood *correlation ratio* is

$$\eta_h^2(x) = \frac{16(x + h)^2 - 30(x + h) + 15}{16(x + h)^2 - 30(x + h) + 15 + 180\sigma^2} .$$

Similar computing gives $\tau_h^2(x) = (1 - x + h)^2$ and

$$\eta_h^2(x) = \frac{16(1 - x + h)^2 - 30(1 - x + h) + 15}{16(1 - x + h)^2 - 30(1 - x + h) + 15 + 180\sigma^2} ,$$

for $1 > x > (1 - h)$. We can easily check that the function $\eta_h^2(x)$ is continuous on $[0, 1]$; it attains its maximum at $x = 0, 1$, while its minimum value is $h^2/(h^2 + 45\sigma^2)$, and is attained at $x = 1/2$.

2. Bump model.

Like in the previous example, for $h \leq x \leq (1-h)$, $\tau_h^2(x) = 12(h^2/3)$, and the expected conditional variance is constant, $\sigma_{Y|X,h}^2(x) = \sigma^2$. We need to compute

$$V_h = \mathrm{Var}\left[-5(X - 1/2) + 5e^{-100(X-1/2)^2} | X \in [x - h, x + h]\right] . \tag{9.1}$$

Let $z = 10(x - 1/2)$ and $\delta = 10h$; then, if $\widetilde{X}$ is uniform $\mathcal{U}[x - h, x + h]$, the variable $Z = 10(\widetilde{X} - 1/2)$ is uniform $\mathcal{U}[z - \delta, z + \delta]$, and (9.1) can be obtained from the simpler formula:

$$\widetilde{V}_\delta = 25\mathrm{Var}\left(-Z/10 + e^{-Z^2}\right) = 25\left\{\mathrm{Var}\, Z/100 - (1/5)\mathbb{E}[(Z - \mathbb{E}\, Z)e^{-Z^2}] + \mathrm{Var}\, e^{-Z^2}\right\} . \tag{9.2}$$

In order to compute $\widetilde{V}_\delta$ we need to consider three cases: $-(5-\delta) \le z \le (5-\delta)$, $z-\delta < -5$, and $z+\delta > 5$.

We obtain, if $-(5-\delta) \le z \le (5-\delta)$ (or $(z+\delta) \le 5$, $(z-\delta) \ge -5$):

$$\begin{aligned}\widetilde{V}_\delta = {}& \frac{\delta^2}{12} - \frac{5}{4\delta}\left[e^{-(z-\delta)^2} - e^{-(z+\delta)^2}\right] + \frac{5z}{2\delta}\sqrt{\pi}\left\{\Phi(\sqrt{2}(z+\delta)) - \Phi(\sqrt{2}(z-\delta))\right\} \\ & + 25\left\{\frac{\sqrt{2\pi}}{4\delta}\left\{\Phi(2(z+\delta)) - \Phi(2(z-\delta))\right\} - \frac{\pi}{4\delta^2}\left\{\Phi(\sqrt{2}(z+\delta)) - \Phi(\sqrt{2}(z-\delta))\right\}^2\right\} .\end{aligned} \tag{9.3}$$

In a similar way we obtain, when z is such that $z-\delta < -5$:

$$\begin{aligned}\widetilde{V}_\delta = {}& \frac{(z+\delta+5)^2}{12\cdot 4} - \frac{5}{2(z+\delta+5)}\left[e^{-25} - e^{-(z+\delta)^2}\right] \\ & + \frac{5}{2}\frac{z+\delta-5}{z+\delta+5}\sqrt{\pi}\left\{\Phi(\sqrt{2}(z+\delta)) - \Phi(-5\sqrt{2})\right\} \\ & + 25\left\{\frac{\sqrt{2\pi}}{2(z+\delta+5)}\left\{\Phi(2(z+\delta)) - \Phi(-10)\right\}\right. \\ & \left. - \frac{\pi}{(z+\delta+5)^2}\left\{\Phi(\sqrt{2}(z+\delta)) - \Phi(-5\sqrt{2})\right\}^2\right\},\end{aligned} \tag{9.4}$$

and for z such that $z+\delta > 5$:

$$\begin{aligned}\widetilde{V}_\delta = {}& \frac{(5-z+\delta)^2}{12\cdot 4} - \frac{5}{2(5-z+\delta)}\left[e^{-(z-\delta)^2} - e^{-25}\right] \\ & + \frac{5}{2}\frac{z-\delta+5}{5-z+\delta}\sqrt{\pi}\left\{\Phi(5\sqrt{2}) - \Phi(\sqrt{2}(z-\delta))\right\} \\ & + 25\left\{\frac{\sqrt{2\pi}}{2(5-z+\delta)}\left\{\Phi(10) - \Phi(2(z-\delta))\right\}\right. \\ & \left. - \frac{\pi}{(5-z+\delta)^2}\left\{\Phi(5\sqrt{2}) - \Phi(\sqrt{2}(z-\delta))\right\}^2\right\}.\end{aligned} \tag{9.5}$$

Finally, in $\widetilde{V}_\delta = f(z,\delta)$, as given by (9.3), (9.4), and (9.5), replace z by $10(x-1/2)$ and δ by $10h$, and obtain the variance V_h. The *correlation ratio* curve is $\eta_h^2(x) = V_h/\left(V_h + \tau_h^2\tau^2\right)$, and it is plotted in Figure (5.2), together with $\rho^2(x)$, ρ^2 and η^2.

3. Twisted pear model

We compute first

$$V_h = \text{Var}\,(\mu(X)|X \in [x-h, x+h]) = a^2e^{2b}\text{Var}\,\left[Xe^{-cX}|X \in [x-h, x+h]\right] . \tag{9.6}$$

Let P_h be the probability of $X \in [x-h, x+h]$, i.e.

$$P_h = \frac{1}{\sqrt{2\pi}}\int_{x-h}^{x+h} e^{-(x-\mu)^2/2\sigma^2}dx = \Phi(\frac{x-\mu+h}{\sigma}) - \Phi(\frac{x-\mu-h}{\sigma}) ;$$

further let $F = e^{(-c\mu+c^2\sigma^2/2)}$, $Z = Xe^{-cX}$, and $w = x-\mu$. We obtain:

$$\begin{aligned}\mathbb{E}\,[Z|X \in [x-h, x+h]] = {}& \frac{F}{P_h}\left\{\frac{\sigma}{\sqrt{2\pi}}\left[e^{-(w+c\sigma^2-h)^2/2\sigma^2} - e^{-(w+c\sigma^2+h)^2/2\sigma^2}\right]\right. \\ & \left. + (\mu - c\sigma^2)\left[\Phi(\frac{w+c\sigma^2+h}{\sigma}) - \Phi(\frac{w+c\sigma^2-h}{\sigma})\right]\right\}\end{aligned}$$

and also:

$$\mathbb{E}[Z^2|X\in[x-h,x+h]]=\frac{F^2e^{c^2\sigma^2}}{P_h}\times$$

$$\left\{\frac{\sigma}{\sqrt{2\pi}}\left[(w+2c\sigma^2-h)e^{-(w+2c\sigma^2-h)^2/2\sigma^2}-(w+2c\sigma^2+h)e^{-(w+2c\sigma^2+h)^2/2\sigma^2}\right]\right.$$

$$+(\sigma^2+(\mu-2c\sigma^2)^2)\left[\Phi(\frac{w+2c\sigma^2+h}{\sigma})-\Phi(\frac{w+2c\sigma^2-h}{\sigma})\right]$$

$$\left.+\frac{2\sigma(\mu-2c\sigma^2)}{\sqrt{2\pi}}\left[e^{-(w+2c\sigma^2-h)^2/2\sigma^2}-e^{-(w+2c\sigma^2+h)^2/2\sigma^2}\right]\right\}.$$

The expected conditional variance, $\sigma^2\tau^2\mathbb{E}(\gamma+\lambda X)^2$, is:

$$D_h=\frac{\sigma^2\tau^2}{P_h}\left\{\left[\Phi(\frac{w+h}{\sigma})-\Phi(\frac{w-h}{\sigma})\right][(\gamma+\lambda\mu)^2+\lambda^2\sigma^2]\right.$$

$$+\frac{2(\gamma+\lambda\mu)\lambda\sigma}{\sqrt{2\pi}}\left[e^{-(w-h)^2/2\sigma^2}-e^{-(w+h)^2/2\sigma^2}\right]$$

$$\left.+\frac{\lambda^2\sigma}{\sqrt{2\pi}}\left[(w-h)e^{-(w-h)^2/2\sigma^2}-(w+h)e^{-(w+h)^2/2\sigma^2}\right]\right\},$$

where $w=x-\mu$.

Finally, the factor $\tau_h^2(x)$ is given by: $\left(\mathbb{E}X_h^2-(\mathbb{E}X_h)^2\right)/\sigma^2$, where:

$$\mathbb{E}X_h=\frac{1}{P_h}\left\{\mu P_h+\frac{\sigma}{\sqrt{2\pi}}\left[e^{-(w-h)^2/2\sigma^2}-e^{-(w+h)^2/2\sigma^2}\right]\right\},$$

and

$$\mathbb{E}X_h^2=\frac{1}{P_h}\left\{P_h(\sigma^2+\mu^2)+\frac{2\mu\sigma}{\sqrt{2\pi}}\left[e^{-(w-h)^2/2\sigma^2}-e^{-(w+h)^2/2\sigma^2}\right]\right.$$

$$\left.+\frac{\sigma}{\sqrt{2\pi}}\left[(w-h)e^{-(w-h)^2/2\sigma^2}-(w+h)e^{-(w+h)^2/2\sigma^2}\right]\right\}.$$

The neighbourhood *correlation ratio* is given by $\eta_h^2(x)=V_h/\left(V_h+\tau_h^2D_h\right)$. Figure (5.3) gives $\eta_h^2(x)$ for $a=.1$, $b=5$, $c=.5$, $\gamma=1$, $\lambda=.5$, $\mu=1.2$, $\sigma=1/3$ and $\tau^2=1/4,\ 1,\ 4,\ 16$.

In the second part of this section we give the proofs of Proposition 6.2, Lemma 6.3, and Proposition 6.4.

Proof of Proposition 6.2. Consider the function of t, $g(t)=t/[t+\tau_h^2(1-t)]$, with derivative: $g'(t)=\tau_h^2/[t+\tau_h^2(1-t)]^2$. Since τ_h^2 is fixed, $\eta_{CA,h}^2=g(\eta_{CO,h}^2)$, and $\widehat{\eta}_{CA,h}^2=g(\widehat{\eta}_{CO,h}^2)$. By Lemma 6.1, the estimator $\widehat{\eta}_{CO,h}^2$ is asymptotically normal. The result now follows by applying the delta method (e.g. see Bickel and Doksum, 1991) to

$$\sqrt{n}\left[g(\widehat{\eta}_{CO,h}^2)-g(\eta_{CO,h}^2)\right].\quad\square$$

Proof of Lemma 6.3. Consider the function $h(t,u,v)=t/uv$ at $\widehat{a},\widehat{b},\widehat{c}$, where:

$$\widehat{a}=\sum_{i=1}^n\left(X_i-\overline{X_h}\right)^2W_i/n=\sum_{i=1}^n(X_i-x_h)^2W_i/n-\sum_{i=1}^nW_i\left(x_h-\overline{X}_h\right)^2/n=\widehat{a}_1-\widehat{a}_2,$$

$$\widehat{b}=\sum_{i=1}^n\left(X_i-\overline{X}\right)^2/n=\sum_{i=1}^n(X_i-\mathbb{E}X)^2/n-\left(\mathbb{E}X-\overline{X}\right)^2/n=\widehat{b}_1-\widehat{b}_2,$$

$$\widehat{c}=\sum_{i=1}^nW_i/n,$$

and therefore $\widehat{a}/\widehat{b}\widehat{c} = \widehat{\tau}_h^2$. Let

$$a = \mathbb{E}\left\{\frac{\sum_{i=1}^n [X_i - x_h]^2 W_i}{n}\right\} = P_h \sigma_{X,h}^2 ,$$
$$b = \mathbb{E}\left\{\frac{\sum_{i=1}^n [X_i - \mathbb{E}X]^2}{n}\right\} = \sigma_X^2 ,$$
$$c = \mathbb{E}\left\{\frac{\sum_{i=1}^n W_i}{n}\right\} = P_h ,$$

and compute the Taylor expansion of the function $h(t,u,v)$ around (a,b,c), i.e. at the point where $h(a,b,c) = a/bc = \tau_h^2$. This gives

$$\begin{aligned} \frac{\widehat{a}}{\widehat{b}\widehat{c}} &= \frac{a}{bc} + (\widehat{a}-a)\frac{1}{\tilde{b}\tilde{c}} - \left(\widehat{b}-b\right)\frac{\tilde{a}}{\tilde{b}^2\tilde{c}} - (\widehat{c}-c)\frac{\tilde{a}}{\tilde{b}\tilde{c}^2} \Longleftrightarrow \\ \sqrt{n}\left(\widehat{\tau}_h^2 - \tau_h^2\right) &= \sqrt{n}\left(\widehat{a}_1 - a\right)\frac{1}{\tilde{b}\tilde{c}} - \sqrt{n}\left(\widehat{b}_1 - b\right)\frac{\tilde{a}}{\tilde{b}^2\tilde{c}} - \sqrt{n}\left(\widehat{c}-c\right)\frac{\tilde{a}}{\tilde{b}\tilde{c}^2} \\ &\quad + \sqrt{n}\left[-\widehat{a}_2\frac{1}{\tilde{b}\tilde{c}} + \widehat{b}_2\frac{\tilde{a}}{\tilde{b}^2\tilde{c}}\right] , \end{aligned} \tag{9.7}$$

where $\tilde{a}$, $\tilde{b}$, and $\tilde{c}$ are random and such that $\tilde{a} \xrightarrow{P} P_h\sigma_{X,h}^2$, $\tilde{b} \xrightarrow{P} \sigma_X^2$, and $\tilde{c} \xrightarrow{P} P_h$. The term

$$\sqrt{n}\left[-\widehat{a}_2\frac{1}{\tilde{b}\tilde{c}} + \widehat{b}_2\frac{\tilde{a}}{\tilde{b}^2\tilde{c}}\right] = -\frac{\sqrt{n}\sum_{i=1}^n W_i\left(x_h - \overline{X}_h\right)^2}{n}\frac{1}{\tilde{b}\tilde{c}} + \sqrt{n}\left(\mathbb{E}X - \overline{X}\right)^2\frac{\tilde{a}}{\tilde{b}^2\tilde{c}}$$

of the expansion (9.7) can be incorporated into the $o_P(1)$ term. The variables

$$\sqrt{n}\left(\widehat{a}_1 - a\right) ,\ \sqrt{n}\left(\widehat{b}_1 - b\right) ,\ \text{ and } \sqrt{n}\left(\widehat{c}-c\right) ,$$

are asymptotically normal by the central limit theorem; therefore they are bounded in probability. Hence, if we replace $\tilde{a}$, $\tilde{b}$, and $\tilde{c}$ with their limiting values, the difference beteween this new value of $\sqrt{n}\left(\widehat{\tau}_h^2 - \tau_h^2\right)$ and the one given by (9.7) is $o_P(1)$. We obtain:

$$\begin{aligned} \sqrt{n}\left(\widehat{\tau}_h^2 - \tau_h^2\right) &= \frac{1}{\sqrt{n}}\left\{\sum_{i=1}^n (X_i - x_h)^2 W_i - nP_h\sigma_{X,h}^2\right\}\frac{1}{P_h\sigma_X^2} \\ &\quad - \frac{1}{\sqrt{n}}\left\{\sum_{i=1}^n (X_i - \mathbb{E}X)^2 - n\sigma_X^2\right\}\frac{P_h\sigma_{X,h}^2}{\sigma_X^4 P_h} \\ &\quad - \frac{1}{\sqrt{n}}\left\{\sum_{i=1}^n W_i - nP_h\right\}\frac{P_h\sigma_{X,h}^2}{\sigma_X^2 P_h^2} + o_P(1) . \end{aligned} \tag{9.8}$$

The result follows after reducing the terms and letting $\tau_h^2 = \sigma_{X,h}^2/\sigma_X^2$ in (9.8). □

Proof of Proposition 6.4. Consider the function: $g(u,v) = u/[u + (1-u)v]$, with partial derivatives: $\partial g/\partial u = v/[u+(1-u)v]^2$, and $\partial g/\partial v = -u(1-u)/[u+(1-u)v]^2$. The bivariate distribution of

$$\left[\sqrt{n}\left(\widehat{\eta}_{CO,h}^2 - \eta_{CO,h}^2\right), \sqrt{n}\left(\widehat{\tau}_h^2 - \tau_h^2\right)\right]$$

is asymptotically bivariate normal, because, for any real λ, γ,

$$\begin{aligned} &\left[\lambda\sqrt{n}\left(\widehat{\eta}_{CO,h}^2 - \eta_{CO,h}^2\right) + \gamma\sqrt{n}\left(\widehat{\tau}_h^2 - \tau_h^2\right)\right] = \\ &n^{-1/2}\sum_{i=1}^n \left\{\lambda\left(1 - \eta_{CO,h}^2\right)\left[e_i^2 - u_i^2\right]W_i + \gamma\tau_h^2\left[d_i^2 W_i - f_i^2 - (W_i - P_h)/P_h\right]\right\} + o_P(1) , \end{aligned}$$

by Lemmas 6.1 and 6.3. The result follows from standard asymptotic theory (e.g. Serfling, 1980), applied to the transformation

$$\sqrt{n}\,[\hat{\eta}^2_{CA,h} - \eta^2_{CA,h}] = \sqrt{n}\,[g(\hat{\eta}^2_{CO,h}, \hat{\tau}^2_h) - g(\eta^2_{CO,h}, \tau^2_h)] \ . \quad \square$$

Acknowledgements. A large part of the computer work was done in the Computing Facility of the Department of Statistics, University of California at Berkeley, while the second author was visiting the department. H. Ibser, University of California at Berkeley, helped with the plots. P. Spector, University of California at Berkeley, and J.-F. Angers and A. Leblanc, Université de Montréal, gave important help and advice in the final stages of computing. The authors gratefully acknowledge this support.

BIBLIOGRAPHY

BICKEL, P. J. AND DOKSUM, K. (1991), *Mathematical statistics: basic ideas and methods,* Englewood Cliffs, N.J.: Prentice-Hall.

BJERVE, S. AND DOKSUM, K. (1993), "Correlation curves: measures of association as functions of covariate values," *Annals of Statistics,* **21**, 890-902.

CLEVELAND, W. S. (1979), "Robust locally weighted regression and smoothing scatterplots," *Journal of the American Statistical Association,* **74**, 829-836.

CLEVELAND, W. S. AND DEVLIN, S. J. (1988), "Locally weighted regression: An approach to regression analysis by local fitting," *Journal of the American Statistical Association,* **83**, 596-610.

DOKSUM, K., BLYTH S., BRADLOW, E., MENG, X.L., AND ZHAO, H. (1994). "Correlation curves as local measures of variance explained by regression," *Journal of the American Statistical Association,* **89** .

DOKSUM, K. AND SAMAROV, A. (1993), "Global functionals and a measure of the explanatory power," IFRSC No. 255-93, *Sloan School of Management, MIT*, to appear in the *Annals of Statistics.*

FRIEDMAN, J.(1984). "A variable span smoother," Department of Statistics Technical Report LCS5, Stanford University, Stanford, CA.

GULLIKSEN, H. (1951), *Theory of Mental Tests,* New York: John Wiley.

HALL, P. AND WEHRLY, J.(1991), "A geometrical method for removing edge effects from kernel-type nonparametric regression," *Journal of the American Statistical Association,* **86**, 665-672.

HÄRDLE, W.(1991), *Applied nonparametric regression,* Cambridge: Cambridge University Press.

HÄRDLE, W. AND BOWMAN, A.(1988), "Bootstrapping in nonparametric regression: local adaptive smoothing and confidence bands," *Journal of the American Statistical Association,* **83**, 102-110.

RÉNYI, A.(1959), "On measures of dependence," *Acta Mathematica Academiae Hungarica,* **10**, 441-451.

SAMAROV, A.(1993), "Exploring regression structure using nonparametric functional estimation," *Journal of the American Statistical Association,* **88**, 836-847.

SERFLING, R. J. (1980), *Approximation theorems of mathematical statistics,* New York: John Wiley.

TARTER, R. J. AND LOCK, M. D. (1991), *Model-free curve estimation,* Chapman and Hall .

Tables

In all tables Ke stands for 'kernel', and LQ stands for 'locally quadratic'. All results are for 200 Monte-Carlo samples of size 200 each, from the indicated model.

TABLE 1: MISE summary statistics for the six estimates of the neighbourhood correlation ratio with $h = .3$.

TABLE 1A: Quadratic model with $\tau = .1$.

Fixed design

ESTIMATOR	MEAN	MEDIAN	SD	IQR
Ke: all in	0.006820	0.005468	0.005061	0.005087
Ke: one out	0.010802	0.008172	0.009658	0.010597
Ke: average	0.008193	0.006035	0.006857	0.007105
LQ: all in	0.006706	0.005317	0.005197	0.005116
LQ: one out	0.011954	0.008589	0.010828	0.010188
LQ: average	0.007678	0.005620	0.006503	0.006062

Random design

ESTIMATOR	MEAN	MEDIAN	SD	IQR
Ke: all in	0.007163	0.005504	0.005376	0.006877
Ke: one out	0.011108	0.007715	0.010886	0.007597
Ke: average	0.008499	0.006276	0.007661	0.006214
LQ: all in	0.007060	0.005846	0.004657	0.005750
LQ: one out	0.012167	0.007940	0.012442	0.009188
LQ: average	0.007939	0.005766	0.007139	0.006454

TABLE 1B: Bump model with 4 values of τ.

$\tau = .5$, Fixed design

ESTIMATOR	MEAN	MEDIAN	SD	IQR
Ke: all in	0.013658	0.011456	0.008729	0.012211
Ke: one out	0.019975	0.017133	0.012016	0.017717
Ke: average	0.016642	0.014189	0.010284	0.014901
LQ: all in	0.000375	0.000228	0.000398	0.000336
LQ: one out	0.000907	0.000510	0.001140	0.000915
LQ: average	0.000587	0.000313	0.000717	0.000561

$\tau = .5$, Random design

ESTIMATOR	MEAN	MEDIAN	SD	IQR
Ke: all in	0.007875	0.006896	0.004768	0.006013
Ke: one out	0.013901	0.012023	0.008742	0.013351
Ke: average	0.009256	0.007869	0.005979	0.008782
LQ: all in	0.007269	0.007074	0.003900	0.005097
LQ: one out	0.017588	0.016199	0.010054	0.013568
LQ: average	0.009651	0.008743	0.005684	0.007175

$\tau = 1$, Fixed design

ESTIMATOR	MEAN	MEDIAN	SD	IQR
Ke: all in	0.015565	0.011495	0.011768	0.017106
Ke: one out	0.027955	0.027405	0.016031	0.026048
Ke: average	0.020839	0.018702	0.013387	0.021911
LQ: all in	0.003716	0.002457	0.004380	0.002801
LQ: one out	0.011462	0.006355	0.012590	0.013520
LQ: average	0.006283	0.003141	0.007544	0.006247

$\tau = 1$, Random design

ESTIMATOR	MEAN	MEDIAN	SD	IQR
Ke: all in	0.015763	0.011897	0.011743	0.015916
Ke: one out	0.026859	0.025767	0.015484	0.025794
Ke: average	0.020415	0.018048	0.012998	0.020522
LQ: all in	0.004203	0.002963	0.003887	0.003590
LQ: one out	0.012724	0.008165	0.011714	0.015798
LQ: average	0.007016	0.004392	0.006586	0.008116

$\tau = 2$, Fixed design

ESTIMATOR	MEAN	MEDIAN	SD	IQR
Ke: all in	0.013744	0.012537	0.007166	0.009806
Ke: one out	0.024261	0.022548	0.012865	0.017658
Ke: average	0.015887	0.014367	0.008467	0.011121
LQ: all in	0.011728	0.010058	0.007543	0.009680
LQ: one out	0.018528	0.017188	0.009591	0.012744
LQ: average	0.010256	0.008661	0.006247	0.007743

$\tau = 2$, Random design

ESTIMATOR	MEAN	MEDIAN	SD	IQR
Ke: all in	0.014765	0.012636	0.008797	0.009437
Ke: one out	0.026038	0.023061	0.015036	0.018043
Ke: average	0.017162	0.014840	0.014047	0.011644
LQ: all in	0.011530	0.009427	0.006762	0.008259
LQ: one out	0.018903	0.016971	0.010474	0.012619
LQ: average	0.010237	0.008618	0.006415	0.007494

$\tau = 4$, Fixed design

ESTIMATOR	MEAN	MEDIAN	SD	IQR
Ke: all in	0.014793	0.011429	0.0116792	0.011263
Ke: one out	0.032218	0.027173	0.021286	0.028196
Ke: average	0.016158	0.013998	0.010333	0.011391
LQ: all in	0.023829	0.018587	0.016436	0.019553
LQ: one out	0.024995	0.019356	0.018167	0.023686
LQ: average	0.015431	0.012828	0.010547	0.013190

$\tau = 4$, Random design

ESTIMATOR	MEAN	MEDIAN	SD	IQR
Ke: all in	0.015844	0.013029	0.010491	0.013578
Ke: one out	0.036157	0.029098	0.025788	0.032586
Ke: average	0.018093	0.015806	0.011661	0.013785
LQ: all in	0.021341	0.017443	0.014352	0.018423
LQ: one out	0.029426	0.023928	0.022321	0.026688
LQ: average	0.016163	0.014216	0.010167	0.012603

TABLE 1C: Twisted pear model with 4 values of τ : Random design

$\tau = .5$

ESTIMATOR	MEAN	MEDIAN	SD	IQR
Ke: all in	0.007875	0.006896	0.004768	0.006013
Ke: one out	0.013901	0.012023	0.008742	0.013351
Ke: average	0.009257	0.007869	0.005979	0.008782
LQ: all in	0.007269	0.007074	0.003900	0.005097
LQ: one out	0.017588	0.016199	0.010053	0.013568
LQ: average	0.009651	0.008743	0.005684	0.007175

$\tau = 1$

ESTIMATOR	MEAN	MEDIAN	SD	IQR
Ke: all in	0.062117	0.061266	0.0193193	0.026910
Ke: one out	0.089327	0.088774	0.025569	0.036980
Ke: average	0.073185	0.073802	0.021693	0.031833
LQ: all in	0.052484	0.051717	0.0175540	0.024908
LQ: one out	0.097796	0.096689	0.030221	0.040760
LQ: average	0.070997	0.070212	0.022632	0.031515

$\tau = 2$

ESTIMATOR	MEAN	MEDIAN	SD	IQR
Ke: all in	0.205789	0.201882	0.042932	0.057747
Ke: one out	0.272224	0.267878	0.051848	0.079549
Ke: average	0.234277	0.229982	0.045733	0.067330
LQ: all in	0.177301	0.176174	0.039877	0.055587
LQ: one out	0.292462	0.290384	0.055562	0.081236
LQ: average	0.227598	0.228066	0.045644	0.065135

$\tau = 4$

ESTIMATOR	MEAN	MEDIAN	SD	IQR
Ke: all in	0.452642	0.456105	0.075092	0.099985
Ke: one out	0.561215	0.559499	0.080929	0.105508
Ke: average	0.498667	0.501966	0.071951	0.095692
LQ: all in	0.386171	0.388151	0.074234	0.101836
LQ: one out	0.585855	0.586533	0.085743	0.119263
LQ: average	0.474100	0.480792	0.072205	0.104879

TABLE 2: Optimal bandwith for estimation at selected quantiles; Q-1, Q-3 stand for first, third quartile respectively.

	Kernel estimate			Locally quadratic estimate		
Quantile	Q-1	Q-3	Median	Q-1	Q-3	Median
Quadratic model, $\tau = .1$						
$x_{.25} = .25$	0.11	0.24	0.18	0.23	0.30	0.30
$x_{.50} = .50$	0.13	0.30	0.27	0.14	0.30	0.27
$x_{.75} = .75$	0.12	0.21	0.16	0.21	0.30	0.30
Bump model, $\tau = .5$						
$x_{.25} = .25$	0.06	0.08	0.06	0.11	0.14	0.13
$x_{.50} = .50$	0.06	0.06	0.06	0.10	0.13	0.12
$x_{.75} = .75$	0.06	0.07	0.06	0.11	0.14	0.13
Bump model, $\tau = 1$						
$x_{.25} = .25$	0.06	0.10	0.08	0.13	0.18	0.15
$x_{.50} = .50$	0.06	0.09	0.07	0.13	0.17	0.15
$x_{.75} = .75$	0.07	0.10	0.08	0.13	0.175	0.15
Bump model, $\tau = 2$						
$x_{.25} = .25$	0.08	0.13	0.105	0.18	0.22	0.18
$x_{.50} = .50$	0.07	0.11	0.095	0.14	0.21	0.18
$x_{.75} = .75$	0.08	0.13	0.11	0.15	0.23	0.20
Bump model, $\tau = 4$						
$x_{.25} = .25$	0.08	0.27	0.125	0.10	0.23	0.16
$x_{.50} = .50$	0.09	0.16	0.13	0.14	0.27	0.21
$x_{.75} = .75$	0.09	0.185	0.14	0.15	0.29	0.23
Twisted pear model, $\tau = .5$						
$x_{.25} = .975$	0.15	0.30	0.255	0.25	0.30	0.30
$x_{.50} = 1.2$	0.15	0.30	0.25	0.245	0.30	0.30
$x_{.75} = 1.425$	0.20	0.30	0.30	0.26	0.30	0.30
Twisted pear model, $\tau = 1$						
$x_{.25} = .975$	0.17	0.30	0.26	0.25	0.30	0.30
$x_{.50} = 1.2$	0.17	0.30	0.29	0.27	0.30	0.30
$x_{.75} = 1.425$	0.215	0.30	0.30	0.23	0.30	0.30
Twisted pear model, $\tau = 2$						
$x_{.25} = .975$	0.155	0.30	0.30	0.21	0.30	0.30
$x_{.50} = 1.2$	0.20	0.30	0.30	0.24	0.30	0.30
$x_{.75} = 1.425$	0.145	0.30	0.30	0.13	0.30	0.30
Twisted pear model, $\tau = 4$						
$x_{.25} = .975$	0.15	0.30	0.30	0.16	0.30	0.30
$x_{.50} = 1.2$	0.125	0.28	0.30	0.13	0.30	0.30
$x_{.75} = 1.425$	0.15	0.30	0.30	0.16	0.30	0.30

TABLE 3: Summary statistics for the six estimators of $\eta^2(x)$ at selected quantiles, when locally optimal bandwith is used.

TABLE 3A: Quadratic model with $\tau = .1$: Random design.

$x = x_{.25} = .25$, $\eta^2(x) = .649805$

Estimator	QUARTILE-1	QUARTILE-3	MEDIAN	MSE
Ke: all in	0.620330	0.718005	0.670615	0.006062
Ke: one out	0.619852	0.717863	0.670573	0.006081
Ke: average	0.595503	0.697033	0.646115	0.006744
LQ: all in	0.631910	0.733443	0.688271	0.007038
LQ: one out	0.548988	0.656359	0.603542	0.011895
LQ: average	0.592358	0.693311	0.644465	0.007297

$x = x_{.5} = .5$, $\eta^2(x) = .1667$

Estimator	QUARTILE-1	QUARTILE-3	MEDIAN	MSE
Ke: all in	0.153575	0.298198	0.225532	0.014882
Ke: one out	0.152426	0.298077	0.224311	0.014906
Ke: average	0.120015	0.248605	0.177665	0.008488
LQ: all in	0.172991	0.336235	0.252178	0.022494
LQ: one out	0.053080	0.106245	0.170541	0.008862
LQ: average	0.123571	0.249546	0.182825	0.009102

$x = x_{.75} = .75$, $\eta^2(x) = .649805$

Estimator	QUARTILE-1	QUARTILE-3	MEDIAN	MSE
Ke: all in	0.616723	0.716867	0.676719	0.005636
Ke: one out	0.615946	0.716211	0.676538	0.005927
Ke: average	0.587656	0.699163	0.654730	0.005953
LQ: all in	0.633267	0.731651	0.691290	0.006473
LQ: one out	0.539504	0.664291	0.613885	0.009621
LQ: average	0.586948	0.695748	0.654298	0.006129

TABLE 3B: Bump model with 2 values of τ : Random design.

$\tau = .5$, $x = x_{.25} = .25$, $\eta^2(x) = .952615$

Estimator	QUARTILE-1	QUARTILE-3	MEDIAN	MSE
Ke: all in	0.948591	0. 959554	0.954086	0. 000063
Ke: one out	0.948661	0.959507	0.954113	0. 000063
Ke: average	0.944077	0.955714	0.950561	0. 000082
LQ: all in	0.953489	0.964414	0. 958695	0.000089
LQ: one out	0.937501	0.951074	0. 944906	0.000169
LQ: average	0.945289	0.957425	0.951739	0.000074

$\tau = .5$, $x = x_{.5} = .5$, $\eta^2(x) = .976808$

Estimator	QUARTILE-1	QUARTILE-3	MEDIAN	MSE
Ke: all in	0.974633	0.980510	0.977695	0.000020
Ke: one out	0.974557	0.980508	0.977666	0.000020
Ke: average	0.972383	0.978831	0.976045	0.000025
LQ: all in	0.977591	0.983008	0.980278	0.000028
LQ: one out	0.970473	0.977536	0.974166	0.000040
LQ: average	0.974368	0.980260	0.977158	0.000023

$\tau = .5$, $x = x_{.75} = .75$, $\eta^2(x) = .988185$

Estimator	QUARTILE-1	QUARTILE-3	MEDIAN	MSE
Ke: all in	0.987194	0.989505	0.988331	0.000004
Ke: one out	0.987196	0.989492	0.988329	0.000004
Ke: average	0.986065	0.988711	0.987365	0.000005
LQ: all in	0.988617	0.990844	0.989742	0.000005
LQ: one out	0.984531	0.987757	0.986332	0.000011
LQ: average	0.986659	0.989231	0.987982	0.000005

$\tau = 2$, $x = x_{.25} = .25$, $\eta^2(x) = .556832$

Estimator	QUARTILE-1	QUARTILE-3	MEDIAN	MSE
Ke: all in	0.524709	0.642698	0.595189	0.010134
Ke: one out	0.525610	0.642100	0.597163	0.010166
Ke: average	0.469006	0.596860	0.536902	0.010622
LQ: all in	0.555089	0.677809	0.627567	0.011430
LQ: one out	0.375833	0.526572	0.442925	0.026233
LQ: average	0.470169	0.599776	0.536685	0.010950

$\tau = 2$, $x = x_{.5} = .5$, $\eta^2(x) = .724696$

Estimator	QUARTILE-1	QUARTILE-3	MEDIAN	MSE
Ke: all in	0.715064	0.771763	0.740538	0.002048
Ke: one out	0.714887	0.771286	0.740195	0.002054
Ke: average	0.688901	0.752252	0.719927	0.002072
LQ: all in	0.732619	0.790789	0.756172	0.002888
LQ: one out	0.652283	0.721688	0.686567	0.004150
LQ: average	0.690500	0.755195	0.723826	0.002022

$\tau = 2$, $x = x_{.75} = .75$, $\eta^2(x) = .839413$

Estimator	QUARTILE-1	QUARTILE-3	MEDIAN	MSE
Ke: all in	0.833328	0.867229	0.851011	0.000824
Ke: one out	0.833087	0.866744	0.850881	0.000824
Ke: average	0.821083	0.856125	0.839553	0.000824
LQ: all in	0.843511	0.876334	0.860178	0.001121
LQ: one out	0.800431	0.838808	0.822420	0.001589
LQ: average	0.821693	0.858330	0.841531	0.000885

TABLE 3C: Twisted pear model with 2 values of τ : Random design

$\tau = .5,\ x = x_{.25} = .975,\ \eta^2(x) = .973326$

Estimator	QUARTILE-1	QUARTILE-3	MEDIAN	MSE
Ke: all in	0.970571	0.976462	0.973632	0.000023
Ke: one out	0.970488	0.976460	0.973626	0.000023
Ke: average	0.969576	0.975849	0.972819	0.000024
LQ: all in	0.971020	0.977257	0.974170	0.000023
LQ: one out	0.967865	0.974023	0.971082	0.000030
LQ: average	0.969471	0.975499	0.972706	0.000024

$\tau = .5,\ x = x_{.5} = 1.2,\ \eta^2(x) = .945311$

Estimator	QUARTILE-1	QUARTILE-3	MEDIAN	MSE
Ke: all in	0.940070	0.953211	0.947849	0.000099
Ke: one out	0.940030	0.953101	0.947841	0.000096
Ke: average	0.938070	0.951774	0.946325	0.000103
LQ: all in	0.941430	0.954624	0.948801	0.000102
LQ: one out	0.935639	0.949536	0.942671	0.000124
LQ: average	0.938372	0.951899	0. 946302	0.000103

$\tau = .5,\ x = x_{.75} = 1.425,\ \eta^2(x) = .879342$

Estimator	QUARTILE-1	QUARTILE-3	MEDIAN	MSE
Ke: all in	0.865533	0.899470	0.883435	0.000731
Ke: one out	0.865775	0.899532	0.883706	0.000731
Ke: average	0.861500	0.894867	0.879486	0.000759
LQ: all in	0.867910	0.902555	0.885603	0.000759
LQ: one out	0.852365	0.886716	0.870807	0.000959
LQ: average	0.860937	0.895046	0.878844	0.000781

$\tau = 2,\ x = x_{.25} = .975,\ \eta^2(x) = .695182$

Estimator	QUARTILE-1	QUARTILE-3	MEDIAN	MSE
Ke: all in	0.673229	0.748298	0.708327	0.004248
Ke: one out	0.673074	0.748246	0.707982	0.004252
Ke: average	0.652206	0.728606	0.693450	0.004593
LQ: all in	0.680261	0.762735	0.719559	0.004539
LQ: one out	0.617446	0.705322	0.664079	0.006790
LQ: average	0.650424	0.730140	0.689360	0.004677

$\tau = 2,\ x = x_{.5} = 1.2,\ \eta^2(x) = .519307$

Estimator	QUARTILE-1	QUARTILE-3	MEDIAN	MSE
Ke: all in	0.471991	0.606067	0.526045	0.010383
Ke: one out	0.472135	0.605244	0.527965	0.010410
Ke: average	0.441966	0.577655	0.495691	0.010584
LQ: all in	0.487593	0.625738	0.539032	0.012027
LQ: one out	0.386342	0.527479	0.449938	0.016534
LQ: average	0.436404	0.570841	0.493195	0.011246

$\tau = 2,\ x = x_{.75} = 1.425,\ \eta^2(x) = .312947$

Estimator	QUARTILE-1	QUARTILE-3	MEDIAN	MSE
Ke: all in	0.287338	0.470723	0.381243	0.019513
Ke: one out	0.288768	0.470245	0.381017	0.019513
Ke: average	0.234926	0.418430	0.329591	0.014379
LQ: all in	0.323809	0.496744	0.409405	0.025315
LQ: one out	0.127937	0.312330	0.232509	0.022270
LQ: average	0.233418	0.409369	0.325515	0.014149

TABLE 4: Summary statistics of the four regression estimates of $\mu(x)$ at selected quantiles, when locally optimal bandwith is used.

TABLE 4A: Quadratic model with $\tau = .1$: Random design.

$x = x_{.25} = .25,\ \mu(x) = .0625$

	Ke: all in	Ke: one out	LQ: all in	LQ: one out
BIAS	0.002698	0.003149	-0.001745	-0.001037
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MSE	0.000634	0.006404	0.000868	0.009191
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MEDIAN	0.062950	0.063860	0.059568	0.059760

$x = x_{.50} = .50,\ \mu(x) = 0$

	Ke: all in	Ke: one out	LQ: all in	LQ: one out
BIAS	0.011352	0.011661	0.003246	0.003715
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MSE	0.000370	0.000380	0.000598	0.000675
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MEDIAN	0.010827	0.011550	0.002141	0.001087

$x = x_{.75} = .75,\ \mu(x) = .0625$

	Ke: all in	Ke: one out	LQ: all in	LQ: one out
BIAS	0.003581	0.004051	-0.002930	-0.002372
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MSE	0.000644	0.000661	0.000916	0.000948
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MEDIAN	0.067264	0.067901	0.061420	0.615163

TABLE 4B: Bump model with 2 values of τ : Random design.

$\tau = .5$, $x = x_{.25} = .25$, $\mu(x) = .759652$

	Ke: all in	Ke: one out	LQ: all in	LQ: one out
BIAS	0.037695	0.035737	0.026432	0.021225
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MSE	0.028846	0.029800	0.039242	0.043323
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MEDIAN	0.783411	0.784975	0.789445	0.780731

$\tau = .5$, $x = x_{.50} = .50$, $\mu(x) = 4.5$

	Ke: all in	Ke: one out	LQ: all in	LQ: one out
BIAS	-0.783227	-0.795204	-0.587847	-0.581178
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MSE	1.083560	1.101290	0.884804	0.883123
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MEDIAN	3.96858	3.93787	4.15900	4.15410

$\tau = .5$, $x = x_{.75} = .75$, $\mu(x) = -1.74035$

	Ke: all in	Ke: one out	LQ: all in	LQ: one out
BIAS	0.047746	0.052182	0.038087	0.045441
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MSE	0.070804	0.072878	0.073207	0.073207
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MEDIAN	-1.73321	-1.73425	-1.73669	-1.73166

$\tau = 2$, $x = x_{.25} = .25$, $\mu(x) = .759652$

	Ke: all in	Ke: one out	LQ: all in	LQ: one out
BIAS	0.107724	0.110239	0.078950	0.085267
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MSE	0.190957	0.184770	0.395841	0.389043
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MEDIAN	0.866162	0.865024	0.832112	0.807872

$\tau = 2$, $x = x_{.50} = .50$, $\mu(x) = 1.3$

	Ke: all in	Ke: one out	LQ: all in	LQ: one out
BIAS	-0.935395	-0.961065	-0.508853	-0.517577
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MSE	1.285460	1.312490	0.867072	0.858126
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MEDIAN	3.59658	3.58374	4.09616	4.10285

$\tau = 2$, $x = x_{.75} = .75$, $\mu(x) = -1.74035$

	Ke: all in	Ke: one out	LQ: all in	LQ: one out
BIAS	0.031500	0.038765	-0.059067	-0.053529
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MSE	0.186880	0.199263	0.296997	0.337206
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MEDIAN	-1.74458	-1.76303	-1.81677	-1.84207

TABLE 4C: Twisted pear model with 2 values of τ : Random design.

$\tau = .5$, $x = x_{.25} = .975$, $\mu(x) = 8.88786$

	Ke: all in	Ke: one out	LQ: all in	LQ: one out
BIAS	0.050393	0.050803	0.005239	0.004391
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MSE	0.026681	0.026972	0.021943	0.022458
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MEDIAN	8.94356	8.94599	8.90458	8.90946

$\tau = .5$, $x = x_{.5} = 1.2$, $\mu(x) = 9.7741$

	Ke: all in	Ke: one out	LQ: all in	LQ: one out
BIAS	-0.019417	-0.019365	0.000504	0.000732
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MSE	0.011420	0.011636	0.011103	0.011720
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MEDIAN	9.75436	9.75629	9.77467	9.77272

$\tau = .5$, $x = x_{.75} = 1.425$, $\mu(x) = 10.3714$

	Ke: all in	Ke: one out	LQ: all in	LQ: one out
BIAS	-0.066721	-0.671324	-0.007192	-0.006001
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MSE	0.013544	0.013811	0.008750	0.009075
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MEDIAN	10.3097	10.3094	10.3666	10.3657

$\tau = 2$, $x = x_{.25} = .975$, $\mu(x) = 8.88786$

	Ke: all in	Ke: one out	LQ: all in	LQ: one out
BIAS	0.053911	0.055003	0.008825	0.011237
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MSE	0.056107	0.053890	0.088241	0.083588
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MEDIAN	8.93182	8.95260	8.87963	8.73920

$\tau = 2$, $x = x_{.50} = 1.2$, $\mu(x) = 9.7741$

	Ke: all in	Ke: one out	LQ: all in	LQ: one out
BIAS	-0.034182	-0.035970	-0.012300	-0.015072
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MSE	0.032061	0.032730	0.049492	0.054732
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MEDIAN	9.72745	9.73415	9.75888	9.76606

$\tau = 2$, $x = x_{.75} = 1.425$, $\mu(x) = 10.3714$

	Ke: all in	Ke: one out	LQ: all in	LQ: one out
BIAS	-0.062849	-0.061583	-0.004470	-0.000552
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MSE	0.047155	0.046177	0.075682	0.076252
	Ke: all in	Ke: one out	LQ: all in	LQ: one out
MEDIAN	10.2959	10.2882	10.3788	10.3626

TABLE 5:
Local ANOVA decomposition and local R-squared at 6 selected quantiles, for 4 estimates of the neighbourhood correlation ratio with $h = \sigma_X$. The data set is Y = food expenditure versus X = income for $n = 7125$ households, Family Expenditure Survey, 1968-1983. The variances are given in 10^6 units.

$x = x_{.1}$

ESTIMATOR	Explained	Residual	Total	Local R-sq.
Ke: all in	5.06	1.09	6.15	0.82
Ke: one out	5.05	1.09	6.15	0.82
LQ: all in	5.18	1.08	6.26	0.83
LQ: one out	5.16	1.08	6.24	0.83

$x = x_{.25}$

ESTIMATOR	Explained	Residual	Total	Local R-sq.
Ke: all in	6.56	2.12	8.68	0.76
Ke: one out	6.55	2.12	8.67	0.76
LQ: all in	6.66	2.09	8.76	0.76
LQ: one out	6.65	2.10	8.75	0.76

$x = x_{.5}$

ESTIMATOR	Explained	Residual	Total	Local R-sq.
Ke: all in	6.38	3.73	10.11	0.63
Ke: one out	6.37	3.73	10.10	0.63
LQ: all in	6.42	3.72	10.14	0.63
LQ: one out	6.40	3.72	10.12	0.63

$x = x_{.75}$

ESTIMATOR	Explained	Residual	Total	Local R-sq.
Ke: all in	3.17	4.21	7.38	0.43
Ke: one out	3.15	4.22	7.36	0.43
LQ: all in	3.18	4.21	7.39	0.43
LQ: one out	3.13	4.22	7.35	0.43

$x = x_{.9}$

ESTIMATOR	Explained	Residual	Total	Local R-sq.
Ke: all in	2.42	5.66	8.07	0.30
Ke: one out	2.35	5.67	8.02	0.29
LQ: all in	2.43	5.65	8.08	0.30
LQ: one out	2.32	5.68	8.00	0.29

Figure 5.1. Local and Global Squared Correlations for the Quadratic Model

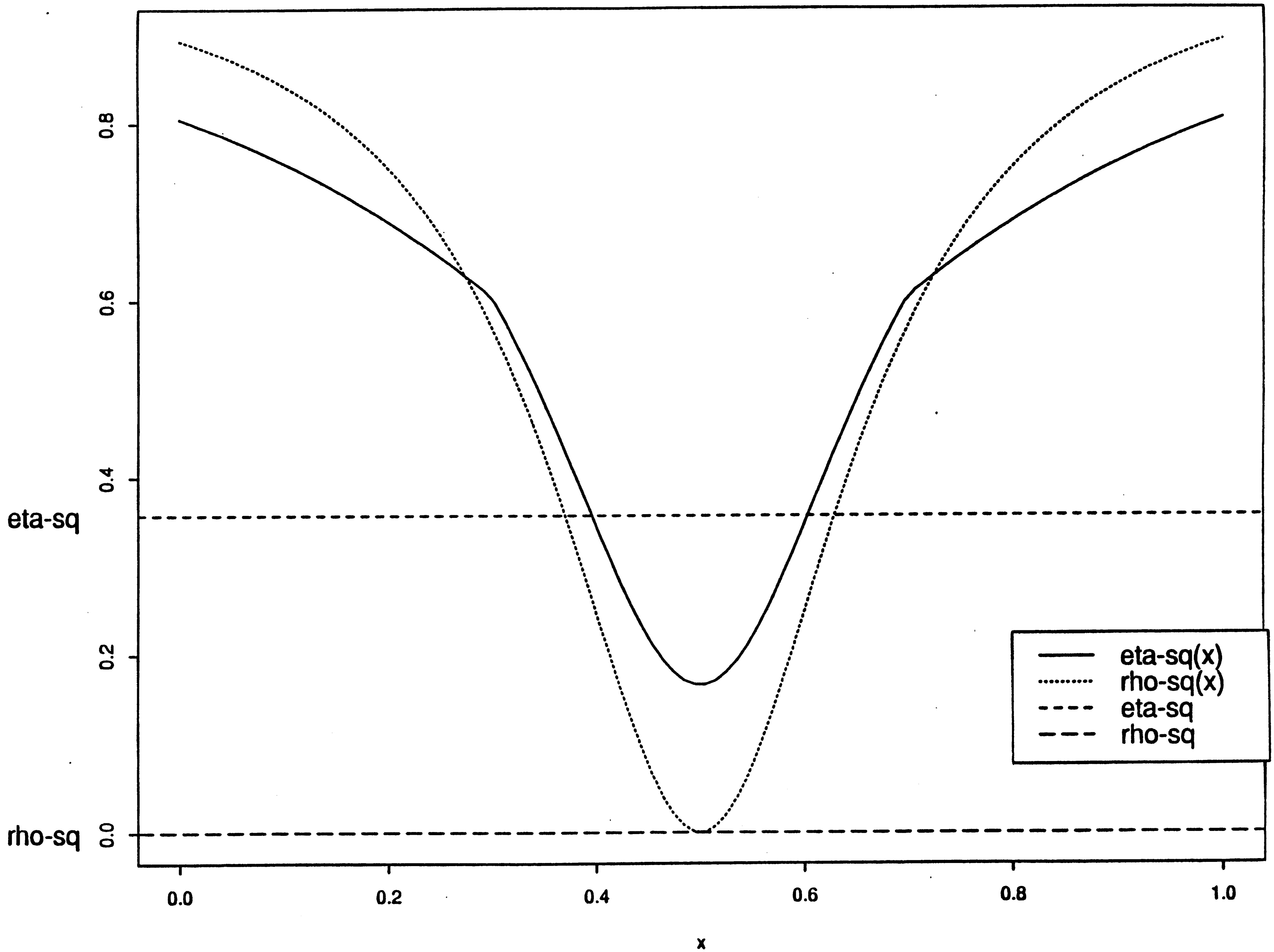


Figure 5.2. Local and Global Squared Correlations for the Bump Model

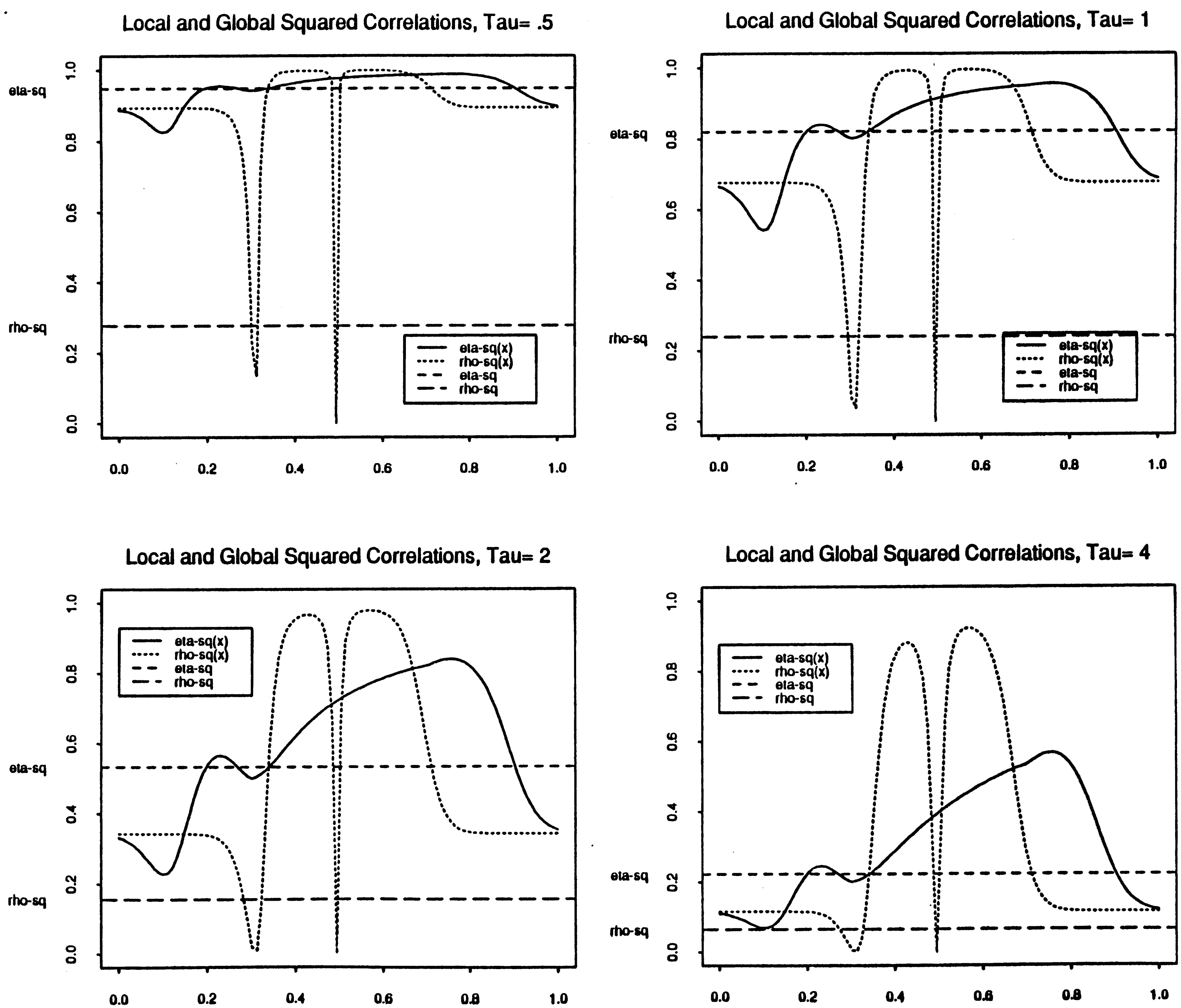


Figure 5.3. Local and Global Squared Correlations for the Twisted Pear Model

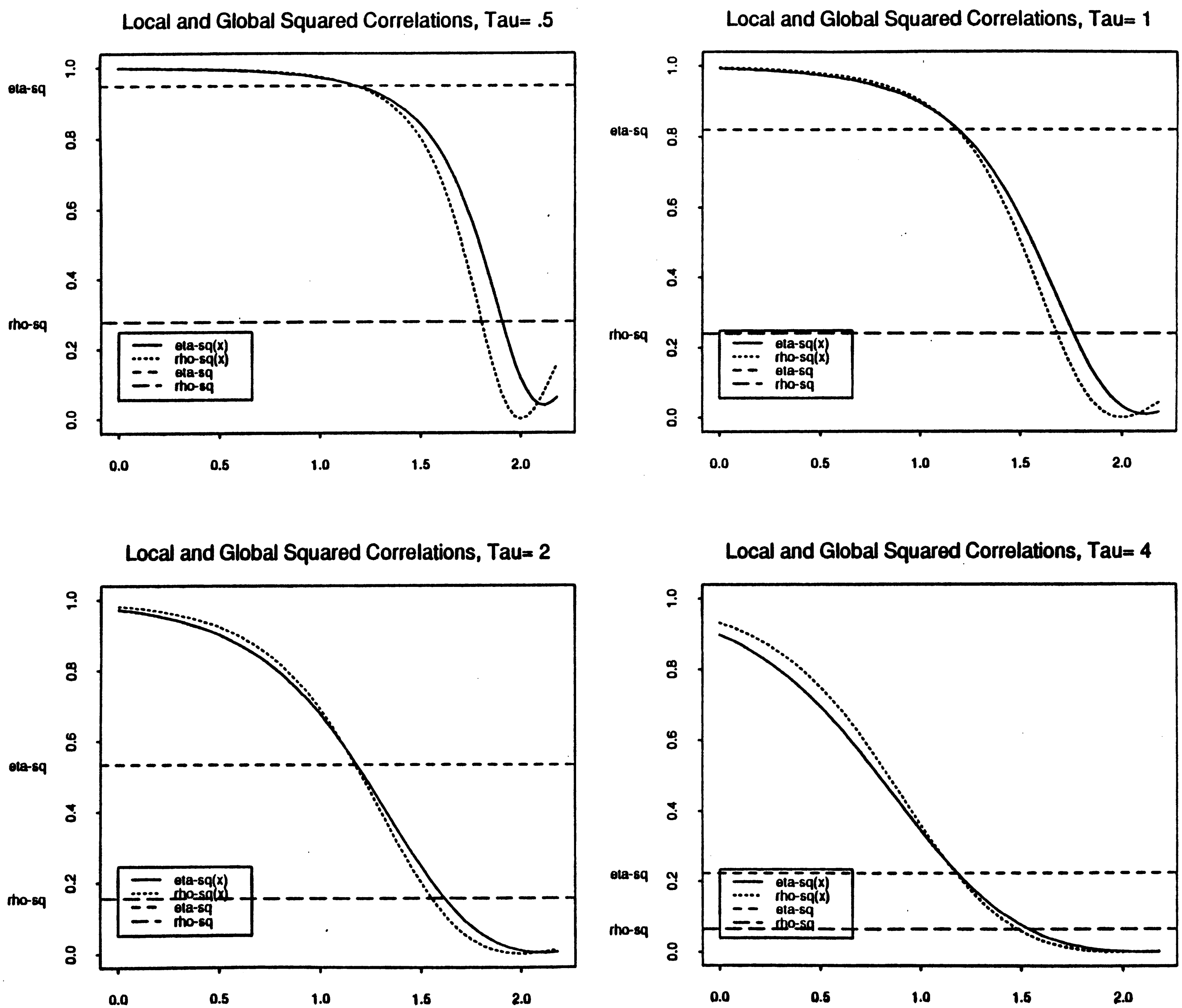
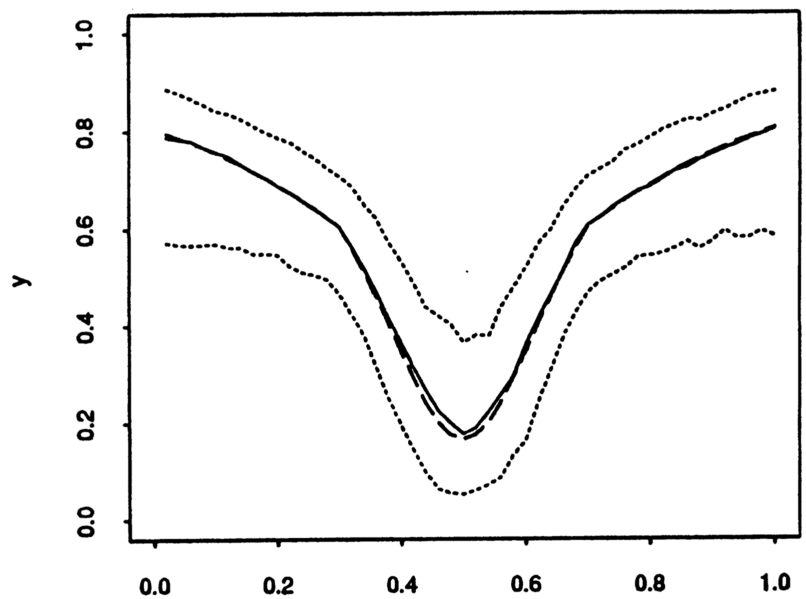


Figure 7.1. Envelope Curves for the Quadratic Model, tau=.1, Random Design

Kernel Estimator, Envelope Curves



Locally Quadratic Estimator, Envelope Curves

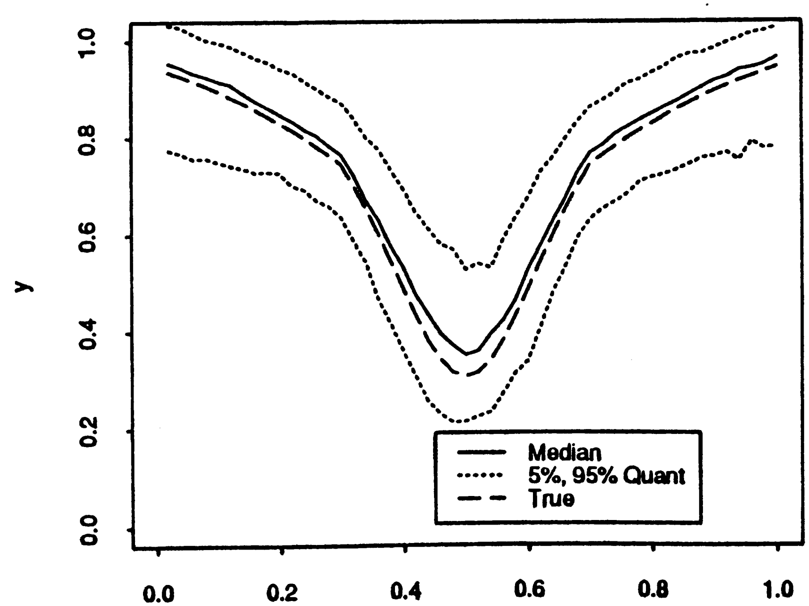


Figure 7.2. Envelope Curves for the Bump Model, Random Design

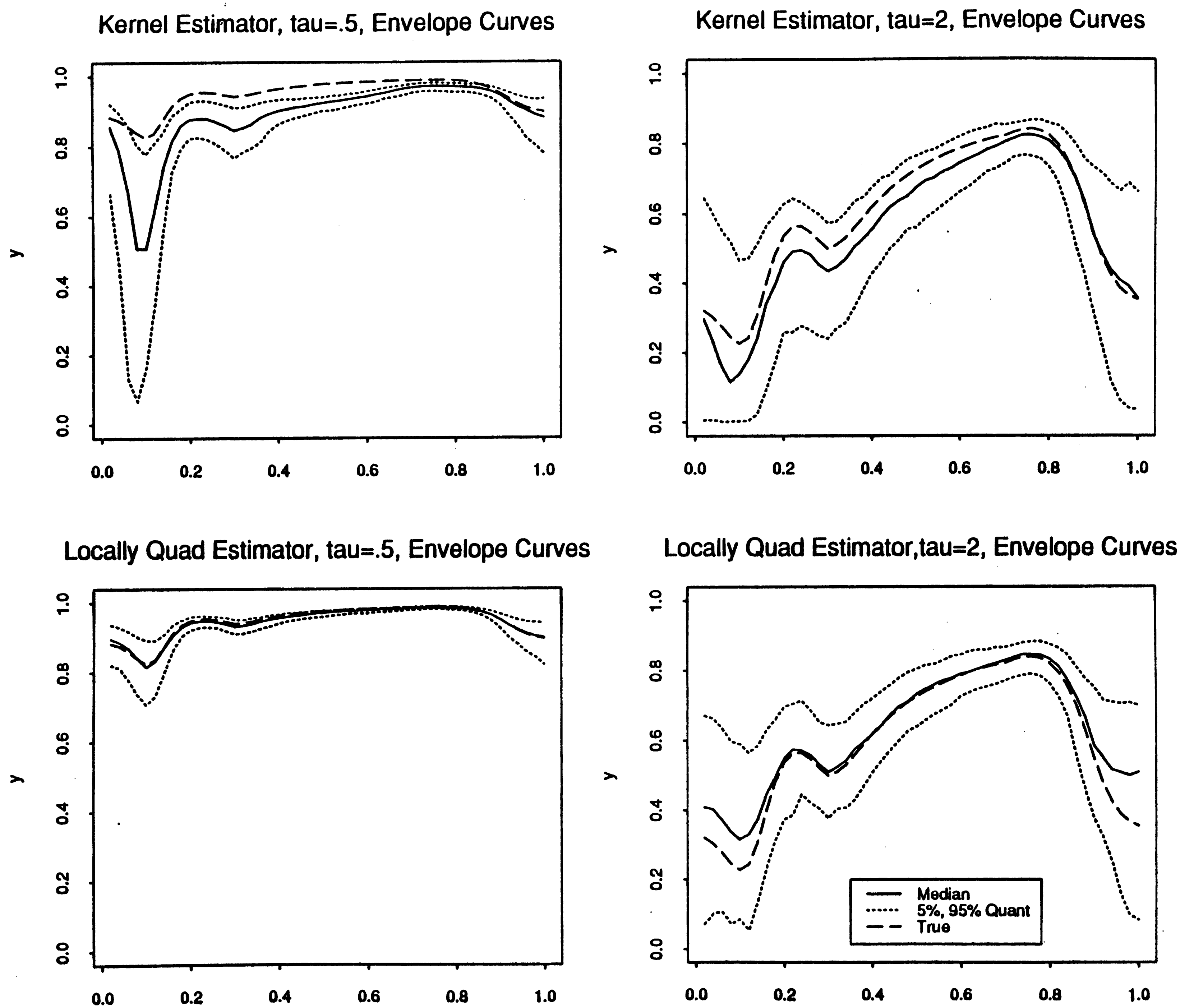


Figure 7.3. Envelope Curves for the Pear Model, Random Design

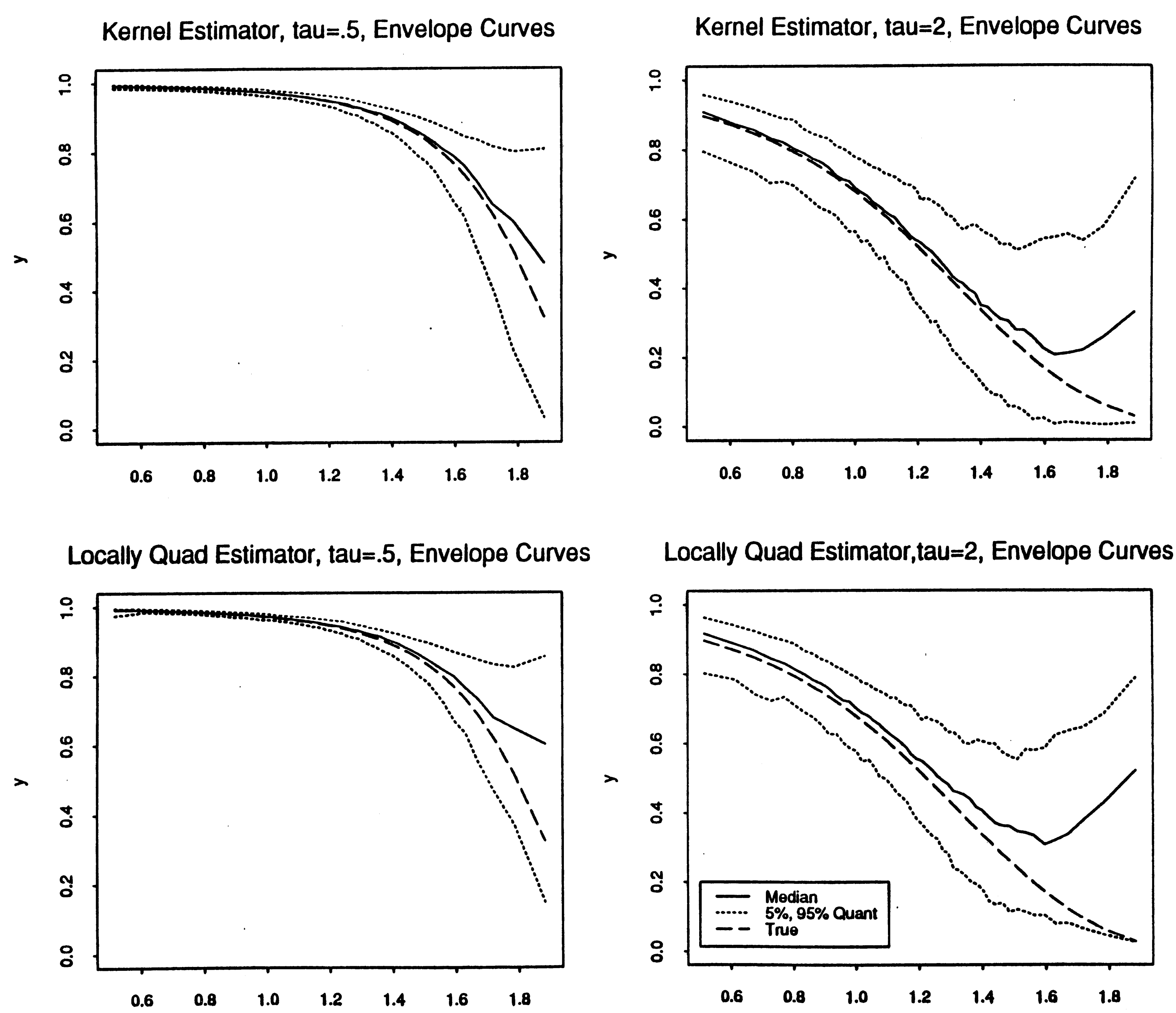


Figure 8.1. Local Squared Correlation Estimates for the Income-Food Expenditure Data