The Annals of Statistics
2001, Vol. 29, No. 5, 1361–1400

NONPARAMETRIC ANALYSIS OF COVARIANCE[1]

BY HOLGER DETTE AND NATALIE NEUMEYER

Ruhr-Universität Bochum

In the problem of testing the equality of k regression curves from independent samples, we discuss three methods using nonparametric estimators of the regression function. The first test is based on a linear combination of estimators for the integrated variance function in the individual samples and in the combined sample. The second approach transfers the classical one-way analysis of variance to the situation of comparing nonparametric curves, while the third test compares the differences between the estimates of the individual regression functions by means of an L^2-distance. We prove asymptotic normality of all considered statistics under the null hypothesis and local and fixed alternatives with different rates corresponding to the various cases. Additionally, consistency of a wild bootstrap version of the tests is established. In contrast to most of the procedures proposed in the literature, the methods introduced in this paper are also applicable in the case of different design points in each sample and heteroscedastic errors. A simulation study is conducted to investigate the finite sample properties of the new tests and a comparison with recently proposed and related procedures is performed.

1. Introduction. An important problem in applied regression analysis is the comparison of a response Y across several groups in the presence of a covariate effect. In general, this model can be written as

$$Y_{ij} = g_i(t_{ij}) + \sigma_i(t_{ij})\varepsilon_{ij}, \qquad i = 1, \dots, k, \; j = 1, \dots, n_i, \tag{1.1}$$

where ε_{ij} are independently identically distributed errors, g_i, σ_i are the regression and variance function in the ith group ($i = 1, \dots, k$) and the covariate t_{ij} varies in the interval [0, 1]. In this paper we are interested in the problem of testing the equality of the mean functions, that is,

$$\begin{aligned} &H_0\colon g_1 = g_2 = \cdots = g_k \text{ versus} \\ &H_1\colon g_i \neq g_j \qquad \text{for some } i,\ j \in \{1, \dots, k\}. \end{aligned} \tag{1.2}$$

Much effort has been devoted to this problem in the literature [see, e.g., Härdle and Marron (1990), Hall and Hart (1990), King, Hart and Wehrly (1991), Delgado (1993), Young and Bowman (1995), Kulasekera (1995), Kulasekera and Wang (1997), Hall, Huber and Speckman (1997) and Dette and Munk (1998)]. Most authors concentrate on the case of two samples ($k = 2$) and a homoscedastic error in all groups. Härdle and Marron (1990) consider a semiparametric approach in the case of equal designs (i.e., $n_1 = \cdots = n_k$,

Received September 1999; revised May 2001.
[1]Supported by the Deutsche Forschungsgemeinschaft, DE 502/9-1, SFB 475.
AMS 2000 *subject classifications*. Primary 62G05.
Key words and phrases. Nonparametric analysis of covariance, variance estimation, comparison of regression curves, goodness of fit, wild bootstrap.

$t_{ij} = t_j, i = 1, \dots, k$). King, Hart and Wehrly (1991) and Hall and Hart (1990) discuss the completely nonparametric homoscedastic [i.e., $\sigma_i^2(t) = \sigma_i^2$, $i = 1, \dots, k$] model in the case of equal design points. While the latter approach can be generalized to the case of unequal design points [see Hall and Hart (1990)], Kulasekera (1995) points out some drawbacks of the test in this case and proposes several alternatives which are applicable to the model (1.1) under the additional assumption of homoscedasticity in all groups. This approach can detect alternatives which converge to the null at a rate of order $1/\sqrt{n}$ but the author also mentions some practical problems of this procedure. On the one hand, the level of the test depends sensitively on the smoothing parameter; on the other hand, larger noises yield levels substantially different from the nominal levels [see also Kulasekera and Wang (1997) for a detailed simulation study and data-driven guidelines for bandwidth selection]. Moreover, a generalization to a heteroscedastic error or a multivariate predictor seems to be difficult. A rather different test was introduced by Young and Bowman (1995) who generalized the one-way analysis of variance to the model (1.1). Under the assumption of normally distributed homoscedastic errors over all groups, these authors proposed a χ^2-approximation for the distribution of the test statistic. Although the finite sample properties of the test under these assumptions look promising, a generalization to the general heteroscedastic, nonnormal case does not appear trivial and the asymptotic properties of this test have not been investigated so far. To our knowledge the problem of testing the equality of the regression functions in the completely heteroscedastic model (1.1) with a univariate predictor and unequal design points was first considered by Dette and Munk (1998), who introduced a simple estimator of the distance

$$\sum_{i<j} \int_0^1 \left[g_i(t) - g_j(t)\right]^2 dt$$

and proved an asymptotic normal law with a $\sqrt{n}$-rate for a corresponding test statistic. As a consequence, this test can only detect alternatives which converge to the null hypothesis at a rate of order $n^{-1/4}$.

In this paper we discuss various tests for the hypothesis (1.2) which are directly applicable in the general model (1.1), do not require any additional assumptions (such as homoscedasticity or equal design points) and improve on the asymptotic efficiency of the test of Dette and Munk (1998). Moreover, the new methods can easily be extended to the case of multivariate predictors. A first method for testing the hypothesis (1.2) is based on a difference between a nonparametric variance estimator in the combined sample $\{Y_{ij} | j = 1, \dots, n_i,\ i = 1, \dots, k\}$ and the corresponding estimators in the individual samples $\{Y_{ij} | j = 1, \dots, n_i\}$ and yields in fact an estimator of an alternative measure of equality. Our second proposal is to use Young and Bowman's (1995) test also in the situation of a heteroscedastic error. Finally, we suggest a generalization of King, Hart and Wehrly's (1991) test to the general setup (1.1), which compares the estimates of the regression functions in the individual samples. This method is closely related to an approach introduced by Rosenblatt (1975) in the context of testing independence and further developed by Härdle and Mammen (1993) and González-Manteiga and Cao (1993) for the problem of testing the parametric form of a regression function. We prove asymptotic normality of all proposed test statistics under the null hypothesis and fixed alternatives with different rates of convergence corresponding to both cases. In Section 2 we introduce the different methods, state the main asymptotic results and discuss various links among the different approaches. In Section 3 we investigate the finite sample properties of some of the proposed tests and perform a comparison with alternative procedures which have been suggested in the literature [see Hall and Hart (1990), Delgado (1993), Kulasekera (1995), Kulasekera and Wang (1997)]. It is demonstrated that a wild bootstrap version of the test based on the difference of variance estimators has excellent finite sample properties and is very often remarkably more powerful than several other tests proposed in the literature, which can detect alternatives converging to the null at a parametric rate. Finally, some of the proofs, which are cumbersome, are given in Section 4.

2. Testing the equality of regression functions by kernel-based methods. To motivate the different methods for testing the hypothesis of the form (1.2) and to investigate the asymptotic distribution of the corresponding test statistics, we need a few regularity assumptions. Let $N = \sum_{i=1}^{k} n_i$ denote the total sample size and assume

$$\frac{n_i}{N} = \kappa_i + O\left(\frac{1}{N}\right), \qquad i = 1, \dots, k, \tag{2.1}$$

for given constants $\kappa_1, \dots, \kappa_k \in (0, 1)$. Let $r_1, \dots, r_k$ denote positive densities on the interval $[0, 1]$ such that the design points t_{ij} satisfy

$$\int_0^{t_{ij}} r_i(t)\, dt = \frac{j}{n_i}, \qquad j = 1, \dots, n_i,\ i = 1, \dots, k \tag{2.2}$$

[see Sacks and Ylvisaker (1970)]. Throughout this paper we will assume the continuity of the variance functions $\sigma_i^2(\cdot)$, $i = 1, \dots, k$, and additionally that the design densities in (2.2) and the regression functions are sufficiently smooth, that is,

$$g_j,\ r_j \in C^{(r)}[0, 1], \qquad j = 1, \dots, k, \tag{2.3}$$

where $r \geq 2$ and $C^{(r)}[0, 1]$ denotes the space of r-times continuously differentiable functions on the interval $[0, 1]$. Let

$$\hat{g}_i(t) = \frac{\sum_{j=1}^{n_i} K\big((t - t_{ij})/h_i\big) Y_{ij}}{\sum_{j=1}^{n_i} K\big((t - t_{ij})/h_i\big)} \tag{2.4}$$

denote the Nadaraya–Watson estimator of the ith regression function g_i and h_i the corresponding bandwidth [see Nadaraya (1964) and Watson (1964)]. We assume that the kernel in (2.4) is supported on a compact interval, say

$[-1, 1]$, and is of order $r \geq 2$ [see Gasser, Müller and Mammitzsch (1985)], that is,

$$\frac{(-1)^j}{j!} \int_{-1}^{1} K(u) u^j \, du = \begin{cases} 1, & j = 0, \\ 0, & 1 \leq j \leq r-1, \\ k_r \neq 0, & j = r, \end{cases} \tag{2.5}$$

where $\int_{-1}^{1} K^2(u) \, du < \infty$.

If the hypothesis of equal regression functions is valid, the total sample could be used to estimate the common regression, that is,

$$\hat{g}(t) = \frac{\sum_{i=1}^{k} \sum_{j=1}^{n_i} K\big((t - t_{ij})/h\big) Y_{ij}}{\sum_{i=1}^{k} \sum_{j=1}^{n_i} K\big((t - t_{ij})/h\big)}, \tag{2.6}$$

where h is an additional bandwidth. For the sake of simplicity, the asymptotic analysis of the statistics proposed below is performed for the case of equal bandwidths $h_i = h, i = 1, \ldots k$, in the estimators (2.4) and (2.6) where the bandwidth h satisfies

$$Nh^2 \to \infty, \qquad h = O(N^{-2/(4r+1)}). \tag{2.7}$$

2.1. *Comparing variance estimators.* Our first measure of equality between the different regression functions is obtained by comparing the nonparametric variance estimators of the individual samples with a nonparametric variance estimator of the pooled sample. To be precise, let

$$\hat{\sigma}_i^2 = \frac{1}{n_i} \sum_{j=1}^{n_i} \big(Y_{ij} - \hat{g}_i(t_{ij})\big)^2, \qquad i = 1, \ldots, k, \tag{2.8}$$

denote the estimator of the variance of the ith sample introduced by Hall and Marron (1990), where $\hat{g}_i$ is the nonparametric estimator of the regression function in the ith sample defined in (2.4). Although these authors considered only a homoscedastic model, it will be shown (see Lemma 4.0) that in the heteroscedastic model $\hat{\sigma}_i^2$ consistently estimates the integrated variance function $\int_0^1 \sigma_i^2(t) r_i(t) \, dt$ of each sample ($i = 1, \ldots, k$). In the following we will consider the analogue of (2.8) for the total sample size

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^{k} \sum_{j=1}^{n_i} \big(Y_{ij} - \hat{g}(t_{ij})\big)^2. \tag{2.9}$$

It is proved in Section 4 that under the hypothesis of equal regression curves this is essentially an estimator for a convex combination of the individual integrated variance functions, that is,

$$\sum_{i=1}^{k} \kappa_i \int_0^1 \sigma_i^2(t) r_i(t) \, dt.$$

For these reasons we propose as a test statistic

$$T_N^{(1)} = \hat{\sigma}^2 - \frac{1}{N}\sum_{i=1}^{k} n_i \hat{\sigma}_i^2.$$

The asymptotic properties of the statistic $T_N^{(1)}$ are listed in the following theorem.

THEOREM 2.1. *Assume that* (2.1), (2.2), (2.3), (2.5) *and* (2.7) *are satisfied.*

(i) *If the hypothesis of equal regression functions is valid, then the statistic* $T_N^{(1)}$ *satisfies*

$$N\sqrt{h}\Big(T_N^{(1)} - B_k^{(1)}h^{2r} - \frac{1}{Nh}D_k^{(1)}\Big) \xrightarrow{\mathscr{D}} \mathscr{N}(0,\ \beta_{k,1}^2),$$

where

$$\begin{aligned} B_k^{(1)} &= k_r^2 \int_0^1 \big((g_1\overline{R})^{(r)} - g_1\overline{R}^{(r)}\big)^2(t)\frac{dt}{\overline{R}(t)} \\ &\quad - k_r^2 \sum_{j=1}^{k} \kappa_j \int_0^1 \big((g_1 r_j)^{(r)} - g_1 r_j^{(r)}\big)^2(t)\frac{dt}{r_j(t)}, \end{aligned} \tag{2.10}$$

$$D_k^{(1)} = \Big[\int_{-1}^{1} K^2(u)\,du - 2K(0)\Big]\sum_{j=1}^{k}\Big(\int_0^1 \frac{\kappa_j\sigma_j^2(t)r_j(t)}{\overline{R}(t)}\,dt - \int_0^1 \sigma_j^2(t)\,dt\Big),$$

the asymptotic variance is given by

$$\begin{aligned} \beta_{k,1}^2 &= 2\int_{-1}^{1}(2K - K*K)^2(u)\,du \\ &\quad \times\Big\{\sum_{j=1}^{k}\int_0^1 \sigma_j^4(t)\Big(\frac{\kappa_j r_j(t)}{\overline{R}(t)} - 1\Big)^2 dt \\ &\qquad + \sum_{j=1}^{k}\sum_{\substack{l=1 \\ l\neq j}}^{k}\int_0^1 \sigma_j^2(t)\sigma_l^2(t)\frac{\kappa_j r_j(t)\kappa_l r_l(t)}{\overline{R}^2(t)}\,dt\Big\}, \end{aligned} \tag{2.11}$$

$K * K$ *is the convolution of the kernel with itself and* $\overline{R}$ *denotes the convex combination of the underlying densities, that is,*

$$\overline{R}(t) = \sum_{j=1}^{k}\kappa_j r_j(t).$$

(ii) *Under the alternative* $g_i \neq g_j$, *for some* $i,\ j \in \{1,\dots,k\}$, *the statistic* $T_N^{(1)}$ *satisfies*

$$\sqrt{N}(T_N^{(1)} - M_{k,1}^2) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \gamma_{k,1}^2),$$

where

$$M_{k,1}^2 = \sum_{j=1}^{k} \sum_{\substack{l=1 \\ l<j}}^{k} \int_0^1 (g_j - g_l)^2(t) \frac{\kappa_j r_j(t) \kappa_l r_l(t)}{\bar{R}(t)} \, dt \tag{2.12}$$

and the asymptotic variance is given by

$$\gamma_{k,1}^2 = 4 \sum_{j=1}^{k} \int_0^1 \left(\sum_{\substack{l=1 \\ l \neq j}}^{k} (g_j(t) - g_l(t)) \frac{\kappa_l r_l(t)}{\bar{R}(t)} \right)^2 \sigma_j^2(t) \kappa_j r_j(t) \, dt. \tag{2.13}$$

2.2. *An ANOVA-type statistic.* The following method for testing the equality of the regression functions was introduced by Young and Bowman (1995) in the context of a homoscedastic normal distribution for the error over all k samples. The corresponding statistic is closely related to the difference of variance estimators introduced in Section 2.1. It will be shown in this section that the method proposed by these authors is also applicable in the general situation of nonnormal heteroscedastic errors. The test statistic of Young and Bowman (1995) is motivated by the classical one-way analysis of variance and given by

$$Y_N = \frac{N}{\hat{s}^2} T_N^{(2)},$$

where

$$T_N^{(2)} = \frac{1}{N} \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left[\hat{g}(t_{ij}) - \hat{g}_i(t_{ij}) \right]^2, \tag{2.14}$$

$\hat{g}, \hat{g}_1, \dots, \hat{g}_k$ are defined in (2.6) and (2.4), respectively, and

$$\hat{s}^2 = \frac{1}{2(N-k)} \sum_{i=1}^{k} \sum_{j=1}^{n_i - 1} (Y_{i,j+1} - Y_{i,j})^2$$

is a pooled mean of the difference-based estimators for the variances in the individual samples [see, e.g., Rice (1984)]. The statistic $\hat{s}^2$ is also a consistent estimator of

$$s^2 = \sum_{j=1}^{k} \kappa_j \int_0^1 \sigma_j^2(t) r_j(t) \, dt, \tag{2.15}$$

which follows by a straightforward calculation [see also Kulasekera and Wang (1997) for a related result under homoscedasticity]. Note that there is a strong link between the statistics $T_N^{(1)}$ and $T_N^{(2)}$. While the statistic $T_N^{(1)}$ is comparing the regression functions through residual sums of squares, the statistic $T_N^{(2)}$ compares the curves through the fitted values. In the case of a fixed design, equal homoscedastic variances in all groups and a normally distributed error, Young and Bowman (1995) proposed a χ^2-approximation of the corresponding test statistic under the null hypothesis. These restrictions allow a rapid and accurate calculation of the p-value [see Young and Bowman (1995) for more details]. It is also worthwhile mentioning that the use of the same smoothing parameters in the estimates of the individual regression function yields a direct cancellation of the bias.

Obviously, the numerator of Y_N given in (2.14) is an estimate for an appropriate measure of equality of the k regression curves and we will also use $T_N^{(2)}$ as a test statistic for the hypothesis (1.2) in the general situation of not necessarily homoscedastic and normally distributed errors. The following result makes this heuristic argument more precise and provides the asymptotic properties of the statistic (2.14). As a by-product, it also proves consistency of the test proposed by Young and Bowman (1995) if the required assumptions for the finite sample size approximation used by these authors are not satisfied. Moreover, critical values could be obtained from an approximation by a normal distribution or a wild bootstrap procedure as proposed in Section 3.

THEOREM 2.2. *Assume that* (2.1), (2.2), (2.3), (2.5) *and* (2.7) *are satisfied*.

(i) *If the hypothesis of equal regression curves is valid*, *then the statistic* $T_N^{(2)}$ *defined in* (2.14) *satisfies*

$$N\sqrt{h}\left(T_N^{(2)} + B_k^{(2)} h^{2r} - \frac{D_k^{(2)}}{Nh}\right) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \beta_{k,2}^2),$$

where $B_k^{(2)} = B_k^{(1)}$ *is defined in* (2.10),

$$D_k^{(2)} = \int_{-1}^{1} K^2(u)\,du \sum_{j=1}^{k} \int_0^1 \left(1 - \kappa_j \frac{r_j(t)}{\overline{R}(t)}\right) \sigma_j^2(t)\,dt \tag{2.16}$$

and the asymptotic variance is given by

$$\begin{aligned} \beta_{k,2}^2 = 2\int_{-1}^{1} (K*K)^2(u)\,du & \\ \times \Biggl\{ \sum_{j=1}^{k} \int_0^1 \sigma_j^4(t) \left(\frac{\kappa_j r_j(t)}{\overline{R}(t)} - 1\right)^2 dt & \\ + \sum_{j=1}^{k} \sum_{\substack{l=1 \\ l\neq j}}^{k} \int_0^1 \sigma_j^2(t)\sigma_l^2(t) \frac{\kappa_j r_j(t) \kappa_l r_l(t)}{\overline{R}^2(t)}\,dt \Biggr\}. & \end{aligned} \tag{2.17}$$

(ii) *Under the alternative* $g_i \neq g_j$, *for some* $i, j \in \{1, \dots, k\}$, *the statistic* $T_N^{(2)}$ *defined in* (2.14) *satisfies*

$$\sqrt{N}\left(T_N^{(2)} - M_{k,2}^2\right) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \gamma_{k,2}^2),$$

where $M_{k,2}^2 = M_{k,1}^2$ *and* $\gamma_{k,2}^2 = \gamma_{k,1}^2$ *are defined in* (2.12) *and* (2.13), *respectively*.

2.3. *Pairwise comparison of regression curves.* Following Rosenblatt (1975), Härdle and Mammen (1993) and González-Manteiga and Cao (1993), an obvious alternative test of the hypothesis (1.2) could be obtained from a pairwise comparison of the estimators of the regression functions. To this end, we consider the statistic

$$T_N^{(3)} = \sum_{i=1}^{k} \sum_{j=1}^{i-1} \int_0^1 \left[\hat{g}_i(t) - \hat{g}_j(t)\right]^2 w_{ij}(t)\,dt, \tag{2.18}$$

where $w_{ij}(\cdot)$ are positive weight functions satisfying $w_{ij} = w_{ji}$, $1 \le j < i \le k$. A similar statistic was considered by King, Hart and Wehrly (1991) in the case of two samples with equal design points (here the integral was approximated by a sum and a constant weight was used). A calculation similar to that given in the proof of Lemma 4.0 shows that

$$E\left[T_N^{(3)}\right] = \begin{cases} \dfrac{1}{Nh} D_k^{(3)} + O\left(\dfrac{1}{N}\right), & \text{under } H_0, \\ \dfrac{1}{Nh} D_k^{(3)} + M_{k,3}^2 + O\left(\dfrac{1}{N}\right), & \text{under } H_1, \end{cases} \tag{2.19}$$

where the constants $D_k^{(3)}$, $M_{k,3}^2$ are defined by

$$D_k^{(3)} = \int K^2(u)\,du \sum_{j=1}^{k} \int_0^1 \frac{\sigma_j^2(t)}{\kappa_j r_j(t)} \left(\sum_{\substack{l=1 \\ l \neq j}}^{k} w_{jl}(t) \right) dt, \tag{2.20}$$

$$M_{k,3}^2 = \sum_{j=1}^{k} \sum_{\substack{l=1 \\ l<j}}^{k} \int_0^1 (g_j - g_l)^2(t) w_{jl}(t)\,dt. \tag{2.21}$$

Note that, in contrast to Theorems 2.1 and 2.2, there does not appear a term of order h^{2r} in (2.19), which is in fact a result of the application of equal bandwidths in the estimates of the regression functions in the individual samples. The following result can be proved using similar arguments as given for the proof of Theorem 2.1 in Section 4.

THEOREM 2.3. *Assume that* (2.1), (2.2), (2.3), (2.5) *and* (2.7) *are satisfied.*

(i) *If the hypothesis of equal regression curves is valid, then the statistic* $T_N^{(3)}$ *defined in* (2.18) *satisfies*

$$N\sqrt{h}\left(T_N^{(3)} - \frac{1}{Nh} D_k^{(3)}\right) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \beta_{k,3}^2),$$

where the asymptotic variance is defined by

$$\beta_{k,3}^2 = \int_0^1 (K*K)^2(u)\,du \times \left\{ \sum_{j=1}^k \int_0^1 \frac{\sigma_j^4(t)}{\kappa_j^2 r_j^2(t)} \left(\sum_{\substack{l=1 \\ l\neq j}}^k w_{jl}(t)\right)^2 dt + \sum_{j=1}^k \sum_{\substack{l=1 \\ l\neq j}}^k \int \frac{\sigma_j^2(t)\sigma_l^2(t)}{\kappa_j r_j(t) \kappa_l r_l(t)} w_{jl}^2(t)\,dt \right\}. \tag{2.22}$$

(ii) *Under the alternative* $g_i \neq g_j$, *for some* $i, j \in \{1, \dots, k\}$, *the statistic* $T_N^{(3)}$ *defined in* (2.18) *satisfies*

$$\sqrt{N}(T_N^{(3)} - M_{k,3}^2) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \gamma_{k,3}^2),$$

where $M_{k,3}^2$ *is defined in* (2.21) *and the asymptotic variance is given by*

$$\gamma_{k,3}^2 = 4 \sum_{i=1}^k \sum_{\substack{j=1 \\ j\neq i}}^k \sum_{\substack{l=1 \\ l\neq i}}^k \int_0^1 \big(g_i(t) - g_l(t)\big)\big(g_i(t) - g_j(t)\big) w_{ji}(t) w_{li}(t) \frac{\sigma_i^2(t)}{\kappa_i r_i(t)}\,dt.$$

REMARK 2.4. It is worthwhile mentioning that there is a strong link among the three statistics $T_N^{(1)}$, $T_N^{(2)}$, $T_N^{(3)}$, which can nicely be explained by looking at the classical one-way analysis of variance model, where

$$X_{ij} \sim \mathscr{N}(\mu_i, \sigma^2), \qquad j = 1, \dots, n_i,\ i = 1, \dots, k.$$

Here the denominator of the corresponding F-test corresponds to the statistic $T_N^{(2)}$ of Young and Bowman (1995) and can be decomposed as

$$\sum_{i=1}^k n_i \left(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot} \right)^2 = \sum_{i=1}^k \sum_{j=1}^{n_i} \left(X_{ij} - \overline{X}_{\cdot\cdot} \right)^2 - \sum_{i=1}^k \sum_{j=1}^{n_i} \left(X_{ij} - \overline{X}_{i\cdot} \right)^2, \tag{2.23}$$

where the first term on the right-hand side is an estimator of the variance from the pooled sample (assuming equal means in all k samples) and the second term is a combination of the variance estimators in the individual samples. Consequently, the right-hand side of (2.23) corresponds to the statistic $T_N^{(1)}$ introduced in Section 2.1. Therefore in linear models both statistics are equivalent, while for nonparametric models there appear to be differences because the cross-product terms involve a nonvanishing bias. Similarly, we have the representation

$$\sum_{i=1}^k n_i (\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})^2 = \frac{1}{2N} \sum_{i=1}^k \sum_{j=1}^k n_i n_j (\overline{X}_{i\cdot} - \overline{X}_{j\cdot})^2,$$

which establishes an analogy between the statistics $T_N^{(2)}$ and $T_N^{(3)}$.

2.4. *Some asymptotic power comparisons.* As a consequence of Theorems 2.1–2.3, we obtain consistent, asymptotic level-α tests by rejecting the hypothesis of equal regression curves whenever

$$N\sqrt{h}\left(T_N^{(i)} - B_k^{(i)}h^{2r} - \frac{D_k^{(i)}}{Nh}\right) > \beta_{k,i}u_{1-\alpha}, \qquad i = 1, 2, 3, \tag{2.24}$$

where $B_k^{(3)} = 0$, $B_k^{(i)}$, $D_k^{(i)}$, $\beta_{k,i}^2$ are defined in Theorems 2.1–2.3 and have to be replaced by consistent estimators. In the following section we will illustrate the performance of a wild bootstrap version of the tests given by (2.24), because the speed of convergence under the null hypothesis is usually rather slow [see also Azzalini and Bowman (1993), Hjellvik and Tjøstheim (1995) and Alcalá, Christóbal and González-Manteiga (1999) for similar observations].

Moreover, the second parts of Theorems 2.1–2.3 provide an important advantage in the application of these tests (compared to most of the procedures proposed in the literature). It is well known that in the problem of testing goodness of fit the essential error is the type II error and a large observed p-value does not give any empirical evidence for the null hypothesis [see, e.g., Berger and Delampady (1987) and Staudte and Sheater (1990)]. The second parts of Theorems 2.1–2.3 now provide an approximation for the type II error of the test by

$$\begin{aligned} P(\text{“rejection”}) &\sim \Phi\left(\frac{\sqrt{N}}{\gamma_{k,i}}\left\{M_{k,i}^2 - \frac{u_{1-\alpha}\beta_{k,i}}{N\sqrt{h}}\right\}\right) \\ &\sim \Phi\left(\frac{\sqrt{N}M_{k,i}^2}{\gamma_{k,i}}\right), \qquad i = 1, 2, 3. \end{aligned} \tag{2.25}$$

We note that the approximation by a normal distribution under fixed alternatives is more reliable than under the null hypothesis, because it is similar to the approximation by a normal distribution in the classical central limit theorem (see the proof in Section 4.3). Moreover, the second parts of Theorems 2.1–2.3 can also be used for testing the precise hypotheses [see Berger and Delampady (1987)]

$$H_0\colon M_{k,i}^2 > \Delta \quad \text{versus} \quad H_1\colon M_{k,i}^2 \le \Delta, \tag{2.26}$$

where Δ is a sufficiently small constant such that, whenever $M_{k,i}^2 \le \Delta$, the experimenter agrees to analyze the data under the assumption of equal regression curves. Note that the rejection of H_0 in (2.26) allows us to show that the regression functions are “close” at a controlled error rate.

REMARK 2.5. The possibility of choosing the weight functions in (2.18) leaves some freedom for the statistic $T_N^{(3)}$ and using

$$w_{ij} = \frac{\kappa_i r_i \kappa_j r_j}{\bar{R}} \tag{2.27}$$

gives a statistic with an asymptotically similar behavior as described in Theorems 2.1 and 2.2 for the tests based on $T_N^{(1)}$ and $T_N^{(2)}$. This weight function is very natural because under the additional assumption of homoscedasticity it maximizes the asymptotic power for comparing the curves g_i and g_j with respect to the choice of the weight function. To be precise, assume that $k = 2$. Then a straightforward calculation [see also the derivation of (2.25)] shows that the probability of rejection is an increasing function of

$$
\begin{aligned}
\frac{(M_{2,3}^2)^2}{\gamma_{2,3}^2} &= \frac{\left(\int_0^1 (g_1-g_2)^2(t) w_{12}(t)\,dt\right)^2}{\sigma^2 \int_0^1 (g_1-g_2)^2(t) w_{12}^2(t)(1/\kappa_1 r_1(t) + 1/\kappa_2 r_2(t))\,dt} \\
&\le \frac{1}{\sigma^2} \int_0^1 (g_1-g_2)^2 \left(\frac{1}{\kappa_1 r_1(t)} + \frac{1}{\kappa_2 r_2(t)}\right)^{-1} dt \qquad (2.28) \\
&= \frac{(M_{2,1}^2)^2}{\gamma_{2,1}^2},
\end{aligned}
$$

where $M_{2,1}^2 = M_{2,2}^2$ and $\gamma_{2,1}^2 = \gamma_{2,2}^2$ are defined in (2.12) and (2.13), respectively, and the second line follows from Cauchy's inequality. Now discussing the equality in (2.28) shows that the maximal power (with respect to the choice of the weight function w_{12}) is obtained by the weight (2.27). For these reasons, the tests based on $T_N^{(2)}$ [Young and Bowman (1995)] and $T_N^{(1)}$ (proposed in Section 2) should be preferred because they automatically adapt to the best possible (but unknown) weight function for the maximization of the power at any fixed alternative.

REMARK 2.6. In the remaining part of this section, we will concentrate on the asymptotic behavior of the different tests with respect to local alternatives. For the sake of transparency, we will concentrate on the case of $k = 2$ samples. There is no difference in the discussion of the general situation of $k \ge 3$ regression functions. We will adopt an approach of Rosenblatt (1975), who considered alternatives of the form

$$g_2(\cdot) = g_1(\cdot) + \delta_N s\left(\frac{\cdot - c}{\gamma_N}\right), \qquad (2.29)$$

where s is a continuously differentiable function of order r and $c \in [0, 1]$ a given constant. In addition, δ_N and γ_N are sequences converging to 0 such that

$$\delta_N^2 \gamma_N = \frac{1}{N\sqrt{h}}, \qquad \delta_N = o(1), \qquad \gamma_N^{-1} = o\left(h^{-2r/(2r-1)}\right) \qquad (2.30)$$

(a typical example in the case $r = 2$ is $h = N^{-2/9}$, $\delta_N = N^{-13/36}$ and $\gamma_N = N^{-1/6}$). For alternatives of the form (2.29) satisfying (2.30), it follows by similar arguments as given in Section 4 for the proof of Theorem 2.1 that

$$N\sqrt{h}\left(T_N^{(i)} - B_k^{(i)} h^{2r} - \frac{D_k^{(i)}}{Nh}\right) \xrightarrow{\mathscr{D}} \mathscr{N}(\mu^{(i)},\ \beta_{k,i}^2), \tag{2.31}$$

where the constants $B_k^{(i)}$, $D_k^{(i)}$ and $\beta_{k,i}^2$ are defined in Theorems 2.1–2.3 (note that $B_k^{(3)} = 0$) and

$$\mu^{(i)} = \begin{cases} \displaystyle\int_0^1 s^2(x)\,dx \cdot \frac{\kappa_1\kappa_2 r_1(c) r_2(c)}{\kappa_1 r_1(c) + \kappa_2 r_2(c)}, & \text{if } i = 1, 2, \\ \displaystyle\int_0^1 s^2(x)\,dx \cdot w_{12}(c), & \text{if } i = 3. \end{cases} \tag{2.32}$$

A similar result is obtained for local alternatives of the form (2.29) with $c = 0$ and $\gamma_N = 1$, that is, $g_1 = g_2 + s \cdot (N\sqrt{h})^{-1/2}$. In this case (2.31) is still valid, with a different expectation in the limit distribution, that is,

$$\mu^{(i)} = \begin{cases} \displaystyle\int_0^1 s^2(x) \frac{\kappa_1\kappa_2 r_1(x) r_2(x)}{\kappa_1 r_1(x) + \kappa_2 r_2(x)}\,dx, & \text{if } i = 1, 2, \\ \displaystyle\int_0^1 s^2(x) w_{12}(x)\,dx, & \text{if } i = 3. \end{cases} \tag{2.33}$$

For an asymptotic analysis of the three testing procedures with respect to these local alternatives, we use the optimal (but unknown) weight function (2.27) for w_{12} in the definition of the statistic $T_N^{(3)}$. The comparison can now easily be performed by looking at the different variances in (2.31) and observing the relation

$$\int (K*K)^2(x)\,dx \le \int K^2(x)\,dx \le \int (2K - K*K)^2(x)\,dx,$$

which has been proved by Biedermann and Dette (2000). From this inequality it follows that $\beta_{2,2}^2 = \beta_{2,3}^2 \le \beta_{2,1}^2$, and, consequently, the procedures based on $T_N^{(2)}$, $T_N^{(3)}$ are asymptotically more efficient as the test based on $T_N^{(1)}$.

However, some care is necessary with this interpretation, because the speed of convergence in (2.31) is rather slow and the asymptotic analysis usually requires a rather large sample size [see Azzalini and Bowman (1993), Hjellvik and Tjøstheim (1995) and Alcalá, Christóbal and González-Manteiga (1999) for similar observations]. For realistic sample sizes the approximation (2.25) indicates a similar behavior of all three methods. Moreover, for moderate sample sizes the bias always has to be taken into account and the superiority of one of the three methods cannot be established in general. In the examples presented in Section 3, we observe a much better performance of the test based on $T_N^{(1)}$.

2.5. *Generalizations: different bandwidths, smoothing techniques and random design.*

REMARK 2.7. Note that we have assumed the equality of all bandwidths in Sections 2.1–2.3, which substantially simplifies the presentation of the asymptotic results and their proofs. Nevertheless, in practice, it is strictly recommended to choose the bandwidth h_i, $i = 1, \dots, k$, and h of each estimator $\hat{g}_i$ and $\hat{g}$ according to the size of the corresponding sample. If there exist constants $b_1, \dots, b_k \in (0, 1)$ such that these bandwidths satisfy (2.7) and

$$\frac{h_i}{h} = b_i + O\left(\frac{1}{N}\right), \qquad i = 1, \dots, k, \tag{2.34}$$

as $N \to \infty$, similar results as given in Theorems 2.1–2.3 can be established, where the constants $B_k^{(i)}$ and $D_k^{(i)}$ additionally depend on the proportions $b_1, \dots, b_k$ (note that $B_k^{(3)}$ does not vanish in this case). For more details we refer to Neumeyer (1999) and Dette and Neumeyer (1999).

REMARK 2.8. It should also be pointed out that the asymptotic results given in Theorems 2.1–2.3 do not depend on the special structures of the smoothing procedures used in the construction of the variance estimators. We used the Nadaraya–Watson estimator for the calculation of residuals because for this choice the proofs given in Section 4 are more transparent. For example, a local polynomial estimator [see Fan (1992) and Fan and Gijbels (1996)] can be treated in the same way with greater mathematical complexity but without changing the structure of the asymptotic results. Although local polynomial estimators have various advantages for the estimation of the regression function, particularly at the boundaries, our simulation results showed that this superiority is not reflected in the problem of testing the equality of regression functions. A heuristical explanation of this observation is that the methods presented in Sections 2.1–2.3 essentially avoid the direct estimation of the regression function and only use estimates for quantities smoothed by linear integral operators. Nevertheless, there are still theoretical advantages to using local smoothing in the definition of the statistics $T_N^{(i)}$, $i = 1, 2, 3$. On the one hand, the use of these estimators allows weaker assumptions on the design densities, because only the continuity of the design density is required (for local polynomials of odd order). On the other hand, the bias of local polynomials of odd order is the same for all curves irrespective of the design pattern. More precisely, if equal bandwidths are used for the local polynomial estimation of the individual regression functions [see (2.4) and (2.6)], the terms $B_k^{(1)}$ and $B_k^{(2)}$ in Theorems 2.1 and 2.2 vanish, while the kernel K in the asymptotic bias $D_k^{(i)}$ and in the asymptotic variance $\beta_{k,i}^2$, $i = 1, 2, 3$, of Theorems 2.1–2.3 has to be replaced by the equivalent higher order kernel corresponding to the local polynomial estimator [see Wand and Jones (1995)]. A similar observation was made by Alcalá, Christóbal and González-Manteiga (1999) in the context of testing for a parametric form of the regression function.

REMARK 2.9. The test statistics $T_N^{(1)}$, $T_N^{(2)}$ and $T_N^{(3)}$ can be directly used for a multivariate predictor and a random design. Under the assumption of a random design, $t_{i1}, \dots, t_{in_i}$ are realizations of i.i.d. random variables $T_{i1}, \dots, T_{in_i}$ with positive density r_i on the interval $[0, 1]$, $i = 1, \dots, k$. In this case, the first parts of the statements of Theorems 2.1–2.3 regarding the asymptotic behavior of the statistics and the hypothesis of equal regression functions remain valid and consistent tests are obtained exactly as in the case of a fixed design. However, it is worthwhile mentioning that under the alternative a different asymptotic variance is obtained in all three cases. Consider, for example, the situation of Theorem 2.1 in the case of $k = 2$ independent random samples. Under a fixed alternative, the asymptotic variance of the statistic $T_N^{(1)}$ is given by

$$\begin{aligned}
\operatorname{var}(T_N^{(1)}) = \frac{4\kappa_1\kappa_2}{N} \int &(g_1 - g_2)^2(t)\big(\kappa_2 r_2(t)\sigma_1^2(t) + \kappa_1 r_1(t)\sigma_2^2(t)\big) \\
&\times \frac{r_1(t)r_2(t)}{(\kappa_1 r_1 + \kappa_2 r_2)^2(t)}\, dt \\
&+ \frac{\kappa_1\kappa_2^2}{N} \operatorname{var}\left((g_1 - g_2)^2(T_{11}) \frac{\kappa_2 r_2^2(T_{11}) + 2\kappa_1 r_1(T_{11}) r_2(T_{11})}{(\kappa_1 r_1 + \kappa_2 r_2)^2(T_{11})} \right) \\
&+ \frac{\kappa_1^2\kappa_2}{N} \operatorname{var}\left((g_1 - g_2)^2(T_{21}) \frac{\kappa_1 r_1^2(T_{21}) + 2\kappa_2 r_1(T_{21}) r_2(T_{21})}{(\kappa_1 r_1 + \kappa_2 r_2)^2(T_{21})} \right) \\
&+ o\left(\frac{1}{N}\right).
\end{aligned}$$

3. Simulation results. In similar problems it was observed by several authors [see, e.g., Azzalini and Bowman (1993), Hjellvik and Tjøstheim (1995) and Alcalá, Christóbal and González-Manteiga (1999)] that the asymptotic normal distribution under the null hypothesis does not provide a satisfactory approximation of the distribution of the statistics $T_N^{(i)}$ for reasonable sample sizes. For these reasons, many authors propose the application of bootstrap procedures in these problems [see, e.g., Hall and Hart (1990) and Härdle and Mammen (1993)]. In this section we study the finite sample performance of a wild bootstrap version of the test (2.24) and compare its power properties with several other procedures suggested in the literature. Some remarks regarding the consistency of this procedure are given in Section 4.4.

Because all simulation results published so far consider the two-sample case with equal homoscedastic variance [i.e., $\sigma_1^2(t) = \sigma_2^2(t) = \sigma^2$], and we are interested in a comparison, we mainly restrict our study to this case. Moreover, we will concentrate on the statistic $T_N^{(1)}$ based on the difference of variance estimators, because it performed better than $T_N^{(2)}$ (see Section 3.2) and it does not require the specification of a weight function (in contrast to the statistic $T_N^{(3)}$). In our study we used in fact an asymptotic equivalent test statistic given by

$$\widetilde{T}_N = \hat{\sigma}^2 - \frac{n_1^2}{N\nu_1}\hat{\sigma}_1^2 - \frac{n_2^2}{N\nu_2}\hat{\sigma}_2^2,$$

where n_1/ν_1, n_2/ν_2 are normalizing constants converging to 1, such that $\hat{\sigma}_i^2$ is unbiased for constant regression g_i [see Hall and Marron (1990)]. More precisely, these constants are defined by

$$\nu_l = n_l - 2\sum_{i=1}^{n_l} w_{ii}^{(l)} + \sum_{i,k=1}^{n_l} (w_{ik}^{(l)})^2, \qquad l = 1, 2,$$

where

$$w_{ik}^{(l)} = \frac{K\big((t_{li} - t_{lk})/h_l\big)}{\sum_{s=1}^{n_l} K\big((t_{li} - t_{ls})/h_l\big)}, \qquad l = 1, 2.$$

We used the common wild bootstrap of residuals based on a nonparametric fit [see Härdle and Mammen (1993)]

$$\hat{\varepsilon}_{ij} = Y_{ij} - \hat{g}(t_{ij}), \qquad j = 1, \dots, n_i,\ i = 1, 2, \tag{3.1}$$

where $\hat{g}$ is the estimator of the regression curve from the total sample defined in (2.6). Let V_{ij}^*, $i = 1, 2$, $j = 1, \dots, n_i$, denote i.i.d. random variables with masses $(\sqrt{5}+1)/2\sqrt{5}$ and $(\sqrt{5}-1)/2\sqrt{5}$ at the points $(1-\sqrt{5})/2$ and $(1+\sqrt{5})/2$ (note that this distribution satisfies $E^*[V_{ij}^*] = 0$, $E^*[V_{ij}^{*2}] = E^*[V_{ij}^{*3}] = 1$). Finally, define $\varepsilon_{ij}^* = V_{ij}^*\hat{\varepsilon}_{ij}$ and the bootstrap sample by

$$Y_{ij}^* = \hat{g}(t_{ij}) + \varepsilon_{ij}^*, \qquad j = 1, \dots, n_i,\ i = 1, 2. \tag{3.2}$$

For the test at level α, the null hypothesis is rejected if $\widetilde{T}_N$ is bigger than the corresponding quantile of the bootstrap distribution of $\widetilde{T}_N$, that is,

$$\widetilde{T}_N > \widetilde{T}^*_{N(\lfloor B(1-\alpha)\rfloor)}, \tag{3.3}$$

where $\widetilde{T}^*_{N(i)}$ denotes the ith-order statistic of the bootstrap sample $\widetilde{T}^*_{N,1}, \dots, \widetilde{T}^*_{N,B}$. In our study we resampled $B = 200$ times and used 1000 simulations for the calculation of the level and power in each scenario. Moreover, we used the same bandwidth for the generation of the bootstrap sample (3.1) and the definition of the test statistic $\widetilde{T}_N$. The consistency of this procedure is indicated in Section 4.4.

We considered two samples at the design points

$$\begin{aligned} t_{1i} &= \frac{i-1}{n_1 - 1}, \qquad && i = 1, \dots, n_1, \\ t_{2j} &= \frac{j}{n_2}, \qquad && j = 1, \dots, n_2, \end{aligned} \tag{3.4}$$

and normally distributed errors in both samples unless stated otherwise, that is,

$$\varepsilon_{1l}, \varepsilon_{2l} \sim \mathcal{N}(0, \sigma^2). \tag{3.5}$$

The kernel was chosen as $K(x) = \frac{3}{4}(1 - x^2)I_{[-1,1]}(x)$ (which yields $r = 2$) and the bandwidths are

$$h_i = \left(\frac{\int_0^1 \sigma_i^2(t)\,dt}{n_i}\right)^{3/10} = \left(\frac{\sigma_i^2}{n_i}\right)^{3/10}, \qquad i = 1, 2, \tag{3.6}$$

$$h = \left(\frac{n_1 \int_0^1 \sigma_1^2(t)\,dt + n_2 \int_0^1 \sigma_2^2(t)\,dt}{(n_1 + n_2)^2}\right)^{3/10} = \left(\frac{\sigma_1^2}{n_1 + n_2}\right)^{3/10}, \tag{3.7}$$

where the last equalities follow in the case of homoscedasticity and $\sigma_1^2 = \sigma_2^2$. Note that we use different bandwidths for the estimators $\hat{g}$, $\hat{g}_1$ and $\hat{g}_2$ in our study.

3.1. *Simulation of the level.* Our first example investigates the approximation of the level by the wild bootstrap version of the test (2.24). First, we considered quadratic regression functions $g_1(t) = g_2(t) = t^2$, standard normal distributed errors and different sample sizes $n_1, n_2 = 10, 20, 30, 50$. The results are summarized in Table 1, which shows the simulated rejection probabilities of the wild bootstrap test with level 10%, 5% and 2.5%. Table 2 shows the corresponding results for the regression functions $g_1(t) = g_2(t) = \cos(\pi t)$. We observe a reasonable approximation of the level by the wild bootstrap procedure in all cases, even in the case of very small samples [see also Hall and Hart (1990), who obtained a similar conclusion for their resampling procedure]. Note that for the more oscillating regression functions $g_i(t) = \cos(\pi t)$ the approximation is slightly worse compared to the more smooth case $g_1(t) = g_2(t) = t^2$, which can be partially explained by a larger bias in the variance estimators $\hat{\sigma}^2$, $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$.

TABLE 1

Simulated level of the test (3.3) *for various sample sizes and standard normal errors; the designs are uniform* [*according to* (3.4)] *and* $g_1(t) = g_2(t) = t^2$

(n_1, n_2)	(10, 10)	(10, 20)	(10, 30)	(10, 50)	(20, 20)
$\alpha = 10\%$	0.099	0.096	0.099	0.105	0.101
$\alpha = 5\%$	0.061	0.051	0.051	0.054	0.054
$\alpha = 2.5\%$	0.032	0.030	0.026	0.023	0.025
(n_1, n_2)	(20, 30)	(20, 50)	(30, 30)	(30, 50)	(50, 50)
$\alpha = 10\%$	0.098	0.108	0.099	0.090	0.108
$\alpha = 5\%$	0.054	0.050	0.048	0.047	0.048
$\alpha = 2.5\%$	0.029	0.028	0.025	0.026	0.025

TABLE 2

Simulated level of the test (3.3) *for various sample sizes and standard normal errors*; *the designs are uniform* [*according to* (3.4)] *and* $g_1(t) = g_2(t) = \cos(\pi t)$

(n_1, n_2)	(10, 10)	(10, 20)	(10, 30)	(10, 50)	(20, 20)
$\alpha = 10\%$	0.098	0.114	0.107	0.092	0.097
$\alpha = 5\%$	0.054	0.056	0.055	0.052	0.053
$\alpha = 2.5\%$	0.032	0.030	0.028	0.028	0.031
(n_1, n_2)	(20, 30)	(20, 50)	(30, 30)	(30, 50)	(50, 50)
$\alpha = 10\%$	0.100	0.096	0.098	0.095	0.101
$\alpha = 5\%$	0.053	0.048	0.050	0.051	0.052
$\alpha = 2.5\%$	0.023	0.026	0.031	0.028	0.027

As pointed out by a referee, it might be of interest to investigate the approximation of the level under a heteroscedastic error distribution. To this end, we considered the quadratic regression functions $g_1(t) = g_2(t) = t^2$ and the variance functions

$$\sigma_1^2(t) = \sigma_2^2(t) = \frac{e^t}{\int_0^1 e^x\,dx}, \tag{3.8}$$

$$\sigma_1^2(t) = \frac{e^t}{\int_0^1 e^x\,dx}, \qquad \sigma_2^2(t) = \frac{e^{2t}}{\int_0^1 e^{2x}\,dx}, \tag{3.9}$$

where the first and second scenarios correspond to the case of equal and unequal variance functions, respectively, and we normalized such that $\int_0^1 \sigma_i^2(t)\,dt = 1$, $i = 1, 2$. The results are listed in Tables 3 and 4 and demonstrate an excellent performance of the wild bootstrap procedure under heteroscedasticity.

3.2. *The test of Kulasekera and Wang* (1997). Kulasekera (1995) proposed a new testing procedure for the hypothesis (1.2) in the case of two samples

TABLE 3

Simulated level of the test (3.3) *for various sample sizes and standard normal but heteroscedastic errors*, *the designs are uniform* [*according to* (3.4)], $g_1(t) = g_2(t) = t^2$ *and the variance functions given by* (3.8)

(n_1, n_2)	(10, 10)	(10, 20)	(10, 30)	(10, 50)	(20, 20)
$\alpha = 10\%$	0.100	0.088	0.094	0.092	0.101
$\alpha = 5\%$	0.057	0.046	0.048	0.059	0.049
$\alpha = 2.5\%$	0.032	0.026	0.022	0.020	0.024
(n_1, n_2)	(20, 30)	(20, 50)	(30, 30)	(30, 50)	(50, 50)
$\alpha = 10\%$	0.088	0.093	0.095	0.092	0.106
$\alpha = 5\%$	0.046	0.047	0.055	0.047	0.048
$\alpha = 2.5\%$	0.023	0.020	0.031	0.021	0.028

TABLE 4

Simulated level of the test (3.3) *for various sample sizes and standard normal but heteroscedastic errors, the designs are uniform* [*according to* (3.4)], $g_1(t) = g_2(t) = t^2$ *and the variance functions given by* (3.9)

$(\boldsymbol{n_1}, \boldsymbol{n_2})$	**(10, 10)**	**(10, 20)**	**(10, 30)**	**(10, 50)**	**(20, 20)**
$\alpha = 10\%$	0.097	0.084	0.087	0.084	0.105
$\alpha = 5\%$	0.046	0.050	0.043	0.041	0.052
$\alpha = 2.5\%$	0.035	0.028	0.019	0.017	0.029
$(\boldsymbol{n_1}, \boldsymbol{n_2})$	**(20, 30)**	**(20, 50)**	**(30, 30)**	**(30, 50)**	**(50, 50)**
$\alpha = 10\%$	0.089	0.086	0.095	0.091	0.103
$\alpha = 5\%$	0.051	0.044	0.050	0.047	0.044
$\alpha = 2.5\%$	0.026	0.021	0.033	0.020	0.030

with homoscedastic errors. Because this test is applicable to different designs in both groups and can detect alternatives converging to the null at a rate $1/\sqrt{n}$, we will discuss it in a little more detail. The test is based on the quasi residuals

$$e_{1i} = Y_{1i} - \hat{g}_2(t_{1i}), \qquad i = 1, \dots, n_1,$$
$$e_{2j} = Y_{2j} - \hat{g}_1(t_{2j}), \qquad j = 1, \dots, n_2,$$

and the corresponding partial sums

$$\mu_i(t) = \sum_{j=1}^{\lfloor n_i t \rfloor} \frac{e_{ij}}{\sqrt{n_i}}, \qquad 0 < t < 1,\ i = 1, 2.$$

The test statistic proposed by Kulasekera (1995) is defined as a suitable function of

$$K_1^{(i)} = \frac{1}{n_i S_{n_i}^2} \sum_{k=1}^{n_i} \mu_i^2\left(\frac{k}{n_i}\right), \qquad i = 1, 2,$$

or

$$K_2^{(i)} = \frac{1}{S_{n_i}^3} \int_0^1 \mu_i^2(t-)\, d\mu_i(t), \qquad i = 1, 2,$$

where $S_{n_i}^2$ is a consistent estimator of $\sigma_i^2, i = 1, 2$. Note that this test does not require equal designs in both groups. Kulasekera and Wang (1997) investigated the functions $W_1 = \min\{K_1^{(1)}, K_1^{(2)}\}$, $W_2 = \min\{|K_2^{(1)}|, |K_2^{(2)}|\}$ and proposed a method for choosing the bandwidth, which, roughly speaking, maximizes the power at a specific alternative. As pointed out in the latter paper, the data-based smoothing parameters inflate the size of the test and the discrepancy from the actual size depends largely on the variability of the responses and the sample size. For these reasons, Kulasekera and Wang (1997) used simulation (for $g_1 = g_2 = 0$) to find the critical points.

TABLE 5

Simulated rejection probabilities of the test (3.3) *for various alternatives given in* (3.10); *the designs are uniform* [*according to* (3.4)] *and the errors are normal with variance* $\sigma^2 = 0.5$

	$n_1 = n_2 = 25$			$n_1 = n_2 = 50$		
Model	$\alpha = 10\%$	$\alpha = 5\%$	$\alpha = 2.5\%$	$\alpha = 10\%$	$\alpha = 5\%$	$\alpha = 2.5\%$
(a)	0.824	0.736	0.621	0.985	0.971	0.952
(b)	0.814	0.738	0.653	0.979	0.964	0.936
(c)	0.764	0.648	0.553	0.966	0.929	0.871
(d)	0.987	0.973	0.952	1.000	1.000	1.000
(e)	0.598	0.505	0.406	0.912	0.865	0.805
(f)	0.983	0.973	0.943	1.000	1.000	0.998

In Table 5 we compare the test (3.3) with the procedure proposed by Kulasekera and Wang (1997). For the sake of comparison, we chose the setup considered in Table 3 of the latter paper, that is, normally distributed errors with variance $\sigma^2 = 0.5$ and the regression functions

$$\begin{aligned} &\text{(a)} \quad g_1(x) = -g_2(x) = 0.5\cos(2\pi x),\\ &\text{(b)} \quad g_1(x) = -g_2(x) = 0.5\sin(2\pi x),\\ &\text{(c)} \quad g_1(x) = g_2(x) - x = \cos(\pi x),\\ &\text{(d)} \quad g_1(x) = g_2(x) - 1 = \cos(\pi x),\\ &\text{(e)} \quad g_1(x) = g_2(x) - x = \cos(2\pi x),\\ &\text{(f)} \quad g_1(x) = g_2(x) - 1 = \cos(2\pi x). \end{aligned} \tag{3.10}$$

Comparing the results of Table 5 with the corresponding results of Kulasekera and Wang (1997) in Table 3 of their paper, we observe that the test proposed in this paper yields a substantial improvement with respect to the power in all cases considered. Note that Kulasekera and Wang (1997) chose the bandwidths such that the power is maximized (at the cost of a simulated level) and we could obtain a further improvement in power for the test (3.3) by applying a similar technique. Although this would have theoretical advantages, we do not recommend this approach in practice, because this data-based choice of the smoothing parameter usually yields a large discrepancy between the size of the test and the actual level.

We note that the test (3.3) can only detect alternatives converging to the null at a rate $(N\sqrt{h})^{-1/2}$ [which gives $N^{-17/40}$ for the choice (3.7)], while Kulasekera and Wang's (1997) test achieves the parametric rate $N^{-1/2}$. At a first glance, this is a contradiction to the results obtained in our simulation. However, these observations can be explained by the fact that the method of Kulasekera and Wang implicitly uses a sample splitting. One sample is used for estimating the regression, while the other sample is used for the calculation of the residuals.

TABLE 6

Simulated rejection probabilities for the test of Young and Bowman (1995) *for various alternatives given in* (3.10), *the designs are uniform* [*according to* (3.4)] *and the errors normal with variance* $\sigma^2 = 0.5$

	$n_1 = n_2 = 25$			$n_1 = n_2 = 50$		
Model	$\alpha = 10\%$	$\alpha = 5\%$	$\alpha = 2.5\%$	$\alpha = 10\%$	$\alpha = 5\%$	$\alpha = 2.5\%$
(a)	0.388	0.236	0.136	0.897	0.801	0.660
(b)	0.772	0.651	0.547	0.990	0.975	0.956
(c)	0.789	0.683	0.592	0.973	0.947	0.902
(d)	0.998	0.996	0.991	1.000	1.000	1.000
(e)	0.833	0.711	0.620	0.997	0.996	0.989
(f)	0.983	0.969	0.940	1.000	1.000	0.998

For the sake of comparison, we also studied the performance of the test of Young and Bowman (1995) in this situation. The results are listed in Table 6. We observe a similar power of both tests in most cases.

3.3. *The tests of Delgado* (1993) *and Dette and Munk* (1998). The test proposed by Dette and Munk (1998) was the first procedure which was applicable in the general model (1.1). This test is based on a simple estimate of an L^2-distance between the regression functions which does not depend on a smoothing parameter. Although this procedure can only detect alternatives which converge to the null at a rate of $N^{-1/4}$, the test has promising finite sample properties with respect to the quality of approximation of the level [see Dette and Munk (1998)]. Moreover, a comparison with Delgado's (1993) test, which can detect alternatives converging to the null at a rate $N^{-1/2}$, indicates that for realistic sample sizes this test is comparable with procedures which are efficient from an asymptotic point of view. Delgado's (1993) test requires equal design points and is based on the sup-norm of a marked empirical process of the pairwise differences from both samples.

To compare the new test (3.3) with these procedures, we considered the setup given in Section 4.2 of Dette and Munk (1998), that is, $n_1 = n_2 = 15, 30, (g_1 - g_2)(t) \equiv 1, , (g_1 - g_2)(t) = \sin(2\pi t)$ and three types of error distributions [see also Hall and Hart (1990), pages 1041–1042]

$$\begin{aligned} &\text{(i)} && (\mathcal{N}_1, \mathcal{N}_2),\\ &\text{(ii)} && (|\mathcal{N}_1| - \sqrt{2/\pi}, |\mathcal{N}_2| - \sqrt{2/\pi}),\\ &\text{(iii)} && (|\mathcal{N}_1| - \sqrt{2/\pi}, \sqrt{2/\pi} - |\mathcal{N}_2|). \end{aligned} \tag{3.11}$$

The results are listed in Table 7 and a comparison of the power at the 5% level shows the following. While Delgado's (1993) test performs better for the smooth alternative $g_1 - g_2 \equiv 1$, Dette and Munk's (1998) test is more efficient for the oscillating alternative. The new test (3.3) has a reasonable performance in both cases. On the one hand, it is substantially more powerful than

TABLE 7

Simulated rejection probabilities of the test (3.3) *in the scenario considered by Dette and Munk* (1998) *Section* 4.2; *the design is uniform* [*according to* (3.4)] *and the error distributions are given by* (3.11)

$(g_1 - g_2)(t)$		1			$\sin(2\pi t)$		
$(\varepsilon_{1i}, \varepsilon_{2i})$		**(i)**	**(ii)**	**(iii)**	**(i)**	**(ii)**	**(iii)**
$n_i = 15$	2.5%	0.552	0.936	0.875	0.176	0.517	0.500
	5%	0.648	0.963	0.913	0.255	0.615	0.604
	10%	0.734	0.977	0.941	0.347	0.703	0.713
$n_i = 30$	2.5%	0.870	0.999	0.992	0.386	0.863	0.772
	5%	0.925	0.999	0.998	0.492	0.912	0.917
	10%	0.954	1.000	1.000	0.608	0.951	0.950

Delgado's test for the oscillating alternative and Dette and Munk's test for the smooth alternative. On the other hand, it is comparable with these procedures in the remaining cases.

Our final example compares the new test with the bootstrap test introduced by Hall and Hart (1990). These authors mainly considered the case of equal design points and briefly mentioned a generalization of their approach to the general case. However, Kulasekera (1995) observed that this generalization is not reliable and recommends the application of Hall and Hart's test only in the case of equal designs. Note that this test can detect alternatives converging to the null at a rate $N^{-1/2}$. For a comparison with our test, we chose the setup of Table 3 in Hall and Hart (1990). The test proposed by these authors depends on a smoothing parameter p and Table 3 in Hall and Hart (1990) lists results for three choices of p. More precisely, the errors are given by (3.11) and the alternatives by $g_1 - g_2 = 1$ and $(g_1 - g_2)(x) = x$, where $g_2 = 0$. The results are given in Table 8 and show that the new test is a serious competitor. In most cases, we observe a better power for the new test (3.3), even if we compare it with the best choice of p in the procedure of Hall and Hart.

4. Proofs.

4.1. *Preliminaries.* We will restrict ourselves to a proof of Theorem 2.1 in the case of $k = 2$ regression functions. The general case $k \geq 3$ and the asymptotic results given in Theorems 2.2 and 2.3 for $T_N^{(2)}$ and $T_N^{(3)}$ follow by exactly the same arguments with an additional amount of algebra and notation. For the sake of a transparent notation, we will omit all indices referring to the number of samples and to the specific statistic discussed in Section 2. In other words, we write B instead of $B_k^{(1)}$, T_N instead of $T_N^{(1)}$ and so on. Recalling the definition of the weights

$$w_{jk}^{(i)} = \frac{K((t_{ij} - t_{ik})/h)}{\sum_{l=1}^{n_i} K((t_{ij} - t_{il})/h)}, \qquad i = 1, 2, \tag{4.1}$$

TABLE 8

Simulated rejection probabilities of the test (3.3) *in the scenario considered by Hall and Hart* (1990), *Table* 3; *the design is uniform* [*according to* (3.4)] *and the error distributions are given by* (3.11)

$(g_1 - g_2)(t)$		1			t		
$(\varepsilon_{1i}, \varepsilon_{2i})$		(i)	(ii)	(iii)	(i)	(ii)	(iii)
$n_i = 15$	2.5%	0.550	0.928	0.874	0.197	0.493	0.493
	5%	0.643	0.957	0.931	0.276	0.612	0.591
	10%	0.738	0.982	0.960	0.384	0.726	0.712
$n_i = 20$	2.5%	0.666	0.989	0.966	0.237	0.624	0.591
	5%	0.752	0.992	0.982	0.335	0.724	0.685
	10%	0.844	0.999	0.990	0.447	0.817	0.779
$n_i = 30$	2.5%	0.848	1.000	0.995	0.369	0.762	0.782
	5%	0.898	1.000	0.998	0.487	0.847	0.832
	10%	0.939	1.000	0.999	0.584	0.906	0.899
$n_i = 50$	2.5%	0.990	1.000	1.000	0.541	0.941	0.940
	5%	0.995	1.000	1.000	0.649	0.969	0.962
	10%	0.999	1.000	1.000	0.741	0.983	0.981

the Nadaraya–Watson estimators (2.4) of the individual regression functions evaluated at the points t_{ij} can be rewritten as

$$\hat{g}_i(t_{ij}) = \sum_{k=1}^{n_i} w_{jk}^{(i)} Y_{ik}, \qquad i = 1, 2.$$

To obtain a similar representation for the estimator in the combined sample, define the weights

$$w_{lk,\,ij} = \frac{K((t_{lk} - t_{ij})/h)}{\sum_{l'=1}^{2} \sum_{k'=1}^{n_{l'}} K((t_{l'k'} - t_{ij})/h)}. \tag{4.2}$$

The Nadaraya–Watson estimator (2.6) evaluated at the points t_{ij} using the total sample can now be written as

$$\hat{g}(t_{ij}) = \sum_{l=1}^{2} \sum_{k=1}^{n_l} w_{lk,\,ij} Y_{lk}. \tag{4.3}$$

We finally introduce the notation $\lambda_1 = \kappa_1/\kappa_2$ and

$$\rho(t) = r_1(t) + \frac{1}{\lambda_1} r_2(t) = \frac{1}{\kappa_1}\big[\kappa_1 r_1(t) + \kappa_2 r_2(t)\big], \tag{4.4}$$

which will be used frequently throughout this section. Our first result establishes the asymptotic expansion for the bias of the estimator $T_N^{(1)}$.

LEMMA 4.0. *Assume that* (2.1), (2.2), (2.3), (2.5) *and* (2.7) *are satisfied. Then*

$$
\begin{aligned}
E[\hat{\sigma}_i^2] &= \int_0^1 \sigma_i^2(t) r_i(t)\,dt + d_i h^{2r} + o(h^{2r}) + O\left(\frac{1}{n_i}\right) \\
&\quad + \frac{1}{n_i h}\left[\int_{-1}^{1} K^2(u)\,du - 2K(0)\right]\int_0^1 \sigma_i^2(t)\,dt,
\end{aligned} \tag{4.5}
$$

where the constant d_i is defined by

$$
d_i = k_r^2 \int_0^1 \left[(g_i r_i)^{(r)}(t) - g_i r_i^{(r)}(t)\right]^2 \frac{1}{r_i(t)}\,dt, \qquad i = 1, \dots, k.
$$

Moreover, if the null hypothesis of equal regression functions is valid, we have, for the estimator (2.9) *in the case* $k = 2$,

$$
\begin{aligned}
E[\hat{\sigma}^2] &= \kappa_1 \int_0^1 \sigma_1^2(t) r_1(t)\,dt + \kappa_2 \int_0^1 \sigma_2^2(t) r_2(t)\,dt \\
&\quad + \frac{1}{Nh}\left[\int_{-1}^{1} K^2(u)\,du - 2K(0)\right]\int_0^1 \frac{\sigma_1^2(t)\kappa_1 r_1(t) + \sigma_2^2(t)\kappa_2 r_2(t)}{(\kappa_1 r_1 + \kappa_2 r_2)(t)}\,dt \\
&\quad + C h^{2r} + o(h^{2r}) + O\left(\frac{1}{N}\right),
\end{aligned} \tag{4.6}
$$

where the constant C is defined by

$$
C = k_r^2 \int_0^1 \left[\left(g_1(\kappa_1 r_1 + \kappa_2 r_2)\right)^{(r)} - g_1(\kappa_1 r_1 + \kappa_2 r_2)^{(r)}\right]^2(t) \frac{dt}{(\kappa_1 r_1 + \kappa_2 r_2)(t)}.
$$

Under the alternative, we obtain, for the estimator (2.9) *in the case* $k = 2$,

$$
\begin{aligned}
E[\hat{\sigma}^2] &= \kappa_1 \int_0^1 \sigma_1^2(t) r_1(t)\,dt + \kappa_2 \int_0^1 \sigma_2^2(t) r_2(t)\,dt + M^2 \\
&\quad + O(h^{2r}) + O\left(\frac{1}{Nh}\right),
\end{aligned} \tag{4.7}
$$

where the constant M^2 is defined by

$$
M^2 = \kappa_1 \kappa_2 \int_0^1 \left(g_1(t) - g_2(t)\right)^2 \frac{r_1(t) r_2(t)}{\kappa_1 r_1(t) + \kappa_2 r_2(t)}\,dt. \tag{4.8}
$$

PROOF. The first part (4.5) of the lemma is obtained from the representation (4.6) by considering equal variance functions and design densities. The proof of (4.6) and (4.7) essentially follows the arguments of Hall and Marron (1990), and we will only mention the main modifications here, which take into account the mixture of two design densities. Define

$$
\Delta_{ij} = g_i(t_{ij}) - \sum_{l=1}^{2} \sum_{k=1}^{n_l} w_{lk,\,ij}\, g_l(t_{lk}). \tag{4.9}
$$

Then the expectation of the variance estimator (2.9) from the total sample splits into two parts, that is,

$$E[\hat\sigma^2]=\frac{1}{N}\sum_{i=1}^{2}\sum_{j=1}^{n_i}\Delta_{ij}^2+\frac{1}{N}\sum_{i=1}^{2}\sum_{j=1}^{n_i}E\left[\left(\sigma_i(t_{ij})\varepsilon_{ij}-\sum_{l=1}^{2}\sum_{k=1}^{n_l}w_{lk,ij}\sigma_l(t_{lk})\varepsilon_{lk}\right)^2\right]. \tag{4.10}$$

A straightforward but tedious calculation shows

$$\begin{aligned}
\Delta_{ij} &= \left[\frac{1}{\rho(t_{ij})}+R(t_{ij})\right]\\
&\quad\times\left[\frac{\kappa_2}{\kappa_i}r_{3-i}(t_{ij})(g_i-g_{3-i})(t_{ij})+\frac{1}{h}\int K\left(\frac{t_{ij}-t}{h}\right)\right.\\
&\quad\times\left((g_1r_1)(t_{ij})-(g_1r_1)(t)+g_i(t_{ij})\{\rho(t)-\rho(t_{ij})\}\right.\\
&\quad\left.\left.+\left(g_2\tfrac{1}{\lambda_1}r_2\right)(t_{ij})-\left(g_2\frac{1}{\lambda_1}r_2\right)(t)\right)dt\right]+O\left(\frac{1}{Nh}\right)+O(h^{2r})
\end{aligned}\tag{4.11}$$

$$= \frac{(\kappa_2/\kappa_i)r_{3-i}(t_{ij})(g_i-g_{3-i})(t_{ij})}{\rho(t_{ij})}+O(h^r)+O\left(\frac{1}{Nh}\right), \tag{4.12}$$

uniformly in $j=1,\dots,n_i$, $i=1,2$. Here the function R is defined by

$$R(x)=\frac{\rho(x)-(1/h)\int K((x-t)/h)\rho(t)\,dt}{\rho^2(x)}=O(h^r), \tag{4.13}$$

where the estimate on the right-hand side follows from the differentiability of the design densities and the moment assumptions on the kernel.

Now the evaluation of the first term in (4.10) gives, for $g_1\neq g_2$,

$$\begin{aligned}
\frac{1}{N}\sum_{i=1}^{2}\sum_{j=1}^{n_i}\Delta_{ij}^2 &= \kappa_1\int_0^1\frac{((1/\lambda_1)r_2)^2(x)r_1(x)}{\rho^2(x)}(g_1-g_2)^2(x)\,dx\\
&\quad+\kappa_2\int_0^1\frac{r_2(x)r_1^2(x)}{\rho^2(x)}(g_1-g_2)^2(x)\,dx\\
&\quad+2\kappa_1\int_0^1\frac{(1/\lambda_1)r_2(x)r_1(x)}{\rho^2(x)}(g_1-g_2)(x)\\
&\quad\times\left(\frac{1}{h}\int K\left(\frac{x-t}{h}\right)\Big\{(g_1r_1)(x)-(g_1r_1)(t)\right.\\
&\qquad\left.+g_1(x)\big(\rho(t)-\rho(x)\big)+\left(g_2\frac{1}{\lambda_1}r_2\right)(x)-\left(g_2\frac{1}{\lambda_1}r_2\right)(t)\Big\}\,dt\right)dx\\
&\quad+2\kappa_1\int_0^1\frac{R(x)}{\rho(x)}\left(\frac{1}{\lambda_1}r_2(x)(g_1-g_2)(x)\right)^2r_1(x)\,dx\\
&\quad+2\kappa_2\int_0^1\frac{r_2(x)r_1(x)}{\rho^2(x)}(g_2-g_1)(x)
\end{aligned}$$

$$\times\left(\frac{1}{h}\int K\left(\frac{x-t}{h}\right)\Big\{(g_1r_1)(x)-(g_1r_1)(t)$$

$$+g_2(x)\big(\rho(t)-\rho(x)\big)+\left(g_2\frac{1}{\lambda_1}r_2\right)(x)-\left(g_2\frac{1}{\lambda_1}r_2\right)(t)\Big\}\,dt\right)dx$$

$$+2\kappa_2\int_0^1\frac{R(x)}{\rho(x)}\big(r_1(x)(g_2-g_1)(x)\big)^2r_2(x)\,dx+O(h^{2r})+O\left(\frac{1}{Nh}\right)$$

$$(4.14)\qquad =\kappa_1\int_0^1\frac{(1/\lambda_1)r_2(x)r_1(x)}{\rho(x)}(g_1-g_2)^2(x)\,dx$$

$$+2\kappa_1\int_0^1\frac{(1/\lambda_1)r_2(x)r_1(x)}{\rho^2(x)}(g_1-g_2)(x)$$

$$\times\left(\frac{1}{h}\int K\left(\frac{x-t}{h}\right)\Big\{(g_1r_1)(x)-(g_1r_1)(t)\right.$$

$$+g_1(x)\big(\rho(t)-\rho(x)\big)+\left(g_2\frac{1}{\lambda_1}r_2\right)(x)$$

$$-\left(g_2\frac{1}{\lambda_1}r_2\right)(t)-(g_1r_1)(x)+(g_1r_1)(t)$$

$$\left.-g_2(x)\big(\rho(t)-\rho(x)\big)-\left(g_2\frac{1}{\lambda_1}r_2\right)(x)+\left(g_2\frac{1}{\lambda_1}r_2\right)(t)\Big\}\,dt\right)dx$$

$$+2\kappa_1\int_0^1R(x)(g_1-g_2)^2(x)r_1(x)\frac{1}{\lambda_1}r_2(x)\,dx+O(h^{2r})+O\left(\frac{1}{Nh}\right)$$

$$=\kappa_1\int_0^1\frac{(1/\lambda_1)r_2(x)r_1(x)}{\rho(x)}(g_1-g_2)^2(x)\,dx$$

$$+2\kappa_1\int_0^1\frac{(1/\lambda_1)r_2(x)r_1(x)}{\rho^2(x)}(g_1-g_2)(x)$$

$$\times\left(\frac{1}{h}\int K\left(\frac{x-t}{h}\right)\big(g_1(x)-g_2(x)\big)\big(\rho(t)-\rho(x)\big)\,dt\right)dx$$

$$+2\kappa_1\int_0^1R(x)(g_1-g_2)^2(x)r_1(x)\frac{1}{\lambda_1}r_2(x)\,dx+O(h^{2r})+O\left(\frac{1}{Nh}\right)$$

$$=\kappa_1\int_0^1\frac{(1/\lambda_1)r_2(x)r_1(x)}{\rho(x)}(g_1-g_2)^2(x)\,dx+O(h^{2r})+O\left(\frac{1}{Nh}\right)$$

$$=M^2+O(h^{2r})+O\left(\frac{1}{Nh}\right),$$

where we used the definition of $1/\lambda_1=\kappa_2/\kappa_1$ in the first equality and the definition of $R(x)$ in (4.13) and the definition of M^2 in (4.8) for the last step. Under the assumption of equal regression curves $g_1=g_2$, (4.11) simplifies and we obtain, observing (2.7) and (4.4),

$$
\begin{aligned}
&\frac{1}{N}\sum_{i=1}^{2}\sum_{j=1}^{n_i}\Delta_{ij}^2 \\
&\quad =\kappa_1\int_0^1\left[\frac{1}{\rho(x)}+R(x)\right]^2 \\
&\qquad \times\left(\frac{1}{h}\int K\left(\frac{x-t}{h}\right)\Big\{(g_1\rho)(x)-(g_1\rho)(t)+g_1(x)\big(\rho(t)-\rho(x)\big)\Big\}\,dt \\
&\qquad +O\left(\frac{1}{Nh}\right)\right)^2 r_1(x)\,dx+\kappa_1\int_0^1\left[\frac{1}{\rho(x)}+R(x)\right]^2 \\
&\qquad \times\left(\frac{1}{h}\int K\left(\frac{x-t}{h}\right)\Big\{(g_1\rho)(x)-(g_1\rho)(t) \\
&\qquad +g_1(x)\big(\rho(t)-\rho(x)\big)\Big\}\,dt+O\left(\frac{1}{Nh}\right)\right)^2\frac{1}{\lambda_1}r_2(x)\,dx+O\left(\frac{1}{N}\right) \qquad (4.15)\\
&\quad =\kappa_1\int_0^1\left[\frac{1}{\rho(x)}+R(x)\right]^2 \\
&\qquad \times\rho(x)\left\{h^r k_r\left[g_1\rho^{(r)}-(g_1\rho)^{(r)}\right](x)+o(h^r)+O\left(\frac{1}{Nh}\right)\right\}^2 dx \\
&\qquad +O\left(\frac{1}{N}\right) \\
&\quad =h^{2r}k_r^2\kappa_1\int_0^1\frac{1}{\rho(x)}\left[g_1\rho^{(r)}-(g_1\rho)^{(r)}\right]^2(x)\,dx+o(h^{2r})+O\left(\frac{1}{N}\right) \\
&\qquad \times h^{2r}C+o(h^{2r})+O\left(\frac{1}{N}\right).
\end{aligned}
$$

For the second term in (4.10), we obtain by a straightforward but cumbersome calculation

$$
\begin{aligned}
U &= \frac{1}{N}\sum_{i=1}^{2}\sum_{j=1}^{n_i}E\left[\left(\sigma_i(t_{ij})\varepsilon_{ij}-\sum_{l=1}^{2}\sum_{k=1}^{n_l}w_{lk,\,ij}\sigma_l(t_{lk})\varepsilon_{lk}\right)^2\right] \\
&= \frac{1}{N}\sum_{i=1}^{2}\sum_{j=1}^{n_i}\sigma_i^2(t_{ij})-\frac{2}{N}\sum_{i=1}^{2}\sum_{j=1}^{n_i}\sigma_i^2(t_{ij})w_{ij,\,ij} \\
&\quad +\frac{1}{N}\sum_{i=1}^{2}\sum_{j=1}^{n_i}\sum_{l=1}^{2}\sum_{k=1}^{n_l}\sigma_l^2(t_{lk})w_{lk,\,ij}^2 \\
&= \sum_{i=1}^{2}\kappa_i\int_0^1\sigma_i^2(t)r_i(t)\,dt+\sum_{i=1}^{2}\frac{1}{Nh}
\end{aligned}
$$

$$
\begin{aligned}
(4.16)\quad & \times\biggl[-2K(0)\int_0^1 \sigma_i^2(t)\frac{(\kappa_i/\kappa_1)r_i(t)}{\rho(t)}\,dt \\
&\quad + \frac{1}{h}\int\int K^2\biggl(\frac{t-x}{h}\biggr)\frac{\sigma_i^2(x)(\kappa_i/\kappa_1)r_i(x)}{\rho(t)}\,dt\,dx\biggr] \\
&\quad + O\biggl(\frac{1}{N}\biggr) + O\biggl(\frac{h^r}{Nh}\biggr) + O\biggl(\frac{1}{(Nh)^2}\biggr) \\
&= \kappa_1\int_0^1 \sigma_1^2(t)r_1(t)\,dt + \kappa_2\int_0^1 \sigma_2^2(t)r_2(t)\,dt + \frac{1}{Nh}\biggl[\int K^2(u)\,du - 2K(0)\biggr] \\
&\quad \times\biggl\{\int_0^1 \frac{\sigma_1^2(x)r_1(x)}{\rho(x)}\,dx + \int_0^1 \frac{\sigma_2^2(x)(1/\lambda_1)r_2(x)}{\rho(x)}\,dx\biggr\} + O\biggl(\frac{1}{N}\biggr),
\end{aligned}
$$

and the assertions (4.6) and (4.7) follow from (4.10), (4.14), (4.15) and (4.16). □

4.2. *Proof of Theorem* 2.1: *the null hypothesis of equal regression functions.* In a first step we introduce the notation [observing (4.1)]

$$
(4.17)\qquad \delta_{ij} = g_i(t_{ij}) - \sum_{k=1}^{n_i} w_{jk}^{(i)} g_i(t_{ik}), \qquad j = 1, \dots, n_i,\ i = 1, 2,
$$

and decompose the centered version of T_N as

$$
(4.18)\qquad T_N - E[T_N] = \hat{\sigma}^2 - \frac{n_1}{N}\hat{\sigma}_1^2 - \frac{n_2}{N}\hat{\sigma}_2 - E[T_N] = R_{1,N} + R_{2,N},
$$

where

$$
\begin{aligned}
R_{1,N} &= \frac{2}{N}\sum_{i=1}^{2}\sum_{j=1}^{n_i}\Delta_{ij}\biggl(\sigma_i(t_{ij})\varepsilon_{ij} - \sum_{l=1}^{2}\sum_{k=1}^{n_l} w_{lk,ij}\sigma_l(t_{lk})\varepsilon_{lk}\biggr) \\
(4.19)\qquad &\quad - \frac{2}{N}\sum_{i=1}^{2}\sum_{j=1}^{n_i}\delta_{ij}\biggl(\sigma_i(t_{ij})\varepsilon_{ij} - \sum_{k=1}^{n_i} w_{jk}^{(i)}\sigma_i(t_{ik})\varepsilon_{ik}\biggr) \\
&= \widehat{T}_N^{(1)} + \widehat{T}_N^{(2)},
\end{aligned}
$$

$$
\begin{aligned}
R_{2,N} &= \frac{1}{N}\sum_{i=1}^{2}\sum_{j=1}^{n_i}\biggl(\sigma_i(t_{ij})\varepsilon_{ij} - \sum_{l=1}^{2}\sum_{k=1}^{n_l} w_{lk,ij}\sigma_l(t_{lk})\varepsilon_{lk}\biggr)^2 \\
&\quad - \frac{1}{N}\sum_{i=1}^{2}\sum_{j=1}^{n_i} E\biggl[\biggl(\sigma_i(t_{ij})\varepsilon_{ij} - \sum_{l=1}^{2}\sum_{k=1}^{n_l} w_{lk,ij}\sigma_l(t_{lk})\varepsilon_{lk}\biggr)^2\biggr] \\
(4.20)\qquad &\quad - \frac{1}{N}\sum_{i=1}^{2}\sum_{j=1}^{n_i}\biggl(\sigma_i(t_{ij})\varepsilon_{ij} - \sum_{k=1}^{n_i} w_{jk}^{(i)}\sigma_i(t_{ik})\varepsilon_{ik}\biggr)^2 \\
&\quad + \frac{1}{N}\sum_{i=1}^{2}\sum_{j=1}^{n_i} E\biggl[\biggl(\sigma_i(t_{ij})\varepsilon_{ij} - \sum_{k=1}^{n_i} w_{jk}^{(i)}\sigma_i(t_{ik})\varepsilon_{ik}\biggr)^2\biggr] = \sum_{j=3}^{7}\widehat{T}_N^{(j)}.
\end{aligned}
$$

Here the random variables $\widehat{T}_N^{(j)}$ are defined by

$$\widehat{T}_N^{(i)} = \frac{1}{N}\sum_{j=1}^{n_i} \alpha_{ij}\varepsilon_{ij}, \qquad i=1,2, \tag{4.21}$$

$$\widehat{T}_N^{(2+i)} = \frac{1}{N}\sum_{j=1}^{n_i} k_{ij}(\varepsilon_{ij}^2-1), \qquad i=1,2, \tag{4.22}$$

$$\widehat{T}_N^{(4+s)} = \frac{1}{N}\sum_{i=1}^{n_s}\sum_{\substack{l=1\\ l\neq i}}^{n_s} r_{il}^{(s)}\varepsilon_{si}\varepsilon_{sl}, \qquad s=1,2, \tag{4.23}$$

$$\widehat{T}_N^{(7)} = \frac{1}{N}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \bar{t}_{ij}\varepsilon_{1i}\varepsilon_{2j} \tag{4.24}$$

and the coefficients α_{ij} are given by

$$\alpha_{ij} = 2\biggl(\Delta_{ij} - \delta_{ij} - \sum_{l=1}^{2}\sum_{k=1}^{n_l}\Delta_{lk}w_{ij,lk} + \sum_{k=1}^{n_i}\delta_{ik}w_{ik}^{(i)}\biggr)\sigma_i(t_{ij}), \qquad j=1,\dots,n_i,\quad i=1,2, \tag{4.25}$$

where Δ_{ij} and δ_{ij} are defined in (4.9) and (4.17), respectively. The coefficients k_{ij}, $r_{il}^{(s)}$, $\bar{t}_{ij}$ in the representation of $R_{2,n}$ are defined as follows:

$$k_{ij} = \biggl(2w_{jj}^{(i)} - 2w_{ij,ij} + \sum_{l=1}^{2}\sum_{k=1}^{n_l} w_{ij,lk}^2 - \sum_{k=1}^{n_i}(w_{ik}^{(i)})^2\biggr)\sigma_i^2(t_{ij}), \qquad i=1,2, \tag{4.26}$$

$$r_{il}^{(s)} = \biggl(2w_{il}^{(s)} - 2w_{si,sl} + \sum_{j=1}^{2}\sum_{k=1}^{n_j} w_{jk,si}w_{jk,sl} - \sum_{k=1}^{n_s} w_{ki}^{(s)}w_{kl}^{(s)}\biggr)\sigma_s(t_{si})\sigma_s(t_{sl}), \qquad s=1,2, \tag{4.27}$$

$$\bar{t}_{ij} = \biggl(-2w_{1i,2j} - 2w_{2j,1i} + 2\sum_{l=1}^{2}\sum_{k=1}^{n_l} w_{lk,1i}w_{lk,2j}\biggr)\sigma_1(t_{1i})\sigma_2(t_{2j}). \tag{4.28}$$

The next lemmas specify the asymptotic behavior of the terms $\widehat{T}_N^{(j)}$ on the right-hand side of (4.19) and (4.20). Note that all terms in these representations are centered, that is, $E[\widehat{T}_N^{(j)}] = 0$, $j=1,\dots,7$.

LEMMA 4.1. *If the assumptions of Theorem* 2.1 *are satisfied*, *we have*, *under the hypothesis of equal regression curves*,

$$\widehat{T}_N^{(j)} \overset{H_0}{=} o_p\biggl(\frac{1}{N\sqrt{h}}\biggr), \qquad j=1,2,$$

and, under the alternative,

$$\operatorname{var}(\widehat{T}_N^{(i)}) \overset{H_1}{=} \frac{4}{N}\frac{(\kappa_1\kappa_2)^2}{\kappa_i}\int_0^1 (g_1-g_2)^2(t)\sigma_i^2(t)\frac{r_1^2(t)r_2^2(t)}{r_i(t)\big(\kappa_1 r_1(t)+\kappa_2 r_2(t)\big)^2}\,dt + o\left(\frac{1}{N}\right), \qquad i=1,2.$$

PROOF. We only prove the assertion for the statistic $\widehat{T}_N^{(1)}$; the remaining case follows by exactly the same arguments. From (4.21) it follows that

$$\operatorname{var}(\widehat{T}_N^{(1)}) = \frac{1}{N^2}\sum_{i=1}^{n_1}\alpha_{1i}^2, \tag{4.29}$$

where, by (4.25) and (4.12),

$$\begin{aligned}\alpha_{1i} &= 2\left\{\frac{(1/\lambda_1)r_2(t_{1i})}{\rho(t_{1i})}(g_1-g_2)(t_{1i})\right.\\ &\qquad \left.-\frac{1}{n_1 h}\sum_{l=1}^{2}\sum_{k=1}^{n_l}\frac{(\kappa_2/\kappa_l)r_{3-l}(t_{lk})}{\rho(t_{lk})}(g_1-g_2)(t_{lk})K\left(\frac{t_{lk}-t_{1i}}{h}\right)\frac{1}{\rho(t_{lk})}\right\}\\ &\qquad \times\sigma_1(t_{1i}) + O(h^r) + O\left(\frac{1}{Nh}\right)\\ &= \frac{2(1/\lambda_1)r_2(t_{1i})\sigma_1(t_{1i})}{\rho(t_{1i})}(g_1-g_2)(t_{1i}) + O(h^r) + O\left(\frac{1}{Nh}\right),\end{aligned} \tag{4.30}$$

uniformly with respect to $i = 1,\dots,n_1$. The last equality in (4.30) uses the fact that the integral approximations of the two sums have the same absolute value with opposite signs. Now (4.29) implies, under the hypothesis of equal regression curves,

$$\operatorname{var}(\widehat{T}_N^{(1)}) = o\left(\frac{1}{N^2 h}\right),$$

and an application of Chebyshev's inequality proves the first part of the lemma. For the second part we obtain, from (4.29) and (4.30),

$$\begin{aligned}\operatorname{var}(\widehat{T}_N^{(1)}) &= \frac{1}{N^2}\sum_{i=1}^{n_1}\left(\frac{2(1/\lambda_1)r_2(t_{1i})\sigma_1(t_{1i})}{\rho(t_{1i})}(g_1-g_2)(t_{1i})\right)^2\\ &\qquad + \frac{n_1}{N^2}\left(O(h^r) + O\left(\frac{1}{Nh}\right)\right)\\ &= \frac{4}{N}\kappa_1\int_0^1 (g_1-g_2)^2(t)\sigma_1^2(t)\frac{r_1(t)((1/\lambda_1)r_2)^2(t)}{\rho^2(t)}\,dt + o\left(\frac{1}{N}\right),\end{aligned}$$

which completes the proof, by the definition of ρ and λ_1. $\square$

LEMMA 4.2. *Under the assumptions of Theorem* 2.1, *we have*

$$\widehat{T}_N^{(j)} = o_p\left(\frac{1}{N\sqrt{h}}\right), \qquad j = 3, 4.$$

PROOF. Note that $E[\widehat{T}_N^{(j)}] = 0$ for $j = 3, 4$. Recalling the definition of the coefficients k_{1i} in (4.26), we obtain

$$\begin{aligned} k_{1i} = &\left\{\frac{2K(0)}{n_1 h_1}\frac{1}{r_1(t_{1i})} - \frac{2K(0)}{n_1 h}\frac{1}{\rho(t_{1i})}\right. \\ &\left. + \frac{1}{n_1 h^2}\int K^2\left(\frac{s - t_{1i}}{h}\right)\frac{1}{\rho(s)}\,ds - \frac{1}{n_1 h_1^2}\int K^2\left(\frac{t - t_{1i}}{h_1}\right)\frac{1}{r_1(t)}\,dt\right\} \\ &\times \sigma_1^2(t_{1i}) + o\left(\frac{1}{n_1 h}\right) = O\left(\frac{1}{Nh}\right), \end{aligned} \tag{4.31}$$

uniformly in $i = 1, \ldots, n_1$. This implies, for the variance of $\widehat{T}_N^{(3)}$,

$$\operatorname{var}(\widehat{T}_N^{(3)}) = \frac{1}{N^2}\sum_{i=1}^{n_1} k_{1i}^2 \operatorname{var}(\varepsilon_{1i}^2) = o\left(\frac{1}{N^2 h}\right)$$

and proves Lemma 4.2 for the case $j = 3$. The remaining case is obtained by exactly the same arguments and is therefore omitted. □

LEMMA 4.3. *Under the assumptions of Theorem* 2.1, *we have*

$$\begin{aligned} \operatorname{var}(\widehat{T}_N^{(4+i)}) = &\frac{2}{N^2 h}\int_0^1 \sigma_i^4(x)\left[1 - \frac{\kappa_i}{\kappa_1}\frac{r_i(x)}{\rho(x)}\right]^2 dx \\ &\times \int (2K - K * K)^2(u)\,du + o\left(\frac{1}{N^2 h}\right), \qquad i = 1, 2. \end{aligned}$$

PROOF. We only sketch a proof of the first part $i = 1$ of the assertion; the remaining case $i = 2$ follows by exactly the same arguments. Recalling the definition of the weights $r_{il}^{(s)}$ in (4.27), we obtain by straightforward algebra

$$\begin{aligned} r_{il}^{(1)} = &\left\{2K\left(\frac{t_{1i} - t_{1l}}{h}\right)\frac{1}{r_1(t_{1i})} - K\left(\frac{t_{1i} - t_{1l}}{h}\right)\frac{2}{\rho(t_{1i})}\right. \\ &+ \frac{1}{h}\int K\left(\frac{s - t_{1i}}{h}\right)K\left(\frac{s - t_{1l}}{h}\right)\frac{1}{\rho(s)}\,ds \\ &\left. - \frac{1}{h}\int K\left(\frac{t - t_{1i}}{h}\right)K\left(\frac{t - t_{1l}}{h}\right)\frac{1}{r_1(t)}\,dt\right\}\frac{\sigma_1(t_{1i})\sigma_1(t_{1l})}{n_1 h} + o\left(\frac{1}{N}\right), \end{aligned}$$

uniformly for $i, l = 1, \ldots, n_1$, and straightforward but tedious algebra shows

$$\begin{aligned}
\operatorname{var}(\widehat{T}_N^{(5)}) &= \frac{1}{N^2}\sum_{i=1}^{n_1}\sum_{\substack{l=1\\ l\neq i}}^{n_1}\Big[(r_{il}^{(1)})^2 + r_{il}^{(1)} r_{li}^{(1)}\Big] \\
&= \frac{2}{N^2}\sum_{i=1}^{n_1}\sum_{l=1}^{n_1}(r_{il}^{(1)})^2 + o\bigg(\frac{1}{N^2 h}\bigg) \\
&= \frac{2}{N^2 h}\Bigg\{4\int K^2(u)\,du\int_0^1 \sigma_1^4(x)\bigg[1-\frac{r_1(x)}{\rho(x)}\bigg]^2 dx \\
&\qquad + \int (K*K)^2(u)\,du \int_0^1 \sigma_1^4(x)\bigg[1-\frac{r_1(x)}{\rho(x)}\bigg]^2 dx \\
&\qquad - 4\int (K*K)(u)K(u)\,du \\
&\qquad \times \int_0^1 \sigma_1^4(x)\bigg[1-\frac{r_1(x)}{\rho(x)}\bigg]^2 dx\Bigg\} + o\bigg(\frac{1}{N^2 h}\bigg) \\
&= \frac{2}{N^2 h}\int_0^1 \sigma_1^4(x)\bigg[1-\frac{r_1(x)}{\rho(x)}\bigg]^2 \cdot \int (2K - K*K)^2(u)\,du + o\bigg(\frac{1}{N^2 h}\bigg). \qquad \square
\end{aligned}$$

LEMMA 4.4. *Under the assumptions of Theorem* 2.1, *we have*

$$\begin{aligned}
\operatorname{var}(\widehat{T}_N^{(7)}) = \frac{4}{N^2 h}&\int (2K - K*K)^2(u)\,du \\
&\times \int_0^1 \frac{r_1(x)(1/\lambda_1) r_2(x)}{\rho^2(x)}\sigma_1^2(x)\sigma_2^2(x)\,dx + o\bigg(\frac{1}{N^2 h}\bigg).
\end{aligned}$$

PROOF. A straightforward calculation shows, for the coefficients $\bar{t}_{ij}$ in (4.28),

$$\begin{aligned}
\bar{t}_{ij} = \Bigg\{&-\frac{2}{n_1 h}K\bigg(\frac{t_{1i}-t_{2j}}{h}\bigg)\bigg(\frac{1}{\rho(t_{1i})}+\frac{1}{\rho(t_{2j})}\bigg) + \frac{2}{n_1 h^2} \\
&\times \int K\bigg(\frac{s-t_{1i}}{h}\bigg)K\bigg(\frac{s-t_{2j}}{h}\bigg)\frac{1}{\rho(s)}\,ds\Bigg\}\sigma_1(t_{1i})\sigma_2(t_{2j}) + o\bigg(\frac{1}{N}\bigg),
\end{aligned} \tag{4.32}$$

uniformly for $i = 1, \ldots, n_1,\ j = 1, \ldots, n_2$, which implies, for the variance of $\widehat{T}_N^{(7)}$,

$$\begin{aligned}
\operatorname{var}(\widehat{T}_N^{(7)}) &= \frac{1}{N^2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\bar{t}_{ij}^2 \\
&= \frac{1}{N^2 h}\int \frac{r_1(x)(1/\lambda_1) r_2(x)}{\rho^2(x)}\sigma_1^2(x)\sigma_2^2(x)\,dx
\end{aligned}$$

$$\times\left\{16\int K^2(u)\,du-16\int(K*K)(u)K(u)\,du+4\int(K*K)^2(u)\,du\right\}$$
$$+o\left(\frac{1}{N^2h}\right)$$

$$=\frac{4}{N^2h}\int(2K-K*K)^2(u)\,du\int_0^1\frac{r_1(x)(1/\lambda_1)r_2(x)}{\rho^2(x)}$$
$$\times\sigma_1^2(x)\sigma_2^2(x)\,dx+o\left(\frac{1}{N^2h}\right). \qquad \square$$

LEMMA 4.5. *Under the assumptions of Theorem* 2.1, *we have*, *for the covariances of the statistics defined in* (4.21)–(4.24),

$$\operatorname{cov}(\widehat{T}_N^{(i)},\widehat{T}_N^{(j)})=0 \quad \textit{if}\,\{i,j\}\neq\{1,3\},\{2,4\}.$$

Under the hypothesis of equal regression functions,

$$\operatorname{cov}(\widehat{T}_N^{(1)},\widehat{T}_N^{(3)})=o\left(\frac{1}{N^2h}\right),\qquad \operatorname{cov}(\widehat{T}_N^{(2)},\widehat{T}_N^{(4)})=o\left(\frac{1}{N^2h}\right),$$

while, *under the alternative* $g_1\neq g_2$,

$$\operatorname{cov}(\widehat{T}_N^{(1)},\widehat{T}_N^{(3)})=o\left(\frac{1}{N}\right),\qquad \operatorname{cov}(\widehat{T}_N^{(2)},\widehat{T}_N^{(4)})=o\left(\frac{1}{N}\right).$$

PROOF. The first part of Lemma 4.5 is obvious. From (4.21) and (4.22) we obtain

$$\operatorname{cov}(\widehat{T}_N^{(1)},\widehat{T}_N^{(3)})=\frac{1}{N^2}\sum_{i=1}^{n_1}\alpha_{1i}k_{1i}E[\varepsilon_{1i}^3],$$

where α_{1i}, k_{1i} are defined in (4.25) and (4.26), respectively. Now (4.30) gives

$$\alpha_{1i}=\begin{cases}O(h^r)+O\left(\dfrac{1}{Nh}\right), & \text{if } g_1=g_2,\\ O(1), & \text{if } g_1\neq g_2,\end{cases}$$

uniformly for $i=1,\dots,n_1$. Similarly, we have, from (4.31),

$$k_{1i}=O\left(\frac{1}{Nh}\right),$$

uniformly for $i=1,\dots,n_1$, which implies

$$\operatorname{cov}(\widehat{T}_N^{(1)},\widehat{T}_N^{(3)})=o\left(\frac{1}{N^2h}\right)$$

under the null–hypothesis and

$$\operatorname{cov}(\widehat{T}_N^{(1)},\widehat{T}_N^{(3)})=o\left(\frac{1}{N}\right)$$

under the alternative $g_1 \neq g_2$. This proves the second part of the assertion for the statistics $\widehat{T}_N^{(1)}$ and $\widehat{T}_N^{(3)}$. The remaining case follows by exactly the same arguments and is therefore omitted. □

PROOF OF THEOREM 2.1(i). Observing Lemma 4.0 and (2.7), we obtain

$$\begin{aligned} N\sqrt{h}\Big(T_N - Bh^{2r} - \frac{1}{Nh}D\Big) &= N\sqrt{h}\Big(T_N - E[T_N]\Big) + o(1) \\ &= N\sqrt{h}\Big(\widehat{T}_N^{(5)} + \widehat{T}_N^{(6)} + \widehat{T}_N^{(7)}\Big) + o_p(1) \\ &= \sum_{s=1}^{2}\Bigg[\sum_{i=1}^{n_s}\sum_{\substack{l=1\\ l\neq i}}^{n_s}\sqrt{h}r_{il}^{(s)}\varepsilon_{si}\varepsilon_{sl}\Bigg] \\ &\quad + \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sqrt{h}\bar{t}_{ij}\varepsilon_{1i}\varepsilon_{2j} + o_p(1), \end{aligned}$$

where the second equality follows from (4.18), (4.19), (4.20) and Lemmas 4.1–4.5 and the constants B and D are defined in (2.10) for $k = 2$. Defining

$$\bar{r}_{il}^{(s)} := \frac{r_{il}^{(s)} + r_{li}^{(s)}}{2}, \quad s = 1, 2,$$

the right-hand side of this equation can be written as a symmetric quadratic form with vanishing diagonal elements, that is,

$$W_N = N\sqrt{h}(\widehat{T}_N^{(5)} + \widehat{T}_N^{(6)} + \widehat{T}_N^{(7)}) = X^T A X,$$

where $X = (X_1, \ldots, X_N)^T = (\varepsilon_{(1)}, \varepsilon_{(2)})^T$, $\varepsilon_{(i)} = (\varepsilon_{i1}, \ldots, \varepsilon_{in_i})$, $i = 1, 2$, the matrix $A = (a_{ij})_{i,\, j=1,\ldots,N}$ is given by $a_{ii} = 0, \quad i = 1, \ldots, N$,

$$a_{ij} := \begin{cases} \sqrt{h}\bar{r}_{ij}^{(1)}, & i, j = 1, \ldots, n_1,\ i \neq j, \\ \sqrt{h}\dfrac{\bar{t}_{i,\, j-n_1}}{2}, & i = 1, \ldots, n_1,\ j = n_1 + 1, \ldots, n_1 + n_2, \\ \sqrt{h}\dfrac{\bar{t}_{j,\, i-n_1}}{2}, & i = n_1 + 1, \ldots, n_1 + n_2,\ j = 1, \ldots, n_1, \\ \sqrt{h}\bar{r}_{i-n_1,\, j-n_1}^{(2)}, & i, j = n_1 + 1, \ldots, n_1 + n_2,\ i \neq j, \end{cases} \tag{4.33}$$

and $r_{ij}^{(s)}$, $\bar{t}_{ij}$ are defined in (4.27) and (4.28), respectively.

To show asymptotic normality of the statistic W_N under the hypothesis of equal regression curves, we apply Theorem 5.2 in de Jong (1987). For the asymptotic variance of W_N, we obtain, from Lemmas 4.3–4.5 and the definition of ρ in (4.4),

$$\sigma_N^2 = \mathrm{var}\Big(N\sqrt{h}(\widehat{T}_N^{(5)} + \widehat{T}_N^{(6)} + \widehat{T}_N^{(7)})\Big) = \beta^2 + o(1) = O(1), \tag{4.34}$$

where $\beta^2 = \beta^2_{2,1}$ is defined in (2.11) for $k = 2$. Observing (4.32), we have

$$
\begin{aligned}
h \sum_{j=1}^{n_2} \bar{t}_{ij}^2 = \frac{1}{n_1} \biggl\{ & \frac{4}{h} \int K^2\biggl(\frac{t_{1i}-t}{h}\biggr)\biggl(\frac{1}{\rho(t_{1i})} + \frac{1}{\rho(t)}\biggr)^2 \sigma_2^2(t) \frac{1}{\lambda_1} r_2(t)\, dt \\
& + \frac{4}{h^3} \int \biggl(\int K\biggl(\frac{s-t_{1i}}{h}\biggr) K\biggl(\frac{s-t}{h}\biggr) \frac{1}{\rho(s)}\, ds \biggr)^2 \frac{1}{\lambda_1} r_2(t) \sigma_2^2(t)\, dt \\
& + \frac{8}{h^2} \int\int K\biggl(\frac{t_{1i}-t}{h}\biggr) K\biggl(\frac{s-t_{1i}}{h}\biggr) K\biggl(\frac{s-t}{h}\biggr) \frac{1}{\lambda_1} r_2(t) \sigma_2^2(t) \\
& \times \biggl(\frac{1}{\rho(t_{1i})} + \frac{1}{\rho(t)}\biggr) \frac{1}{\rho(s)}\, ds\, dt \biggr\} \sigma_1^2(t_{1i}) + o\biggl(\frac{1}{N}\biggr) = O\biggl(\frac{1}{N}\biggr),
\end{aligned}
$$

and a similar argument implies

$$
h \sum_{j=1}^{n_1} (\bar{r}_{ij}^{(1)})^2 = O\biggl(\frac{1}{N}\biggr).
$$

From these estimates it follows that

$$
\sum_{j=1}^{N} a_{ij}^2 = h \sum_{j=1}^{n_1} (\bar{r}_{ij}^{(1)})^2 + \frac{h}{4} \sum_{j=1}^{n_2} \bar{t}_{ij}^2 = O\biggl(\frac{1}{N}\biggr), \qquad i = 1, \dots, n_1, \tag{4.35}
$$

and an analogous argument shows that (4.35) is also valid for $i = n_1+1, \dots, N$. Therefore conditions 1 and 2 in de Jong's (1987) theorem are satisfied with $K_N = \log N$. To establish the remaining condition 3 in the latter theorem, we note that by Gerschgorin's theorem the eigenvalues $\mu_1, \dots, \mu_N$ of the matrix A can be estimated as

$$
\begin{aligned}
\max_{i=1}^{n_1} |\mu_i| \le \max_{i=1}^{n_1} \sum_{l=1}^{N} |a_{il}| &\le \sqrt{h} \sum_{l=1}^{n_1} |\bar{r}_{il}^{(1)}| + \frac{\sqrt{h}}{2} \sum_{l=1}^{n_2} |\bar{t}_{il}| \\
&= O(\sqrt{h}),
\end{aligned}
$$

where we used the definition of $\bar{r}_{il}^{(s)}$ and (4.27) and (4.32) in the last estimate. Similarly, we obtain

$$
\max_{i=1}^{n_2} |\mu_i| = O(\sqrt{h}),
$$

which implies $\max_{i=1}^{N} \mu_i^2/\sigma_N^2 = o(1)$. The assertion of Theorem 2.1(i) now follows from de Jong's (1987) theorem and (4.34), that is,

$$
\sigma_N^{-1} W_N \xrightarrow{\mathscr{D}} \mathscr{N}(0, 1).
$$

4.3. *Proof of Theorem* 2.1: *fixed alternatives.* If $g_1 \neq g_2$ we have, from Lemmas 4.0–4.4,

$$\begin{aligned} \sqrt{N}\big(T_N - M^2\big) &= \sqrt{N}\big(T_N - E[T_N]\big) + o(1) \\ &= \sqrt{N}\big(\widehat{T}_N^{(1)} + \widehat{T}_N^{(2)}\big) + o_p(1) \\ &= \frac{1}{\sqrt{N}} \sum_{i=1}^{2} \sum_{j=1}^{n_i} \alpha_{ij}\varepsilon_{ij} + o_p(1) = W_N + o_p(1), \end{aligned} \tag{4.36}$$

where the last equality defines W_N. The assertion now follows from the standard central limit theorem using Liapounov's condition. To this end, we note that under a fixed alternative

$$\sigma_N^2 = \operatorname{var}(W_N) = \gamma^2 + o(1),$$

where we used Lemmas 4.1 and 4.5 and $\gamma^2 = \gamma_{2,1}^2$ is defined in (2.13) for $k = 2$. For the coefficients α_{ij} in (4.36), we have, from (4.30) for the case $i = 1$ and a similar argument in the case $i = 2$,

$$\alpha_{ij}^4 = \frac{16((\kappa_2/\kappa_i)r_{3-i})^4(t_{ij})}{\rho^4(t_{ij})} \sigma_i^4(t_{ij})(g_1 - g_2)^4(t_{ij}) + o(1), \qquad i = 1, 2,$$

which implies Liapounov's condition, that is,

$$\begin{aligned} &\frac{1}{\sigma_N^4} \sum_{i=1}^{2} \sum_{j=1}^{k} E\left|\frac{\alpha_{ij}\varepsilon_{ij}}{\sqrt{N}}\right|^4 \\ &\quad \leq \frac{16 n_1}{N^2 \sigma_N^4} \bigg(\int_0^1 \frac{((1/\lambda_1) r_2)^4(t) r_1(t)}{\rho^4(t)} \sigma_1^4(t)(g_1 - g_2)^4(t)\, dt \\ &\qquad\qquad + \int_0^1 \frac{r_1^4(t)(1/\lambda_1) r_2(t)}{\rho^4(t)} \sigma_2^4(t)(g_1 - g_2)^4(t)\, dt \bigg) + o(1) = o(1), \end{aligned}$$

and completes the proof of Theorem 2.1(ii). □

4.4. *Some comments on the consistency of the wild bootstrap.* In this section we briefly indicate the consistency of the wild bootstrap procedure used in the simulation study of Section 3. For the sake of brevity, we restrict ourselves to the statistic $T_N^{(1)}$ based on a difference of variance estimators for $k = 2$ samples. Corresponding results for $T_N^{(2)}, T_N^{(3)}$ and $k \geq 3$ regression functions can be proved following a similar pattern. Recall that we again omit all indices referring to the number of samples and to the specific statistics in our notation (e.g., we use B instead of $B_k^{(1)}$, T_N instead of $T_N^{(1)}$, etc.).

To be precise, let $\hat{g}(\cdot, h)$, $\hat{g}_i(\cdot, h)$ denote the estimates of the regression functions from the joint and individual samples defined in (2.4) and (2.6), respectively, where the dependency on the bandwidth $h > 0$ is now stated explicitly.

The residuals for the bootstrap sample are given by (see the discussion in Section 3)

$$\hat{\varepsilon}_{ij} = Y_{ij} - \hat{g}(t_{ij}, b), \qquad j = 1, \dots, n_i,\ i = 1, 2, \tag{4.37}$$

$\varepsilon^*_{ij} = V^*_{ij}\hat{\varepsilon}_{ij}$, where b is a bandwidth not necessarily equal to the bandwidth h used in the statistic T_N of Section 3.1 and the V^*_{ij} are i.i.d. with $E^*[V^*_{ij}] = 0$, $E^*[(V^*_{ij})^2] = 1$. Throughout this section E^* denotes the conditional expectation given the total sample $\{Y_{ij} \mid j = 1, \dots, n_i,\ i = 1, 2\}$ and all quantities formed from the bootstrap sample

$$Y^*_{ij} = \hat{g}(t_{ij}, b) + \varepsilon^*_{ij}, \qquad j = 1, \dots, n_i,\ i = 1, 2, \tag{4.38}$$

will be denoted with an extra * (e.g., T^*_N, $\hat{g}^*$, etc.). Under the additional assumption

$$h^{2r+1} = o(b^{2r}), \tag{4.39}$$

we sketch a proof of

$$d_2\left[N\sqrt{h}\left(T^*_N - Bh^{2r} - \frac{D}{Nh}\right), \mathscr{N}(0,\ \beta^2)\right] \xrightarrow{P} 0, \tag{4.40}$$

where B, D and $\beta^2 = \beta^2_{2,1}$ are defined in Theorem 2.1 and $d_2[\cdot,\ \cdot]$ denotes the Mallows distance [see Mallows (1972)]. Applying Lemma 8.8 of Bickel and Freedman (1981), it follows that (4.40) can be established by showing

$$d_2\big[N\sqrt{h}(T^*_N - E^*(T^*_N)), \mathscr{N}(0,\ \beta^{*2})\big] \xrightarrow{P} 0, \tag{4.41}$$

$$N\sqrt{h}\left|E^*(T^*_N) - Bh^{2r} - \tfrac{D}{Nh}\right| \xrightarrow{P} 0, \tag{4.42}$$

$$\beta^{*2} \xrightarrow{P} \beta^2, \tag{4.43}$$

where

$$\beta^{*2} = \operatorname{var}^*(N\sqrt{h}T^*_N) \tag{4.44}$$

is the conditional variance of $N\sqrt{h}T^*_N$.

PROOF OF (4.41). This follows along the lines of the proof of Theorem 2.1 in Sections 4.1–4.2 and is therefore omitted.

PROOF OF (4.43). Let $T^{*(j)}_N$ denote the bootstrap versions of the statistics $\hat{T}^{(j)}_N$ introduced in (4.21)–(4.24) such that

$$T^*_N - E^*(T^*_N) = \sum_{j=1}^{7} T^{*(j)}_N \tag{4.45}$$

[compare with (4.18), (4.19) and (4.20)]. Conditionally on $\{Y_{ij} \mid j = 1, \dots, n_i, i = 1, 2\}$, it follows by straightforward algebra

$$\begin{aligned} \operatorname{var}^*(T_N^{*(i)}) &= o_p\left(\frac{1}{N^2 h}\right), \qquad i = 1, 2, 3, 4, \\ \operatorname{cov}^*(T_N^{*(i)}, T_N^{*(j)}) &= o_p\left(\frac{1}{N^2 h}\right), \qquad i \neq j \end{aligned} \tag{4.46}$$

(just repeat the steps in the proofs of Lemmas 4.1, 4.2 and 4.5). Moreover,

$$\operatorname{var}^*(T_N^{*(5)}) = \frac{1}{N^2} \sum_{i \neq l} \frac{(r_{il}^{(1)})^2 + r_{il}^{(1)} r_{li}^{(1)}}{\sigma_1^2(t_{1,i}) \sigma_1^2(t_{1,l})} \hat{\varepsilon}_{1i}^2 \hat{\varepsilon}_{1l}^2 + o\left(\frac{1}{N}\right)$$

where $r_{il}^{(1)}$ is defined in (4.27) and its asymptotic expansion is derived in the proof of Lemma 4.3. It now follows by a straightforward calculation that the expectation of the left hand-side is given by

$$\frac{2}{N^2 h} \int_0^1 \sigma_1^4(x) \left[1 - \frac{r_1(x)}{\rho(x)}\right]^2 dx \int (2K - K * K)^2(u)\, du + o\left(\frac{1}{N^2 h}\right)$$

and a tedious calculation for the variance establishes

$$\operatorname{var}^*(T_N^{*(5)}) - \operatorname{var}(T_N^{(5)}) = o_p\left(\frac{1}{N^2 h}\right).$$

Similar arguments for $T_N^{*(6)}$, $T_N^{*(7)}$ and (4.45), (4.46) establish the assertion (4.43), that is,

$$\beta^{*2} - \beta^2 = N^2 h \sum_{j=5}^{7} \left\{\operatorname{Var}^*(T_N^{*(j)}) - \operatorname{Var}(T_N^{(j)})\right\} + o_p(1) = o_p(1).$$

PROOF OF (4.42). Recall the definition of the weights $w_{ij}^{(l)}$, $l = 1, 2$, $w_{lk,\,ij}$ in (4.1) and (4.2), respectively. To reflect the particular dependency on the bandwidth, we denote these quantities $w_{ij}^{(l)}(h)$, $w_{lk,ij}(h)$ and so on. A straightforward calculation shows that, for $i = 1, 2$,

$$\begin{aligned} E^*[\sigma_i^{*2}] &= \frac{1}{n_i} \sum_{j=1}^{n_i} \hat{\delta}_{ij} + \frac{1}{n_i} \sum_{j=1}^{n_i} \hat{\varepsilon}_{ij}^2 \\ &\quad + \frac{1}{n_i} \sum_{j=1}^{n_i} \left\{-2 w_{jj}^{(i)}(h) + \sum_{k=1}^{n_i} \left(w_{jk}^{(i)}(h)\right)^2\right\} \hat{\varepsilon}_{ij}^2 \end{aligned} \tag{4.47}$$

and

$$\begin{aligned} E^*[\hat{\sigma}^{*2}] &= \frac{1}{N} \sum_{i=1}^{2} \sum_{j=1}^{n_i} \left\{\widehat{\Delta}_{ij}^2 + \hat{\varepsilon}_{ij}^2\right\} \\ &\quad + \frac{1}{N} \sum_{i=1}^{2} \sum_{j=1}^{n_i} \left\{-2 w_{ij,\,ij}(h) + \sum_{l=1}^{2} \sum_{k=1}^{n_l} w_{ij,\,lk}^2(h)\right\} \hat{\varepsilon}_{ij}^2. \end{aligned} \tag{4.48}$$

Here $\hat{\delta}_{ij}$ and $\widehat{\Delta}_{ij}$ denote the analogues of the quantities δ_{ij} and Δ_{ij} defined in (4.9) and (4.17), respectively, where the regression functions $g(t)$ and $g_i(t)$ have been replaced by the estimate from the combined sample $\hat{g}(t, b)$. Combining (4.47) and (4.48) yields

$$
\begin{aligned}
E^*[T_N^*] &= E^*[\hat{\sigma}^{*2}] - \frac{n_1}{N} E^*[\sigma_1^{*2}] - \frac{n_2}{N} E^*[\sigma_2^{*2}] \\
&= \frac{1}{N} \sum_{i=1}^{2} \sum_{j=1}^{n_i} (\hat{\Delta}_{ij}^2 - \hat{\delta}_{ij}^2) + \frac{1}{N} \sum_{i=1}^{2} \sum_{j=1}^{n_i} \frac{k_{ij}}{\sigma_i^2(t_{ij})} \hat{\varepsilon}_{ij}^2
\end{aligned}
\tag{4.49}
$$

where k_{ij} is defined in (4.26). Observing that under the null hypothesis of equal curves $g = g_1 = g_2$, $\hat{\varepsilon}_{ij} = \sigma_i(t_{ij})\varepsilon_{ij} + g(t_{ij}) - \hat{g}(t_{ij}, b)$, we obtain, for the second term in (4.49),

$$
\begin{aligned}
\frac{1}{N} \sum_{i=1}^{2} \sum_{j=1}^{n_i} \frac{k_{ij}}{\sigma_i^2(t_{ij})} \hat{\varepsilon}_{ij}^2 - \frac{1}{Nh} D &= \frac{1}{N} \sum_{i=1}^{2} \sum_{j=1}^{n_i} k_{ij} \varepsilon_{ij}^2 - \frac{1}{Nh} D + o_p\left(\frac{1}{N\sqrt{h}}\right) \\
&= \frac{1}{N} \sum_{i=1}^{2} \sum_{j=1}^{n_i} k_{ij} E[\varepsilon_{ij}^2] - \frac{1}{Nh} D + o_p\left(\frac{1}{N\sqrt{h}}\right) \\
&= o_p\left(\frac{1}{N\sqrt{h}}\right),
\end{aligned}
\tag{4.50}
$$

where the first equality follows from the uniform consistency of the estimate $\hat{g}(\cdot, b)$ [see, e.g., Mack and Silverman (1982)], the second equality follows from (4.31) in the proof of Lemma 4.2 and the third equality is obtained by a similar argument as given in the proof of Lemma 4.0 observing the definition of k_{ij}, (4.31) and $E[\varepsilon_{ij}^2] = 1$. For the first term in (4.49) it can be proved by similar arguments

$$
\frac{1}{N} \sum_{i=1}^{2} \sum_{j=1}^{n_i} (\widehat{\Delta}_{ij}^2 - \hat{\delta}_{ij}^2) - Bh^{2r} = o_p\left(\frac{1}{N\sqrt{h}}\right)
\tag{4.51}
$$

and a combination of (4.49), (4.50) and (4.51) yields (4.42), which completes the proof of assertion (4.40). □

Acknowledgments. This paper contains parts of the second author's Diplomarbeit [Neumeyer (1999)] at the Ruhr-Universität, Bochum, Germany. The authors are grateful to Isolde Gottschlich who typed most parts of this paper with considerable technical expertise. We are also indebted to the Associate Editor and two unknown referees for constructive comments, which led to a substantial improvement of an earlier version of this paper.

REFERENCES

ALCALÁ, J. T., CHRISTÓBAL, J. A. and GONZÁLEZ-MANTEIGA, W. (1999). Goodness-of-fit test for linear models based on local polynomials. *Statist. Probab. Lett.* **42** 39–46.

AZZALINI, A. and BOWMAN, A. W. (1993). On the use of nonparametric regression for checking linear relationships. *J. Roy. Statist. Soc. Ser. B* **55** 549–557.

BERGER, J. O. and DELAMPADY, M. (1987). Testing precise hypotheses. *Statist. Sci.* **2** 317–352.

BICKEL, P. J. and FREEDMAN, D. A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* **9** 1196–1217.

BIEDERMANN, S. and DETTE, H. (2000). Testing linearity of regression models with dependent errors by kernel based methods. *Test* **3** 417–438.

DELGADO, M. A. (1993). Testing the equality of nonparametric regression curves. *Statist. Probab. Lett.* **17** 199–204.

DETTE, H. and MUNK, A. (1998). Nonparametric comparison of several regression functions: exact and asymptotic theory. *Ann. Statist*. **26** 2339–2368.

DETTE, H. and NEUMEYER, N. (1999). Nonparametric analysis of covariance. Technical Report 262, Dept. Math., Ruhr-Universität, Bochum, Germany.

FAN, J. (1992). Design-adaptive nonparametric regression. *J. Amer. Statist. Assoc*. **87** 998–1004.

FAN, J. and GIJBELS, I. (1996). *Local Polynomial Modelling and Its Applications*. Chapman and Hall, London.

GASSER, T., MÜLLER, H.-G. and MAMRITZSCH, V. (1985). Kernels for nonparametric curve estimation. *J. Roy. Statist. Soc. Ser. B* **47** 238–252.

GONZÁLEZ-MANTEIGA, W. and CAO, R. (1993). Testing hypothesis of general linear model using nonparametric regression estimation. *Test* **2** 161–189.

HÄRDLE, W. and MAMMEN, E. (1993). Comparing nonparametric versus parametric regression fits. *Ann. Statist*. **21** 1926–1947.

HÄRDLE, W. and MARRON, J. S. (1990). Semiparametric comparison of regression curves. *Ann. Statist*. **18** 63–89.

HALL, P. and HART, J. D. (1990). Bootstrap test for difference between means in nonparametric regression. *J. Amer. Statist. Assoc*. **85** 1039–1049.

HALL, P. and MARRON J. S. (1990). On variance estimation in nonparametric regression. *Biometrika* **77** 415–419.

HALL, P., HUBER, C. and SPECKMAN, P. L. (1997). Covariate-matched one-sided tests for the difference between functional means. *J. Amer. Statist. Assoc*. **92** 1074–1083.

HJELLVIK, V. and TJØSTHEIM, D. (1995). Nonparametric tests of linearity for time series. *Biometrika* **82** 351–368.

DE JONG, P. (1987). A central limit theorem for generalized quadratic forms. *Probab. Theory Related Fields* **75** 261–277.

KING, E. C., HART, J. D. and WEHRLY, T. E. (1991). Testing the equality of regression curves using linear smoothers. *Statist. Probab. Lett*. **12** 239–247.

KULASEKERA, K. B. (1995). Comparison of regression curves using quasi residuals. *J. Amer. Statist. Assoc*. **90** 1085–1093.

KULASEKERA, K. B. and WANG, J. (1997). Smoothing parameter selection for power optimality in testing regression curves. *J. Amer. Statist. Assoc*. **92** 500–511.

MACK, Y. P. and SILVERMAN, B. W. (1982). Weak and strong uniform consistency of kernel regression estimates. *Z. Wahrsch. Verw. Gebiete* **61** 405–415.

MALLOWS, C. L. (1972). A note on asymptotic joint normality. *Ann. Math. Statist*. **43** 508–515.

NADARAYA, E. A. (1964). On estimating regression. *Theory Probab. Appl.* **10** 186–190.

NEUMEYER, N. (1999). Nichtparametrische Tests auf Gleichheit von Regressionsfunktionen. Diplomarbeit, Ruhr-Universität, Bochum, Germany.

RICE, J. (1984). Bandwidth choice for nonparametric regression. *Ann. Statist*. **12** 1215–1230.

ROSENBLATT, M. (1975). A quadratic measure of deviation of two-dimensional density estimates and a test of independence. *Ann. Statist*. **3** 1–14.

SACKS, J. and YLVISAKER, D. (1970). Designs for regression problems for correlated errors. *Ann. Math. Statist*. **41** 2057–2074.

STAUDTE, R. G. and SHEATER, S. J. (1990). *Robust Estimation and Testing*. Wiley, New York.

WAND, M. P. and JONES, M. C. (1995). *Kernel Smoothing*. Chapman and Hall, London.

WATSON, G. S. (1964). Smooth regression analysis. *Sankhyā Ser. A* **26** 359–372.

YOUNG, S. G. and BOWMAN, A. W. (1995). Non-parametric analysis of covariance. *Biometrics* **51** 920–931.

FAKULTÄT FÜR MATHEMATIK
RUHR-UNIVERSITÄT
44780 BOCHUM
GERMANY
E-MAIL: holger.dette@ruhr-uni-bochum.de

FAKULTÄT FÜR MATHEMATIK
RUHR-UNIVERSITÄT
44780 BOCHUM
GERMANY
E-MAIL: natalic.neumeyer@ruhr-uni-bochum.de