The Annals of Statistics
1998, Vol. 26, No. 6, 2339–2368

NONPARAMETRIC COMPARISON OF SEVERAL REGRESSION FUNCTIONS: EXACT AND ASYMPTOTIC THEORY

BY AXEL MUNK[1] AND HOLGER DETTE

Ruhr-Universität Bochum

A new test is proposed for the comparison of two regression curves f and g. We prove an asymptotic normal law under fixed alternatives which can be applied for power calculations, for constructing confidence regions and for testing precise hypotheses of a weighted L^2 distance between f and g. In particular, the problem of nonequal sample sizes is treated, which is related to a peculiar formula of the area between two step functions. These results are extended in various directions, such as the comparison of k regression functions or the optimal allocation of the sample sizes when the total sample size is fixed. The proposed pivot statistic is not based on a nonparametric estimator of the regression curves and therefore does not require the specification of any smoothing parameter.

1. Introduction. The comparison of two (or more) regression curves is a fundamental problem in applied regression analysis. Usually these curves correspond to the means of a control (C) and a treatment (T) outcome where the predictor variable is an adjustable parameter, such as time or concentration of a drug ingredient. In many cases of practical interest, a prepost design is appropriate to compare these treatments; that is, after rescaling the covariable into the unit interval we end up with a sample of n independent pairs of observations,

$$(X_i, Y_i) = (f(t_i) + \varepsilon_i, g(t_i) + \eta_i), \qquad i = 1, \dots, n, \tag{1}$$

where the design points $\{t_i\}_{i=1,\dots,n}$ are fixed, such that $0 \le t_1 < \cdots < t_n \le 1$, and the regression functions f and g correspond to the expected response of treatment and control, respectively. In this case it is common practice to base any decision concerning $f - g$ on the individual differences

$$D_i = X_i - Y_i, \qquad i = 1, \dots, n,$$

which turns the problem into a simple one-sample problem with independent observations D_i, $i = 1, \dots, n$. Hence we may apply any test for the null hypothesis that a regression function (here $f - g$) vanishes in order to make a decision about

$$H_0\colon f = g \quad \text{versus} \quad K_0\colon f \ne g. \tag{2}$$

Of course, under the assumption of a linear model, we end up with the classical analysis of covariance, provided the error distribution is normal [Hocking

Received June 1997; revised September 1998.
[1]Supported in part by Stipendium of the Deutsche Forschungsgemeinschaft.
AMS 1991 *subject classifications*. Primary 62G05; secondary 62G10, 62G30, 62G07.
Key words and phrases. Comparison of regression curves, nonparametric analysis of covariance, heteroscedastic errors, k-sample problem.

(1985)]. However, this method often has shortcomings, because the assumption of a specified linear model or the normal assumption may be violated and hence various authors started to deal with the nonparametric set-up (1). In this model, various tests for the hypotheses in (2) were suggested during the last decade, mostly motivated by assessing the goodness of fit of a specific linear model. We refer to the work of Cox, Koh, Wahba and Yandell (1988), Eubank and Spiegelmann (1990), Azzalini and Bowman or Härdle and Mammen (1993) and the references given there.

In this paper we generalize the above set-up to the comparison of two (or several) independent groups corresponding to C and T, typically with different sample sizes m and n. Hence, a reduction to the individual differences D_i is not possible any more. Although this problem is rather important in many fields where empirical methods are applied, the area of testing hypotheses as in (2) under this assumption is relatively new. Härdle and Marron (1990) and King, Hart and Wehrly (1991) compare two regression functions that have been estimated using kernel estimators. Hall and Hart (1990) provide a bootstrap test in that set-up; however, all these approaches require the assumption of equal sample sizes $n = m$.

The aim of this paper is to provide a consistent test for the problem (2) in the situation of two independent samples with possibly unequal sample sizes. Our approach is based on measuring the discrepancy between f and g by a weighted L^2-distance $M^2 = M^2(f, g) = ||f - g||^2$. This approach is somewhat related to King, Hart and Wehrly's work [(1991), page 240]. In fact, it turns out that their test statistic, considered on page 245, is a special case of our test in the case of equal sample sizes $n = m$. However, we will demonstrate that the unequal sample case is much more difficult and leads to a rather different variance of the pivot statistic compared to the equal sample case, although asymptotic normality still holds true. For example, in the case of both designs being asymptotically uniform, we find the curious result that the variance depends additionally on the numerator and denominator r and s of the reduced fraction of $r/s = m/n$. This phenomenon does not vanish asymptotically as one might expect. In Section 2 an explicit expression of the limiting variance will be given. For the case of nonuniform designs, we could not find such an explicit expression; however, a central limit theorem can still be proved, which allows for a practical performance of the corresponding test.

Particularly, we show asymptotic normality in all cases under the null hypothesis H_0 in (2) as well as under fixed alternatives $||f - g||^2$. This allows the computation of confidence intervals for M^2 as well as testing precise hypotheses

$$H_\pi\colon M^2 \geq \pi \quad \text{versus} \quad K_\pi\colon M^2 \leq \pi,$$

where π is the minimal distance between regression functions considered relevant. These results are generalized to the comparison of more than two (independent) treatments.

The paper will be organized as follows. In Section 2 we define the test statistic and present as the main result its asymptotic normality under fixed alternatives. The estimation of M^2 in the unequal sample case is crucial for our analysis. The proposed test does not depend on any choice of bandwidth or other smoothing parameters and is very simple to perform because of the asymptotic normal law and the fact that all statistics under consideration can be represented as quadratic forms. Section 3 is devoted to additional generalizations and remarks. In Sections 4.1 and 4.2 we report a small simulation study on size and power of our test and a comparison with other procedures. In Section 4.3 we present a data example and reanalyze briefly the rain data discussed by Hall and Hart (1990). In order to improve readability for those readers who are not interested in technical details, all proofs will be deferred to the Appendix.

2. The estimator: some asymptotic and finite sample theory. Consider two independent samples of m and n observations:

$$X(t_{1,i}), Y(t_{2,j}) \in L^2, \qquad i = 1, \dots, m, \;\; j = 1, \dots, n,$$

where (w.l.o.g.) the range of the (deterministic) regressor t is the unit interval $[0, 1]$. Let further $t_{1,i} < t_{1,i+1}$ ($t_{2,j} < t_{2,j+1}$), $i = 1, \dots, m-1$ ($j = 1, \dots, n-1$); that is, we will not allow for repeated measurements, and define $t_{1,0} = t_{2,0} := 0$, $t_{1,m+1} = t_{2,n+1} := 1$. We assume that we may rewrite each random variable as

$$\begin{aligned} X_i &= X(t_{1,i}) = f(t_{1,i}) + \varepsilon(t_{1,i}) = f_i + \varepsilon_i, \qquad i = 1, \dots, m, \\ Y_j &= Y(t_{2,j}) = g(t_{2,j}) + \eta(t_{2,j}) = g_j + \eta_j, \qquad j = 1, \dots, n, \end{aligned}$$

where the regression functions and random variables satisfy $f, g \in L^2[0, 1]$ and $\varepsilon, \eta \in L^2$, such that

$$E[\varepsilon(t)] = E[\eta(t)] = 0, \qquad V[\varepsilon(t)] = \sigma_\varepsilon^2(t), \qquad V[\eta(t)] = \sigma_\eta^2(t).$$

Throughout this paper, the mean functions f and g and variance functions σ_ε^2 and σ_η^2 are required to be Hölder continuous of order $\gamma > 1/2$, that is,

$$\text{(C1)} \qquad f, g, \sigma_\varepsilon^2, \sigma_\eta^2 \in H_\gamma[0, 1] \quad \text{for some } \gamma > 1/2.$$

For the ratio of the sample sizes we assume

$$\text{(C2)} \qquad m \le n \quad \text{such that} \lim_{n, m \to \infty} m/n \to \lambda \in (0, 1].$$

We use the following notation throughout $a \vee b = \max\{a, b\}$, $a \wedge b = \min\{a, b\}$. For the moment, we restrict our consideration to asymptotically uniform designs (see Section 3 for the general case),

$$\text{(C3)} \qquad \begin{aligned} \max_{i=0,\dots,m} \left| \Delta_{1,i} - \frac{1}{m} \right| &= o((m \vee n)^{-1}), \\ \max_{j=0,\dots,n} \left| \Delta_{2,j} - \frac{1}{n} \right| &= o((m \vee n)^{-1}), \end{aligned}$$

where $\Delta_{1,i} = t_{1,i+1} - t_{1,i}$ and $\Delta_{2,j} = t_{2,j+1} - t_{2,j}$ denote the differences of successive locations of measurements. Observe that (C2) implies $o(n^{-1}) = o(m^{-1}) = o((m \vee n)^{-1})$.

2.1. *Estimation of the L^2 distance.* As a measure of discrepancy between f and g we propose

$$M^2 = ||f - g||^2 = \int_0^1 (f(t) - g(t))^2\, dt,$$

which allows us to rewrite the hypotheses in (2) as

$$H_0\colon M^2 = 0 \quad \text{versus} \quad K_0\colon M^2 \neq 0.$$

In the unequal sample case, the following decomposition of M^2 is crucial for the estimation of M^2. Rewriting

$$M^2(f, g) = \sum_{i=0}^{m} \sum_{j=0}^{n} \int_0^1 \mathbf{I}_{[t_{1,i},\, t_{1,i+1})}(t) \mathbf{I}_{[t_{2,j},\, t_{2,j+1})}(t)\, (f(t) - g(t))^2\, dt \tag{3}$$

suggests estimating M^2 as

$$\hat{M}^2 := \sum_{i=0}^{m} \sum_{j=0}^{n} \lambda_{ij} (X_{i+1} - Y_{j+1})(X_i - Y_j),$$

where $X_0 := X_1,\ Y_0 := Y_1$ and $X_{m+1} := X_m,\ Y_{n+1} := Y_n,\ m, n \geq 1$. Here the weights λ_{ij} are defined by

$$\begin{aligned} \lambda_{ij} &:= \int_0^1 \mathbf{I}_{[t_{1,i},\, t_{1,i+1})}(t)\, \mathbf{I}_{[t_{2,j},\, t_{2,j+1})}(t)\, dt \\ &= (t_{1,i+1} \wedge t_{2,j+1} - t_{1,i} \vee t_{2,j}) \mathbf{I}_{\{t_{1,i+1} \wedge t_{2,j+1} > t_{1,i} \vee t_{2,j}\}}. \end{aligned} \tag{4}$$

The estimator $\hat{M}^2$ can be motivated as follows. Each product of differences $D_{ij} = (X_{i+1} - Y_{j+1})(X_i - Y_j)$ has expectation $(f(t_{1,i+1}) - g(t_{2,j+1}))(f(t_{1,i}) - g(t_{2,j}))$, which has to be weighted with λ_{ij} in order to approximate

$$\mathbf{I}_{[t_{1,i},\, t_{1,i+1})}(t) \mathbf{I}_{[t_{2,j},\, t_{2,j+1})}(t)\, (f(t) - g(t))^2$$

in (3). Integration with respect to t yields $\hat{M}^2$.

LEMMA 1. *Under conditions* (C1)–(C3) *we have*

$$E[\hat{M}^2 - M^2] = o((n \vee m)^{-1/2}).$$

In the next theorem we assume that the variances

$$\sigma_\varepsilon^2(t) \equiv \sigma_\varepsilon^2 \quad \text{and} \quad \sigma_\eta^2(t) \equiv \sigma_\eta^2 \tag{5}$$

are constant but not necessarily equal. Further, we rewrite the ratio of the sample sizes as

$$m/n = r/s \leq 1 \quad \text{such that } gcd(r, s) = 1, \tag{6}$$

where $gcd(r, s)$ denotes the greatest common divisor of r and s. If $m = n$ we abbreviate $r = s = 1$. If condition (6) holds, we define, for integers $n \geq m$,

$$\iota(r, s) = (r+s)\frac{1 - r^2 + 3rs}{3rs^2}. \tag{7}$$

THEOREM 1. *Under conditions* (C1)–(C3) *and* (5) *we have for* $m, n \to \infty$ *such that*

$$m/(n+m) \equiv \kappa \in (0, 1/2], \tag{8}$$

that

$$(n+m)^{1/2}(\hat{M}^2 - M^2) \Longrightarrow_{\mathscr{D}} \mathscr{N}(0, \xi^2),$$

where $\mathscr{N}(0, \xi^2)$ *denotes the centered normal distribution with variance* ξ^2 *given by*

$$\xi^2(M^2) = \xi^2 = \kappa^{-1}(\sigma_\varepsilon^4 + 4\sigma_\varepsilon^2 M^2) + (1-\kappa)^{-1}(\sigma_\eta^4 + 4\sigma_\eta^2 M^2) + 2\sigma_\varepsilon^2\sigma_\eta^2\iota(r, s).$$

In particular, if $n/m = q \in \mathbb{N}$ *we have*

$$\xi^2 = (q+1)(\sigma_\varepsilon^4 + 4\sigma_\varepsilon^2 M^2) + (1 + q^{-1})(\sigma_\eta^4 + 4\sigma_\eta^2 M^2) + 2\sigma_\varepsilon^2\sigma_\eta^2(1 + q^{-1})$$

and if $n = m$ *this reduces to*

$$\xi^2 = 2(\sigma_\varepsilon^2 + \sigma_\eta^2)^2 + 8M^2(\sigma_\varepsilon^2 + \sigma_\eta^2).$$

REMARK 1. Two remarks are appropriate when condition (C3) is restricted to an equidistant design, that is,

$$t_{1,i} = i/m, \; t_{2,j} = j/n, \qquad i = 1, \dots, m; \; j = 1, \dots, n. \tag{9}$$

On the one hand, it is shown in Lemma A1 and the proof of Theorem 1 (see the Appendix) that in this case ξ^2 is exactly the finite sample variance of $(n+m)^{1/2}\hat{M}^2$ if $f \equiv g \equiv$ const. On the other hand, it follows from the proof of Theorem 1 (see the Appendix) that assumption (8) in Theorem 1 can be relaxed to sequences satisfying $m \wedge n \to \infty$, such that $m/(n+m) \to \kappa \in (0, 1/2]$. In this case one has to replace $\iota(r, s)$ in Theorem 1 by $\lim_{m\wedge n \to \infty} \iota(r_m, s_n)$ where (r_m, s_n) defines a sequence of integers which fulfills $(r_m, s_n) \to (r, s)$, such that (6) holds.

REMARK 2. The function ι occuring in the limiting variance ξ^2 of Theorem 1 deserves a more detailed discussion. Note that ξ^2 does not depend solely on the ratio of the standardized sample size κ, as one might expect from various other two-sample limit theorems. Suprisingly, we find that the numerator r and denominator s of the reduced fraction m/n determines the asymptotic variance additionally. Note further that this phenomenon does not vanish asymptotically.

REMARK 3. In order to obtain confidence intervals or tests concerning $||f - g||^2$, it remains to estimate the unknown variances $\sigma_\varepsilon^2, \sigma_\eta^2$ occuring in the limiting variance ξ^2 in Theorem 1. To this end, standard estimators of the variance in nonparametric regression can be utilized. For an overwiew and comparsion of various estimators, see Carter and Eagleson (1992) and Buckley, Eagleson and Silverman (1988) or Dette, Munk and Wagner (1997, 1998). A very simple method was suggested by Rice (1984) which leads to an estimator for σ_ε^2 and σ_η^2 as

$$\hat{\sigma}_\varepsilon^2 := \frac{1}{2(m-1)} \sum_{i=2}^{m} (X_i - X_{i-1})^2, \; \hat{\sigma}_\eta^2 := \frac{1}{2(n-1)} \sum_{j=2}^{n} (Y_j - Y_{j-1})^2.$$

A similar estimator was suggested by Gasser, Sroka and Jennen–Steinmetz (1986). Asymptotically more efficient estimators can be found in Hall and Marron (1990) or Hall, Kay and Titterington (1990), whereas Buckley, Eagleson and Silverman (1988) in the normal case and Ullah and Zinde-Walsh (1992) in the nonnormal case gave exact minimax estimators with respect to the MSE for finite sample sizes.

REMARK 4. It is worthwhile to mention that in the situation (1) the estimator $\hat{M}^2$ is based on the differences $D_i = X_i - Y_i$. More precisely, under the assumption (1) we have $m = n$, $t_{1,i} = t_{2,i} (i = 1, \dots, m)$ and

$$\lambda_{ij} = \begin{cases} \Delta_{1,i}, & \text{if } i = j, \\ 0, & \text{if } i \neq j, \end{cases}$$

which gives

$$\hat{M}^2 = \Delta_{1,0} D_1^2 + \Delta_{1,m} D_m^2 + \sum_{i=1}^{m-1} \Delta_{1,i} D_{i+1} D_i.$$

In this case the result of Theorem 1 is also correct for dependent samples, that is,

$$\sqrt{m}(\hat{M}^2 - M^2) \Rightarrow_{\mathscr{D}} \mathscr{N}(0, \tau^4 + 4M^2\tau^2),$$

where $\tau^2 = V(D_i) = V(X_i - Y_i)$.

REMARK 5. Note, that our approach allows an immediate generalization for a weighted L^2 distance

$$M^2 = ||f - g||_W^2 = \int_0^1 (f(t) - g(t))^2 W(t)\, dt,$$

where $W \in H_\gamma[0, 1]$, $\gamma > 1/2$, denotes a positive weighting function. Results for general W are immediate extensions of the previous results.

3. Further applications and extensions.

3.1. *Allocation of sample sizes.* In the following we discuss briefly the optimal allocation of sample sizes when the total sample size is fixed, say $N = m+n$, in order to minimize the variance of the test statistic $(n+m)^{1/2}\hat{M}^2$.

Observe that even in the case $\sigma_\varepsilon^2 = \sigma_\eta^2 = 1$ and $M^2 = 0$, the variance ξ^2 is not minimized when $m = n$ as one might expect. Assume in the following an equidistant design as in (9) and note that in this case the asymptotic variance ξ^2 of $\sqrt{m+n}\hat{M}^2$ equals the finite sample variance if $f \equiv g \equiv$ const. In order to determine the optimal allocation of sample sizes n and m for a given total sample size $N = n + m$ which minimizes ξ^2, the following proposition will be helpful.

PROPOSITION 1. *The function* $\iota(r, s)$ *defined in* (7) *satisfies*

$$\begin{aligned} \iota(1,1) &= 2, \\ \iota(1,s) &= 1 + s^{-1}, \\ \iota(r, r+l) &\to 4/3 \quad \text{as } r \to \infty,\ \forall\, l \in \mathbb{N}, \\ \iota(r,s) &\to 1 \quad \text{as } s \to \infty,\ \forall\, r \in \mathbb{N}. \end{aligned} \tag{10}$$

In particular, we obtain for sequences $r_n \to r$, $s_n \to s$, $r_n \le s_n$ *such that* $r_n/s_n \to \lambda$,

$$\lim_{r_n/s_n \to \lambda} \iota(r_n, s_n) = 1 + \tfrac{1}{3}(s^{-2}\kappa^{-1} - \lambda^2 + 2\lambda),$$

which reduces for irrational λ *to*

$$\lim_{r_n/s_n \to \lambda} \iota(r_n, s_n) = 1 - 1/3\lambda^2 + 2/3\lambda.$$

PROOF. The first four identities follow directly from the definition of $\iota(r,s)$ and the last two identities can be derived from (19) in the Appendix. □

Note that the second part of the preceding proposition refers to a peculiar property of the variance of the statistic $\hat{M}^2$, denoted by $\gamma(m,n) = V(\sqrt{n+m}\hat{M}^2)$. More precisely, consider an equidistant design as in (9); then it follows from Remark 1 that the assertion of Theorem 1 remains valid for sequences (m,n) such that $\lim_{m,n\to\infty} m/(m+n) = \kappa = \lambda/(\lambda+1) \in (0, 1/2]$. Now sequences (m,n) with the same limit κ may yield different asymptotic variances. For example, we have

$$\lim_{m\to\infty} \gamma(m,m) = 2(\sigma_\varepsilon^2 + \sigma_\eta^2)^2 + 8M^2(\sigma_\varepsilon^2 + \sigma_\eta^2)$$

$(r_m = s_m = r = s = 1$ in Proposition 1) and

$$\lim_{m\to\infty} \gamma(m,m+1) = 2(\sigma_\varepsilon^2 + \sigma_\eta^2)^2 + 8M^2(\sigma_\varepsilon^2 + \sigma_\eta^2) - \tfrac{4}{3}\sigma_\varepsilon^2\sigma_\eta^2$$

$(r_m = m, s_m = m+1$ in Proposition 1). On the other hand, whenever λ is irrational, all sequences will yield the same asymptotic variance.

In the case $\sigma_\varepsilon^2 = \sigma_\eta^2 = 1$ and $M^2 = 0$, the variance ξ^2 simplifies to

$$\begin{aligned} \xi^2 &= \kappa^{-1} + (1-\kappa)^{-1} + 2\iota(r,s) \\ &= 4 + \tfrac{2}{3}s^{-2} - \tfrac{2}{3}r^2 s^{-2} + \tfrac{2}{3}(rs)^{-1} + \tfrac{7}{3}rs^{-1} + sr^{-1}. \end{aligned} \tag{11}$$

If $r, s \to \infty$, it can easily be seen that ξ^2 attains its infimum in the set $\{(r,s)|r \le s\} \cup \{(1,1)\}$ when $r/s \to 1$. The statistical interpretation of this observation is as follows. An efficient allocation of sample sizes (m, n) is obtained when the sample sizes are nearly equal, such that m/n cannot be reduced. This is obtained for $n = m+1$ if N is odd and for $n = m+2$ if N is even. We find that

$$\lim_{r,s\to\infty} \inf_{r,s\in\mathbb{N},\, r<s} \xi^2 = 6\tfrac{2}{3}.$$

Suprisingly, when $m = n$ we find that $\xi^2 = 8$. Hence we obtain a relative efficiency of $\sqrt{5/6} \approx 0.912$ compared to the asymptotically optimal allocation with variance $6\frac{2}{3}$. We mention that similar results can be obtained for arbitrary M^2 under the assumption of homogeneous variances. In summary, we find that in the class of equidistant designs in both groups, respectively, asymptotically a nonsymmetric allocation is more efficient than the choice of equal sample sizes. Note that the above results do, of course, also apply for the case of *asymptotically* equidistant designs in the sense of (C3) because the limting variance in Theorem 1 is still valid along subsequences of (m, n) s.t. $m/(m+n) \equiv \kappa$.

We found numerically that this asymptotic result holds true even in the finite sample case; that is, the variance is minimized for any $N \ge 3$ when $m+1 = n$, for N odd, and when $m = n+2$, for N even. Consider as an example the case where $N = 12$. The optimal allocation is $m = 5$, $n = 7$ where $\xi^2 = 1656/245 \approx 6.76$, whereas for $n = m = 6$ we have $\xi^2 = 8$, which corresponds to a relative efficiency of 0.94. Particulary, when N is even, this result is certainly somewhat suprising, because at a first glance one might expect that $n = m = N/2$ would be the variance minimizing allocation in this case. Hence the above results indicate that a good allocation of sample sizes in order to minimize the variance is a strictly interlacing design with approximately equally spaced locations of measurements.

In order to get more detailed insight into the effect of the sample sizes on the exact variance of the pivot statistic, note that from (11) it follows that the variance for the case $n = m$ can be minimized by reducing one of the sample sizes, as long as the design remains equidistant. For example, when $n = 10$, we find from Table 1 that the variances are always smaller than for the case where $n = m = 10$, as long as $m \ge 4$. At first glance this might be somewhat suprising. Recall, however, that we restricted our considerations to equidistant designs. Hence, Table 1 indicates that the combination of two equidistant designs is typically not very efficient.

TABLE 1
Variances ξ^2 for sample sizes $m = 1, \ldots, 10$ where $n = 10$, $\sigma_\varepsilon^2 = \sigma_\eta^2 = 1$, $M^2 = 0$

(m,n)	(10,10)	(9,10)	(8,10)	(7,10)	(6,10)	(5,10)	(4,10)	(3,10)	(2,10)	(1,10)
ξ^2	8	6.685	6.750	6.751	6.898	7.500	7.420	8.002	9.600	14.300

3.2. *Nonconstant intrinsic variability.* In many practical applications it is more realistic to admit a dependency of the second moment of the errors ε, η on the regression variable t. In this case Theorem 1 has an immediate generalization. The proof follows the same pattern as in Theorem 1 (see the Appendix) and is therefore omitted.

THEOREM 2. *Under assumptions* (C1)–(C3) *we have for sequences of* m, n, *such that* (8) *holds that*

$$(n+m)^{1/2}\big(\hat{M}^2 - M^2\big) \Longrightarrow_{\mathscr{D}} \mathscr{N}(0, \tilde{\xi}^2) \quad \textit{as } n, m \to \infty,$$

where the asymptotic variance is given as

$$\tilde{\xi}^2 = \kappa^{-1}\big(||\sigma_\varepsilon^2||^2 + 4||(f-g)\sigma_\varepsilon||^2\big) + (1-\kappa)^{-1}\big(||\sigma_\eta^2||^2 + 4||(f-g)\sigma_\eta||^2\big) \\ + 2||\sigma_\varepsilon\sigma_\eta||^2\iota(r,s).$$

Note that in the case of nonconstant variances, a different estimation of the limiting variance $\tilde{\xi}^2$ is required. A simple estimator can be obtained by modifying Rice's (1984) idea of local residuals. This yields

$$\frac{1}{4(m-3)}\sum_{i=2}^{m-2}(X_i - X_{i-1})^2(X_{i+2} - X_{i+1})^2$$

as an estimator of $||\sigma_\varepsilon^2||^2$ and

$$\frac{1}{4(n-3)}\sum_{j=2}^{n-2}(Y_j - Y_{j-1})^2(Y_{j+2} - Y_{j+1})^2$$

as an estimator of $||\sigma_\eta^2||^2$. Furthermore, $||(f-g)\sigma_\varepsilon||^2$ can be estimated by

$$\tfrac{1}{2}\sum_{i=0}^{m-3}\sum_{j=0}^{n}\lambda_{ij}(X_{i+1} - Y_{j+1})(X_i - Y_j)(X_{i+3} - X_{i+2})^2,$$

$||(f-g)\sigma_\eta||^2$ by

$$\tfrac{1}{2}\sum_{i=0}^{m}\sum_{j=0}^{n-3}\lambda_{ij}(X_{i+1} - Y_{j+1})(X_i - Y_j)(Y_{j+3} - Y_{j+2})^2$$

and finally $||\sigma_\varepsilon\sigma_\eta||^2$ by

$$\tfrac{1}{4}\sum_{i=0}^{m}\sum_{j=0}^{n}\lambda_{ij}(X_{i+1} - X_i)^2(Y_{j+1} - Y_j)^2.$$

Plugging these estimators in $\tilde{\xi}^2$ gives an estimator $\tilde{\xi}^2_{m,n}$ for the limiting variance in Theorem 2. A straightforward calculation (similar to that in the proof of Lemma 1 in the Appendix) shows that $E[\tilde{\xi}^2_{m,n}] = \tilde{\xi}^2 + o(n^{-\gamma})$.

3.3. *The k-sample problem.* We now consider the problem of comparing $k \geq 2$ independent samples,

$$X(t_{j,i_j}), \qquad i_j = 1, \dots, n_j, \; j = 1, \dots, k,$$

such that

$$X(t_{j,i_j}) := X_{j,i_j} = f_j(t_{j,i_j}) + \varepsilon_j(t_{j,i_j}), \qquad j = 1, \dots, k,$$

where $f_1, \dots, f_k \in H_\gamma[0,1]$, $\gamma > 1/2$, denote the regression functions and the random errors satisfy $E[\varepsilon_j(t)] = 0$, $0 < \sigma_j^2(t) := V[\varepsilon_j(t)] < \infty$, such that $\sigma_j^2 \in H_\gamma[0,1]$, $j = 1, \dots, k$. Finally, we may assume that

$$n_1 \leq n_2 \leq \cdots \leq n_k \quad \text{such that } \frac{n_j}{n_{j+1}} \to \lambda_j \in (0,1] \text{ as } n_j \to \infty, \; j = 1, \dots, k-1$$

and that the analogue to condition (C3) holds for any $i = 1, \dots, k$. Define

$$n_i/n_j = r_i/r_j \leq 1 \quad \text{such that } gcd(r_j, r_i) = 1, \; i < j, \; i, j = 1, \dots, k \tag{12}$$

and consider, in analogy to the two-sample case, as a measure of discrepancy between k regression functions $f_1, \dots, f_k$, the squared deviation of each individual regression function f_j, $j = 1, \dots, k$ from the mean $\bar{f} = (1/k) \sum f_j$, that is,

$$\begin{aligned} M_k^2 = M^2(f_1, \dots, f_k) &:= k \sum_{j=1}^{k} \int_0^1 \big(f_j(t) - \bar{f}(t)\big)^2 \, dt \\ &= \sum_{i<j} \|f_i - f_j\|^2. \end{aligned}$$

Now we estimate M_k^2 by means of estimating $M^2(f_i, f_j) = \|f_i - f_j\|^2$ by

$$\hat{M}_{ij}^2 = \hat{M}^2(f_i, f_j), \qquad i < j,$$

exactly as in the two-sample case. Therefore, we obtain as a simple estimator for M_k^2,

$$\hat{M}_k^2 := \sum_{i<j} \hat{M}_{ij}^2$$

and the following result provides its asymptotic distribution under fixed alternatives.

THEOREM 3. *Let $N = \sum_{j=1}^{k} n_j$ and consider sequences of $n_j \to \infty$, such that $n_j/N \equiv \kappa_j$, $j = 1, \dots, k$. Assume that $\sigma_j^2(t) = \sigma_j^2$ $(j = 1, \dots, k)$; then we have, under the assumptions stated above,*

$$\sqrt{N}\big(\hat{M}_k^2 - M_k^2\big) \Longrightarrow_{\mathscr{D}} \mathscr{N}(0, \xi_k^2) \quad \text{as } N \to \infty,$$

where the asymptotic variance is given by

$$\xi_k^2 = (k-1)^2 \sum_{i=1}^{k} \kappa_i^{-1} \sigma_i^4 + 4k^2 \sum_{i=1}^{k} \kappa_i^{-1} \sigma_i^2 \int_0^1 (f_i(t) - \bar{f}(t))^2 \, dt$$
$$+ 2 \sum_{i<j} \sigma_i^2 \sigma_j^2 \iota(r_i, r_j).$$

Note that in the case of nonconstant variance functions, an analogous result to Theorem 2 holds also in the k-sample case, which is omitted for the sake of brevity.

3.4. *The case of nonuniform designs.* In this section we investigate the case when condition (C3), of an asymptotically uniformly distributed design in the treatment groups, fails. To simplify the notation we restrict discussion again to the case of two treatment groups, that is, $k = 2$, with homoscedastic variances σ_ε^2 and σ_η^2, respectively. We use the notation of Section 3.3 and assume that both designs are generated by a regular density in the sense of Sacks and Ylvisaker (1970); that is,

$$\int_{t_{k,j_k-1}}^{t_{k,j_k}} h_k(t) = \frac{1}{n_k}, \qquad j_k = 1, \ldots, n_k; \; k = 1, 2 \tag{13}$$

for possibly different design densities h_1, h_2, fullfilling the condition

$$\inf_{t \in [0,1]} h_k(t) > 0, \qquad k = 1, 2, \tag{14}$$

s.t. $h_1, h_2 \in H_\gamma[0,1]$; $\gamma > \frac{1}{2}$. The corresponding cdf's are denoted as H_k; $k = 1, 2$. Let us first consider the simplest case, where $H = H_1 = H_2$ and the design points in both groups are equal (i.e., $n = n_1 = n_2$), which is covered by the next theorem.

THEOREM 4. *Under assumptions* (13) *and* (14) *in the case of equal design points* $t_{1,i} = t_{2,i}$, $i = 1, \ldots, n$ *and equal design densities* h *we have that*

$$\sqrt{2n}\big(\hat{M}^2 - M^2\big) \Longrightarrow_{\mathscr{D}} \mathscr{N}(0, \xi^2), \tag{15}$$

where

$$\xi^2 = 2\big(\sigma_\varepsilon^2 + \sigma_\eta^2\big)^2 \int_0^1 \frac{1}{h(t)} \, dt + 8\big(\sigma_\varepsilon^2 + \sigma_\eta^2\big) \int_0^1 \left(\frac{f_1(t) - f_2(t)}{h \circ H^{-1}(t)} \right)^2 dt,$$

which reduces under H_0: $f_1 = f_2$ *to*

$$\xi^2 = 2\big(\sigma_\varepsilon^2 + \sigma_\eta^2\big)^2 \int_0^1 \frac{1}{h(t)} \, dt.$$

As we have seen in Theorem 1, unequal sample sizes cause significant complication even for uniform designs in both groups. From the proof of Theorem 1 we find that this difficulty arises from the weighting factor

$$\psi_{n_1, n_2} = (n_1 + n_2) \sum_{i, j=0}^{n_1, n_2} \lambda_{ij}^2,$$

occuring in the variance of III and IV in the decomposition of $\hat{M}^2$. For arbitrary designs it seems to be impossible to find an explicit limiting expression for the term ψ_{n_1, n_2} or a characterization of those subsequences of (n_1, n_2) such that an analoguous result to Lemma A1 holds. Note that for such a result, it is necessary that ψ_{n_1, n_2} be uniformly bounded, which can in fact be shown. However, for practical purposes, the knowledge of such a limiting expression is not important, because ψ_{n_1, n_2} can be computed explicitly for finite sample sizes without much effort. Furthermore, the remaining expressions in the variance of $\sqrt{n+m}(\hat{M}^2 - M^2)$ converge, as the following theorem shows.

THEOREM 5. *Under the above assumptions, the random variable*

$$\sqrt{n_1 + n_2}(\hat{M}^2 - M^2)$$

is asymptotically normally distributed with expectation 0, *s.t.*

$$\begin{aligned} V\big[\sqrt{n_1 + n_2}(\hat{M}^2 - M^2)\big] &- 2(n_1 + n_2) \sum_{i, j} \lambda_{ij}^2 \sigma_\varepsilon^2 \sigma_\eta^2 \\ &= \kappa^{-1}\left(\sigma_\varepsilon^4 \int L_1^2(s)\, ds + 4\sigma_\varepsilon^2 M_1^2\right) \\ &\quad + (1-\kappa)^{-1}\left(\sigma_\eta^4 \int L_2^2(s)\, ds + 4\sigma_\eta^2 M_2^2\right) + o(1), \end{aligned}$$

where

$$L_k(s) = \left\{ \frac{d}{dt} H_i^{-1}(t)\Big|_{t=s} \right\}, \qquad k = 1, 2,$$

and

$$M_k^2 = \int_0^1 (f_1(s) - f_2(s))^2 L_k^2(s)\, ds$$

denote the weighted distance between f_1 *and* f_2 *with respect to the weighting function* L_k^2, $k = 1, 2$, *respectively*.

The proof follows by a careful inspection of the proof of Theorem 1 together with the arguments in the proof of Theorem 4 and will therefore be omitted. Theorem 5 can now be utilized to perform a test for the hypothesis H_0: $f = g$. Observe that under H_0 we have $M_1^2 = M_2^2 = 0$ and therefore it remains to estimate

$$2(n_1 + n_2) \sum_{i, j=0}^{n_1, n_2} \lambda_{ij}^2 \sigma_\varepsilon^2 \sigma_\eta^2 + \kappa^{-1} \sigma_\varepsilon^4 \int_0^1 L_1^2(s)\, ds + (1-\kappa)^{-1} \sigma_\eta^4 \int_0^1 L_2^2(s)\, ds.$$

To this end, use the estimators suggested in Section 3.2 for σ_ε^2 and σ_η^2. If the limiting design density is not known, a good approximation of $\int_0^1 L_k^2(s)\,ds$ can be obtained by $n_k \sum_{i_k=1}^{n_k} \Delta_{k,i_k}^2$, ($\Delta_{k,i_k} = t_{k,i_k} - t_{k,i_k-1}$, $k = 1, 2$), which corresponds to the exact expressions occuring in the variance of $\hat{M}^2$. See Section 4 for some simulation results of this method.

REMARK 6. We mention that it can be shown that the assumption (13) can be weakened to

$$\max_{i_k=1,\dots,n_k} \left| \int_{t_{k,i_k}}^{t_{k,i_k+1}} h_k(s)\,ds - \frac{1}{n_k} \right| = o((n_1 \vee n_2)^{-1}), \qquad k = 1, 2$$

without changing the limit law in Theorems 4 and 5. Furthermore, similar expressions as in Theorem 2, when the variances are not homoscedastic, and as in Theorem 3 for the k-sample case can be derived.

4. Example and simulation results.

4.1. *Simulation of the level and power.* A test based on Theorem 1 or 2 rejects the hypothesis H_0: $f = g$ if

$$\sqrt{n+m}\,\hat{M}^2 > \hat{\xi}_{m,n} u_{1-\alpha}, \tag{16}$$

where $u_{1-\alpha}$ denotes the $(1-\alpha)$ quantile of the standard normal distribution and $\hat{\xi}_{m,n}^2$ (or $\tilde{\xi}_{m,n}^2$) is an estimator of the variances ξ^2 or $\tilde{\xi}^2$ as described in Remark 3 or Section 3.2. In order to investigate the finite sample behavior of this test, we performed a small simulation study for the case of constant variances σ_ε^2 and σ_η^2. Three different set-ups were considered in this study:

(A) $f(z) = z + 1$; $g(z) = z + 1$ ($M^2 = 0$),

(B) $f(z) = z$; $g(z) = z + 1$ ($M^2 = 1$),

(C) $f(z) = 1/(z+1)$; $g(z) = z + 1$ ($M^2 = 5/6$)

and the variances $\sigma_\varepsilon^2, \sigma_\eta^2$ were chosen from the set $\{0.25, 0.5\}$. For both samples we chose an equidistant design with a total sample sizes $N = n + m = 50, 100, 200$. We considered three different partitions of the total sample size which correspond approximately to $m/n = 1, 1/2, 1/3$, that is,

$n+m$	(m,n)	(m,n)	(m,n)
50	(25, 25)	(17, 33)	(13, 37)
100	(50, 50)	(34, 66)	(25, 75)
200	(100, 100)	(67, 133)	(50, 150).

The errors were assumed to be normal and 5000 simulations were carried out for each scenario. The results are displayed in Tables 2–4 and show the simulated level of the pivot statistic $\sqrt{n+m}(\hat{M}^2 - M^2)/\hat{\xi}_{m,n}$ for the 70%, 80%,

TABLE 2
Simulated level for the regression functions $f(z) = z + 1$, $g(z) = z + 1$, $M^2 = 0$

			$m: n = 1$				$m: n \approx 1/2$				$m: n \approx 1/3$			
$m+n$	σ_ε^2	σ_η^2	0.70	0.80	0.90	0.95	0.70	0.80	0.90	0.95	0.70	0.80	0.90	0.95
	0.25	0.25	0.708	0.808	0.909	0.961	0.717	0.811	0.905	0.957	0.752	0.840	0.933	0.970
50	0.25	0.50	0.727	0.821	0.917	0.962	0.710	0.808	0.908	0.956	0.737	0.829	0.923	0.962
	0.50	0.25	0.713	0.808	0.907	0.956	0.684	0.792	0.909	0.958	0.744	0.837	0.927	0.972
	0.25	0.25	0.716	0.814	0.914	0.962	0.714	0.800	0.905	0.957	0.741	0.837	0.926	0.968
100	0.25	0.50	0.704	0.797	0.902	0.956	0.701	0.805	0.903	0.957	0.728	0.823	0.919	0.962
	0.50	0.25	0.713	0.817	0.915	0.963	0.692	0.799	0.908	0.956	0.718	0.818	0.918	0.966
	0.25	0.25	0.697	0.804	0.908	0.955	0.687	0.790	0.898	0.951	0.716	0.811	0.916	0.956
200	0.25	0.50	0.713	0.801	0.904	0.957	0.694	0.793	0.899	0.953	0.717	0.810	0.911	0.963
	0.50	0.25	0.709	0.801	0.909	0.961	0.696	0.794	0.904	0.958	0.710	0.813	0.908	0.957

90% and 95% quantile. The actual level was estimated by counting the number of outcomes larger than the corresponding normal quantiles and dividing by 5000. Simulations for the lower quantiles show similar results and are suppressed for the sake of brevity. Table 2 shows that the normal approximation can be used for moderate sample sizes ($n + m \geq 50$). Moreover, Tables 3 and 4 show that performance of the approximation is independent of the specific alternative $M^2 \geq 0$. This suggests approximating the power of the test in (16) by

$$P_{(f,g):\,\|f-g\|^2=M^2}\left\{\sqrt{n+m}\,\hat{M}^2 > \hat{\xi}_{m,n} u_{1-\alpha}\right\} \sim 1 - \Phi\left(\frac{u_{1-\alpha}\xi(0) - \sqrt{n+m}\,M^2}{\xi(M^2)}\right), \tag{17}$$

where $\xi^2(M)$ is defined in Theorem 1. In order to get an impression of the quality of the approximation (17) we consider a numerical example. For the case

TABLE 3
Simulated level for the regression functions $f(z) = z$, $g(z) = z + 1$, $M^2 = 1$

			$m: n = 1$				$m: n \approx 1/2$				$m: n \approx 1/3$			
$m+n$	σ_ε^2	σ_η^2	0.70	0.80	0.90	0.95	0.70	0.80	0.90	0.95	0.70	0.80	0.90	0.95
	0.25	0.25	0.709	0.804	0.903	0.956	0.695	0.795	0.899	0.950	0.732	0.831	0.920	0.958
50	0.25	0.50	0.710	0.799	0.899	0.951	0.702	0.800	0.902	0.951	0.726	0.818	0.908	0.951
	0.50	0.25	0.698	0.797	0.900	0.953	0.694	0.790	0.894	0.948	0.739	0.827	0.917	0.963
	0.25	0.25	0.697	0.792	0.899	0.951	0.715	0.806	0.904	0.955	0.728	0.818	0.913	0.960
100	0.25	0.50	0.707	0.806	0.910	0.961	0.692	0.789	0.902	0.947	0.722	0.817	0.910	0.954
	0.50	0.25	0.703	0.797	0.902	0.956	0.699	0.802	0.903	0.950	0.727	0.819	0.919	0.966
	0.25	0.25	0.704	0.806	0.908	0.958	0.694	0.799	0.903	0.953	0.724	0.811	0.910	0.956
200	0.25	0.50	0.709	0.810	0.909	0.955	0.699	0.803	0.904	0.955	0.712	0.812	0.909	0.955
	0.50	0.25	0.700	0.800	0.905	0.956	0.698	0.794	0.898	0.951	0.721	0.809	0.909	0.957

TABLE 4
Simulated level for the regression functions $f(z) = 1/(z+1)$, $g(z) = z+1$, $M^2 = 5/6$

			m: $n = 1$				m: $n \approx 1/2$				m: $n \approx 1/3$			
$m+n$	σ_ε^2	σ_η^2	**0.70**	**0.80**	**0.90**	**0.95**	**0.70**	**0.80**	**0.90**	**0.95**	**0.70**	**0.80**	**0.90**	**0.95**
	0.25	0.25	0.711	0.808	0.913	0.959	0.703	0.796	0.906	0.954	0.709	0.812	0.914	0.957
50	0.25	0.50	0.717	0.809	0.909	0.959	0.698	0.801	0.907	0.953	0.712	0.809	0.903	0.949
	0.50	0.25	0.714	0.809	0.904	0.952	0.693	0.794	0.897	0.949	0.720	0.812	0.912	0.960
	0.25	0.25	0.712	0.804	0.903	0.952	0.719	0.812	0.911	0.960	0.712	0.808	0.907	0.956
100	0.25	0.50	0.712	0.809	0.913	0.960	0.700	0.796	0.901	0.950	0.711	0.807	0.906	0.950
	0.50	0.25	0.711	0.796	0.908	0.958	0.697	0.804	0.904	0.953	0.707	0.808	0.910	0.958
	0.25	0.25	0.709	0.805	0.906	0.957	0.693	0.794	0.903	0.953	0.705	0.801	0.907	0.956
200	0.25	0.50	0.711	0.809	0.913	0.956	0.705	0.802	0.902	0.953	0.706	0.799	0.902	0.951
	0.50	0.25	0.705	0.802	0.907	0.954	0.690	0.797	0.903	0.954	0.712	0.801	0.899	0.951

$M^2 = 1$, $\sigma_\varepsilon^2 = \sigma_\eta^2 = 1$, $\alpha = 0.01$ and sample sizes $(m, n) = (25, 25)$, $(17, 33)$, $(13, 37)$ we find from (17) the approximations 0.547, 0.573 and 0.549. The simulated power in these cases (5000 simulations) was 0.557, 0.574 and 0.545, respectively. Other simulations showed similar results and are suppressed for the sake of brevity. As a rule of thumb, we recommend the approximation for the power (17) whenever $n, m \geq 25$. Observe that (17) provides a very simple method to determine the required sample size in order to control the type II error of the proposed test.

For an illustration of the small sample behavior of the proposed test we also considered the sample size $n + m = 30$, where $n = m = 15$ (case a) and $m = 10$, $n = 20$ (case b). Figure 1 shows, for example, a QQ-plot based on 1000 simulations for the regression functions $f(z) = z + 1$, $g(z) = z$, that is, $M^2 = 1$, where $\sigma_\varepsilon^2 = \sigma_\eta^2 = 0.25$. We observed that the distribution of the statistic $\sqrt{m+n}(\hat{M}^2 - M^2)$ is left skewed. However, the approximation of the upper quantiles is still reasonable. For example, we observe in case b 0.681, 0.783, 0.893 and 0.947 as approximations for the 70%, 80%, 90% and 95% probability. Other scenarios gave similar results and are omitted for the sake of brevity.

Finally, the asymptotic results of Theorems 4 and 5 for different designs were investigated for the case a ($M^2 = 0$). Table 5 shows results for the design densities $h_1(x) = 1$ and $h_2(x) = (2x)^{-1/2}$. Here we observe a slight loss in the accuracy of the normal approximation, especially in the case m: $n \approx 1/3$. We mention that in this case condition (14) is not even fulfilled. However, the simulations indicate that the limit law in Theorem 5 is still valid.

4.2. *Comparison with other procedures.* It might also be of interest to compare the power of the test in (16) with the tests proposed by Delgado (1993) and Hall and Hart (1990) which were developed for the case of equal sample sizes and equal control variables, that is, $t_{1,i} = t_{2,i}$ $(i = 1, \dots, m = n)$. Because Delgado (1993) demonstrated a similar power behavior of his and Hall and Hart's (1990) test, we restrict our comparison to Delgado's test. Roughly speaking,

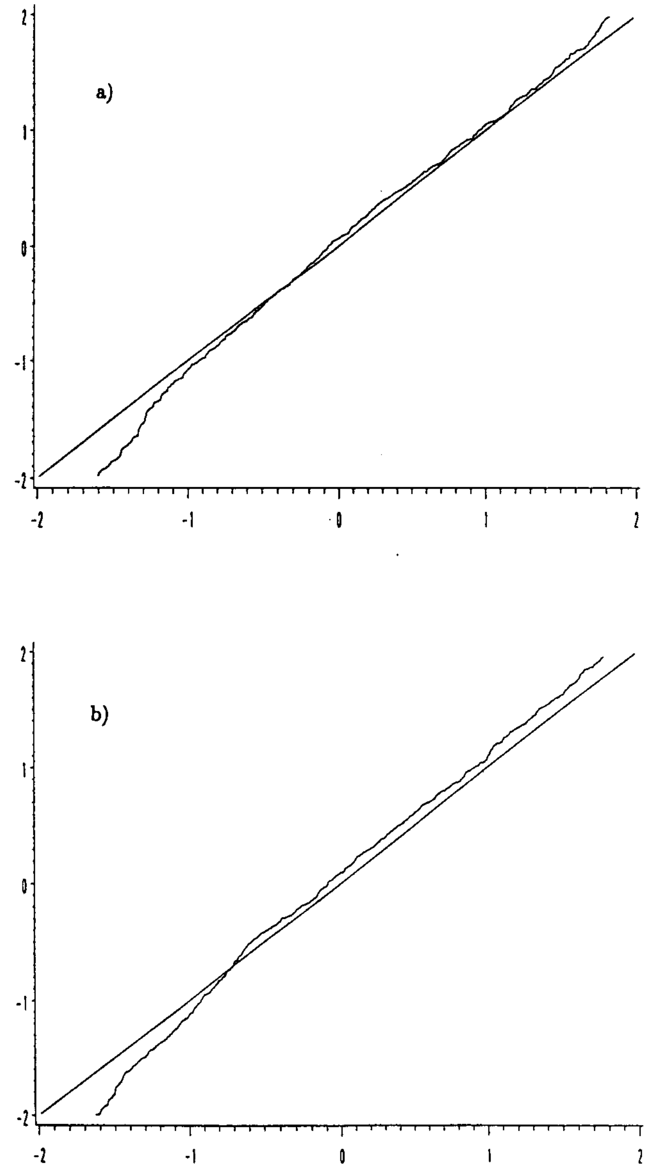

FIG. 1. *QQ-plot based on* 1000 *simulations of the statistic* $\sqrt{n+m}(\hat{M}^2 - M^2)/\hat{\xi}_{m,n}$ *for the regression functions* $f(z) = z + 1$, $g(z) = z$, *where* $n = m = 15$ (*case a*) *and* $m = 10$, $n = 20$ (*case b*). *The horizontal line gives the quantiles of the standard normal distribution.*

we found two situations which characterize the features of both tests. In the case of a nearly linear difference between f and g, Delgado's (1993) and Hall and Hart's (1990) tests are superior. When $f - g$ is more wiggly, for example, an oscilllating function, the test in (16) has larger power. In order to illustrate these findings, we displayed the simulation results for the scenario (b)

TABLE 5
Simulated level for the regression functions $f(z) = z + 1$, $g(z) = z + 1$, $M^2 = 0$ and different design densities $h_1(x) = 1$, $h_2(x) = (2x)^{-1/2}$

			$m\colon n = 1$				$m\colon n \approx 1/2$				$m\colon n \approx 1/3$			
$m+n$	σ_ε^2	σ_η^2	0.70	0.80	0.90	0.95	0.70	0.80	0.90	0.95	0.70	0.80	0.90	0.95
	0.25	0.25	0.715	0.789	0.885	0.934	0.702	0.784	0.893	0.929	0.709	0.790	0.881	0.927
50	0.25	0.50	0.709	0.793	0.881	0.932	0.706	0.788	0.889	0.934	0.707	0.792	0.878	0.928
	0.50	0.25	0.706	0.787	0.879	0.938	0.707	0.789	0.887	0.932	0.711	0.796	0.871	0.925
	0.25	0.25	0.710	0.793	0.884	0.936	0.708	0.789	0.891	0.932	0.717	0.796	0.888	0.929
100	0.25	0.50	0.700	0.793	0.872	0.931	0.721	0.806	0.889	0.938	0.707	0.789	0.881	0.931
	0.50	0.25	0.700	0.786	0.875	0.932	0.712	0.790	0.889	0.934	0.705	0.789	0.883	0.931
	0.25	0.25	0.708	0.792	0.889	0.944	0.715	0.796	0.883	0.941	0.712	0.790	0.884	0.931
200	0.25	0.50	0.711	0.806	0.892	0.951	0.709	0.796	0.891	0.944	0.714	0.794	0.892	0.939
	0.50	0.25	0.706	0.790	0.882	0.942	0.697	0.784	0.881	0.938	0.698	0.786	0.876	0.934

($M^2 = 1 \equiv f - g$), which corresponds to the case (iii) in Delgado (1993) and the case of a more oscillating difference $f(x) - g(x) = \sin(2\pi x)$ ($M^2 = 1/2$). In accordance with Delgado's (1993) Monte Carlo study, we considered three choices for the distributions of the errors (ε_j, η_j), namely,

$$\text{(i)} \qquad (\mathcal{N}_1, \mathcal{N}_2),$$

$$\text{(ii)} \qquad (|\mathcal{N}_1| - \sqrt{2/\pi}, |\mathcal{N}_2| - \sqrt{2/\pi}),$$

$$\text{(iii)} \qquad (|\mathcal{N}_1| - \sqrt{2/\pi}, \sqrt{2/\pi} - |\mathcal{N}_2|),$$

where $\mathcal{N}_1$ and $\mathcal{N}_2$ are independent standard normal variables. The results are listed in Tables 6 and 7 and show that the superiority of one of the two procedures depends sensitively on the specific alternative. If $f - g \equiv 1$, the test of Delgado has more power, especially in the case of normally distributed errors in both samples. On the other hand, for a more oscillating difference $f(x) - g(x) = \sin(2\pi x)$, the test (16) proposed in this paper yields a larger power than Delgado's (1993) test, particulary in the case (iii).

TABLE 6
Simulated power of Delgado's (1993) *test for the alternative* $f - g \equiv 1, f(x) - g(x) = \sin(2\pi x)$*

$f-g$	1			$\sin(2\pi x)$		
(ε, η)	(i)	(ii)	(iii)	(i)	(ii)	(iii)
$m = 15$	0.564	0.942	0.985	0.048	0.083	0.076
	0.744	0.982	0.999	0.128	0.223	0.211
$m = 30$	0.867	0.999	0.999	0.064	0.245	0.216
	0.952	1.000	1.000	0.193	0.513	0.481

*First row: significance level $\alpha = 0.01$; second row: $\alpha = 0.05$.

TABLE 7
Simulated power the test in (16) *for the alternative* $f - g \equiv 1$, $f(x) - g(x) = \sin(2\pi x)^*$

$f-g$ (ε, η)	1 (i)	1 (ii)	1 (iii)	$\sin(2\pi x)$ (i)	$\sin(2\pi x)$ (ii)	$\sin(2\pi x)$ (iii)
$m = 15$	0.401	0.835	0.889	0.191	0.517	0.514
	0.533	0.904	0.951	0.304	0.651	0.653
$m = 30$	0.561	0.976	0.993	0.272	0.748	0.729
	0.704	0.989	0.998	0.432	0.854	0.838

*First row: significance level $\alpha = 0.01$; second row: $\alpha = 0.05$.

4.3. *Example.* In this section we reanalyze an example which has been discussed previously by Hall and Hart (1990). They compared the towns of Coweeta and Lewiston, North Carolina, for the concentration of sulfate in rain as a function of time in a 261-week period between 1979 and 1983. These data were obtained as a part of the National Acid Deposition Project where measurements were taken weekly. After adjusting the acid concentration to covariate "amount of rainfall" Hall and Hart (1990) reduced this data set to a problem with equal sample sizes by means of dropping those weeks where no data were available at both towns simultaneously. After dropping the missing values, 189 of 261 weeks finally remained in the study, a loss of about 28% of the data. We refer to Hall and Hart (1990), page 1045 for an illustrative scatterplot of the adjusted data together with kernel regression estimates. In summary, these authors found that there is strong evidence of a difference between the two curves.

In Lewiston there were $m = 215$ weeks of data and in Coweeta $n = 220$ weeks. In fact, the observations were taken rather equidistantly over the period of 261 weeks, hence approximation (C3) by a uniform design is reasonable. For the application of the test (16) we divided the covariable by 261 in order to adjust to a $[0, 1]$-range. Observe that this leaves the pivot statistic invariant when we adjust the variance estimator $\hat{\xi}^2_{m,n}$ correspondingly. We compute $(r, s) = (43, 44)$ and hence $\iota(r, s) = 1.33351$. Note that this value is very well approximated by (10) in Proposition 1. Assuming constant variances of the logarithms of the acid concentration in each town, we computed $\hat{\sigma}^2_\eta = 0.3476$ and $\hat{\sigma}^2_\varepsilon = 0.6516$. Finally, $\hat{M}^2 = 0.2224$ and $T_{m,n} = \sqrt{n+m}\,\hat{M}^2/\hat{\xi}_{m,n} = 3.5558$, that is, we observe a P-value $< 2 \cdot 10^{-4}$. Note that Hall and Hart's [(1990), Section 3.4] analysis gives a similar result. With a bootstrap method developed in this paper, these authors obtained a P-value $< 1.5 \cdot 10^{-4}$, which gives even more evidence against the equality of both regression curves. However, they reduced this data set by dropping missing values, that is, their anaylsis was restricted to those weeks where both measurements of the log concentration were available. In this case $m = n = 189$, that is, $\iota(1, 1) = 2$. For illustrative purposes, let us finally compare this outcome with our test restricted to this situation. We obtain $\hat{\sigma}^2_\varepsilon = 0.6072$, $\hat{\sigma}^2_\eta = 0.3538$ and $\hat{M}^2 = 0.12275$. Hence

$T_{189,189} = 1.756$, which gives a P-value of about 0.04. In summary, there is certainly evidence against equality of f and g. Observe, nevertheless, that this test would not reject at a 1% error rate. This difference from the above analysis utilizing the full data set can be explained by the fact that the distance M^2 is estimated much larger in the first case ($\hat{M}^2 = 0.2224$) than in the second case ($\hat{M}^2 = 0.1227$), whereas the standard deviations $\hat{\xi}_{m,n}$ are nearly found to be the same (1.3045 in the first case and 1.3590 in the second case). Comparing this with Hall and Hart's (1990) result, we see that in the case of equal sample sizes, their bootstrap test provides stronger evidence against equality of the regression curves, indicating a higher power in this situation.

APPENDIX

Proofs.

PROOF OF LEMMA 1. Recall that f, g are Hölder continous of order $\gamma > 1/2$. Taking into account that $\#\{\lambda_{ij}\colon \lambda_{ij} \neq 0, i = 1, \dots, m,\ j = 1, \dots, n\} \leq n + m$ [where λ_{ij} was defined in (4)] we estimate

$$\begin{aligned}
&\left|M^2 - E\hat{M}^2\right| \\
&\quad \leq C(f, g) \int_0^1 \sum_{i,j} \mathbf{I}_{[t_{1,i}, t_{1,i+1})}(t)\mathbf{I}_{[t_{2,j}, t_{2,j+1})}(t)\left\{|t_{1,i} - t|^\gamma + |t_{2,j} - t|^\gamma\right\} dt \\
&\qquad + o((n \vee m)^{-\gamma}) \\
&\quad = o((n \vee m)^{-\gamma}),
\end{aligned}$$

where $C(f, g)$ denotes a generic constant depending only on f and g. $\square$

LEMMA A1. *Consider sequences of m, n, such that* (8) *and* (6) *hold. If we assume an asymptotic equidistant design, that is, condition* (C3) *holds, then*

$$\lim_{n,m\to\infty} (n + m) \sum_{i,j=0}^{m,n} \lambda_{ij}^2 = \iota(r, s),$$

where λ_{ij} was defined in (4). *If the case of an equidistant design $t_{1,i} = i/m$, $t_{2,j} = j/n$, it holds that*

$$(n + m) \sum_{i,j=0}^{m,n} \lambda_{ij}^2 = \iota(r, s).$$

PROOF. Throughout this paper, let for $x, y \in \mathbb{Z}$, $x \bmod y$ be defined as the smallest nonnegative integer (possibly 0) such that $x - (x \bmod y)$ is divisible by y. We first consider the case of an equidistant design. We have $r/m = s/n$ and hence it is sufficient to restrict to the case $i/m, j/n \in [0, r/m]$ because

$$(n + m) \sum_{i,j=0}^{m,n} \lambda_{ij}^2 = (n + m)n/s \sum_{i,j=0}^{r,s} \lambda_{ij}^2.$$

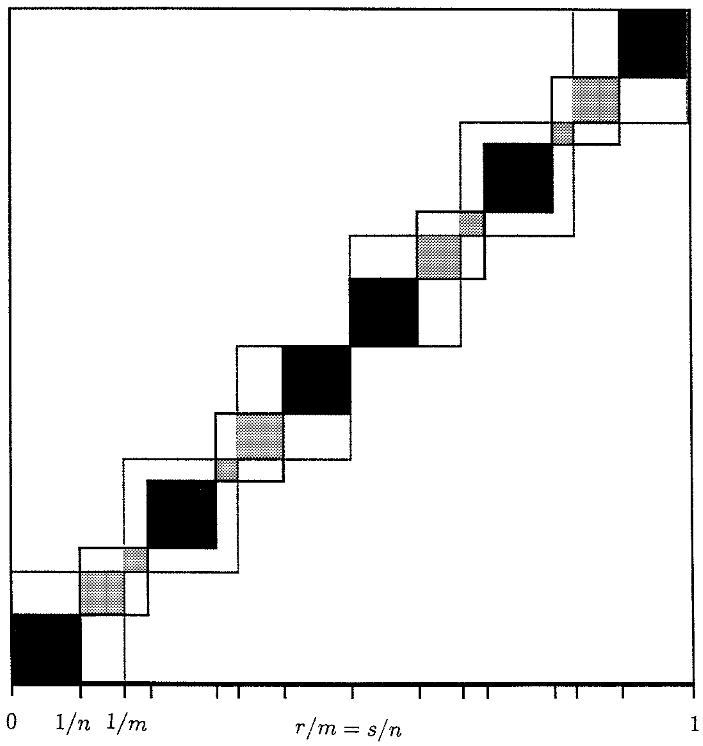

FIG. 2. *The function* $\sum_{i,\,j=0}^{m,n} \lambda_{ij}^2$ *for* $m = 6,\ n = 10,\ r = 3,\ s = 5.$

For the following derivation it is helpful to look at Figure 2. For any $x \in \mathbb{R}$, define $[x]$ to be the largest integer less than or equal to x (possibly 0). We have

$$\begin{aligned} \sum_{i,\,j=0}^{r,\,s} \lambda_{ij}^2 &= 2\sum_{i=1}^{r}\left(\frac{i}{m} - \frac{1}{n}\left[\frac{is}{r}\right]\right)^2 + \frac{s-r+1}{n^2} \\ &= 2n^{-2}\sum_{i=1}^{r-1}\left(i\frac{s}{r} - \left[i\frac{s}{r}\right]\right)^2 + \frac{s-r+1}{n^2} \qquad (18) \\ &= 2n^{-2}\sum_{i=1}^{r-1}\left(\frac{(is) \bmod r}{r}\right)^2 + \frac{s-r+1}{n^2}, \end{aligned}$$

where the first equality in (18) can be seen as follows. Observe that $\sum_{i,\,j=0}^{m,n} \lambda_{ij}^2$ corresponds to the area of the shaded region in Figure 2. In the first expression on the r.h.s. of (18), only those λ_{ij}^2 occur which are not equal to n^{-2} (light shaded squares); that is, those partitions of $[0, 1]$ are counted where $t_{1,\,i} - t_{2,\,j} < n^{-1}$. Note that in this case $t_{1,\,i} - t_{2,\,j} = i/m - 1/n[is/r]$. The factor 2 on the r.h.s. of (18) is obtained by symmetry. Hence exactly $(s - r + 1)$ subintervals remain with length n^{-1} which are entirely contained in an interval of length m^{-1} (dark shaded squares). Observe now that $(is) \bmod r$, $i = 0, \dots, r-1$ represents all residue classes of $\mathbb{Z}/r\mathbb{Z}$ and hence we obtain

$$
\begin{aligned}
(n+m)\sum_{i,\,j=0}^{m,\,n}\lambda_{ij}^2 &= \frac{n+m}{s}\left(2n^{-1}\sum_{i=1}^{r-1}\frac{i^2}{r^2}+\frac{s-r+1}{n}\right)\\
&= (r+s)\frac{1-r^2+3rs}{3rs^2}\\
&= 1+\frac{1}{3s^2}-\frac{r^2}{3s^2}+\frac{1}{3rs}+\frac{2r}{3s} = \iota(r,s).
\end{aligned} \tag{19}
$$

Now we are in the position to prove the general result for asymptotic uniform designs according to condition (C3). We show first that we can find M, N such that for all $m \ge M$ and $n \ge N$ the following property holds:

$$
t_{1,\,i} < t_{2,\,j} \iff \frac{i}{m} < \frac{j}{n} \quad \forall (i,j) \in \{1,\dots,m\}\times\{1,\dots,n\}. \tag{20}
$$

To this end observe that

$$
\begin{aligned}
(n+m)\inf_{(i,\,j):\,i/m\neq j/n}\left|\frac{i}{m}-\frac{j}{n}\right| &= \inf_{(i,\,j):\,i/m\neq j/n}\left|i\left(1+\frac{s}{r}\right)-j\left(1+\frac{r}{s}\right)\right|\\
&= \inf_{\substack{i=0,\dots,r-1\\ j=1,\dots,s}}\left|i\frac{s(r+s)}{rs}-j\frac{(s+r)r}{rs}\right|\\
&= \inf_{\substack{i=0,\dots,r-1\\ j=1,\dots,s}}|is-jr|\,\frac{r+s}{rs} \ge \frac{1}{s}+\frac{1}{r} > 0.
\end{aligned} \tag{21}
$$

Condition (C3) implies that $\max_{i=0,\dots,m}|t_{1,\,i}-i/m| = o((n\vee m))^{-1}$ and $\max_{j=0,\dots,n}|t_{2,\,j}-j/n| = o((n\vee m))^{-1}$. This proves, together with (21), statement (20). Define now $\tilde\lambda_{ij}$ as the weight λ_{ij} in the case of equidistant designs, that is, when $t_{1,\,i}=i/m$, $t_{2,\,j}=j/n$. Now we may estimate for m, n such that $m \ge M$, $n \ge N$ as follows. Observe first that

$$
\begin{aligned}
&\#\{(i,j)\colon \lambda_{ij}^2 \neq 0\} \le n+m,\\
&\#\{(i,j)\colon \tilde\lambda_{ij}^2 \neq 0\} \le n+m
\end{aligned}
$$

and let $\{1,\dots,m\}\times\{1,\dots,n\} = I_{m,\,n}$. We have for sufficiently large m, n,

$$
\begin{aligned}
\left|\sum_{i,\,j=1}^{n,\,m}(\lambda_{ij}^2-\tilde\lambda_{ij}^2)\right| \;&\le\; (n+m)\sup_{(i,\,j)\in I_{m,\,n}}|\lambda_{ij}^2-\tilde\lambda_{ij}^2|\\
&=\; (n+m)\sup_{i,\,j\in I_{m,\,n}}\left|\left(\int_0^1 \mathbf{I}_{[i/m,\,(i+1)/m)}(t)\mathbf{I}_{[j/n,\,(j+1)/n)}(t)\,dt\right)^2\right.\\
&\qquad\left.-\left(\int_0^1 \mathbf{I}_{[t_{1,\,i},\,t_{1,\,i+1})}(t)\mathbf{I}_{[t_{2,\,j},\,t_{2,\,j+1})}(t)\,dt\right)^2\right|
\end{aligned}
$$

$$
\begin{aligned}
&\le\ (n+m)\sup_{i,\,j\in I_{m,n}}\left|\int_0^1\Big(\mathbf{I}_{[t_{1,i},\,t_{1,i+1})}(t)\mathbf{I}_{[t_{2,j},\,t_{2,j+1})}(t)\right.\\
&\qquad\qquad\left.-\,\mathbf{I}_{[i/m,\,(i+1)/m)}(t)\mathbf{I}_{[j/n,\,(j+1)/n)}(t)\Big)\,dt\right|O((m\vee n)^{-1})\\
&\stackrel{(20)}{=}\ \sup_{i,\,j\in I_{m,n}}\Big|\big(t_{1,i+1}\wedge t_{2,j+1}-t_{1,i}\vee t_{2,j}\big)\\
&\qquad\qquad\times\mathbf{I}_{\{t_{1,i+1}\wedge t_{2,j+1}>t_{1,i}\vee t_{2,j}\}}-\big((i+1)/m\wedge(j+1)/n-i/m\vee j/n\big)\\
&\qquad\qquad\qquad\times\mathbf{I}_{\{(i+1)/m\wedge(j+1)/n>i/m\vee j/n\}}\Big|O(1)\\
&=\ o((n\vee m)^{-1}).
\end{aligned}
$$

This proves the assertion. □

LEMMA A2. *Assume* (C1)–(C3) *and let* $f_i=f(t_{1,i})$, $i=1,\dots,m$, *and* $g_j=g(t_{2,j})$, $j=1,\dots,n$. *Then we have*

$$
\lim_{n,m\to\infty}(n+m)\sum_i^m\Big(\sum_j^n\lambda_{ij}g_j\Big)^2=\kappa^{-1}\int_0^1 g^2(t)\,dt,
$$

$$
\lim_{n,m\to\infty}(n+m)\sum_j^n\Big(\sum_i^m\lambda_{ij}f_i\Big)^2=(1-\kappa)^{-1}\int_0^1 f^2(t)\,dt.
$$

PROOF. Define step functions $f_{\langle m\rangle}$ and $g_{\langle n\rangle}$ as

$$
f_{\langle m\rangle}(t)=\sum_{i=0}^m f(t_{1,i})\mathbf{I}_{[t_{1,i},\,t_{1,i+1})}(t),\ g_{\langle n\rangle}(t)=\sum_{j=0}^n g(t_{2,j})\mathbf{I}_{[t_{2,j},\,t_{2,j+1})}(t). \tag{22}
$$

We have

$$
\begin{aligned}
m\sum_i^m\Big(\sum_j^n\lambda_{ij}g_j\Big)^2&=m\sum_{i=1}^m\Big(\int_{t_{1,i}}^{t_{1,i+1}}g_{\langle n\rangle}(t)\,dt\Big)^2=m\sum_{i=1}^m\Delta_{1,i}^2 g_{\langle n\rangle}^2(\xi_i)\\
&=\int_0^1 g^2(t)\,dt+o(1),
\end{aligned}
$$

where $\xi_i\in[t_{1,i},t_{1,i+1})$ $i=1,\dots,m$, and the last equality follows by Lebesgue's convergence theorem. The proof of the second statement is similar and therefore omitted. □

PROOF OF THEOREM 1. We assume for the moment that the designs are equidistant, that is, $\Delta_{1,i}=1/m$, $\Delta_{2,j}=1/n$ and let $m/n=\lambda$. Observe that

$$
\sum_{j=1}^n\lambda_{ij}=\Delta_{1,i}=1/m,\ \ i=1,\dots,m,\qquad \sum_{i=1}^m\lambda_{ij}=\Delta_{2,j}=1/n,\ \ j=1,\dots,n.
$$

We have

$$
\begin{aligned}
&\text{(I)} && \hat{M}^2 - M^2 = \sum_i \Delta_{1,i}\,(f_i+\varepsilon_i)(f_{i+1}+\varepsilon_{i+1}) - \int_0^1 f^2(t)\,dt\\
&\text{(II)} && \qquad + \sum_j \Delta_{2,j}\,(g_j+\eta_j)(g_{j+1}+\eta_{j+1}) - \int_0^1 g^2(t)\,dt\\
&\text{(III)} && \qquad - \sum_{i,j} \lambda_{ij}(f_{i+1}+\varepsilon_{i+1})(g_j+\eta_j) + \int_0^1 f(t)g(t)\,dt\\
&\text{(IV)} && \qquad - \sum_{i,j} \lambda_{ij}(f_i+\varepsilon_i)(g_{j+1}+\eta_{j+1}) + \int_0^1 f(t)g(t)\,dt.
\end{aligned}
$$

The fact that

$$\lambda_{ij} \neq 0 \quad \text{if and only if } \frac{i}{\lambda} - 1 < j < \frac{i+1}{\lambda}$$

and Lemma 1 allow us to rewrite asymptotically $\hat{M}^2 - M^2$ as a sum of centered random variables

$$\hat{M}^2 - M^2 = \sum_{i=2}^m \Psi_i + E[\hat{M}^2 - M^2] = \sum_{i=2}^m \Psi_i + o_p(n^{-1/2}),$$

where

$$
\begin{aligned}
\Psi_i := {} & \Delta_{1,i}(f_i\varepsilon_{i-1} + f_{i-1}\varepsilon_i + \varepsilon_i\varepsilon_{i-1})\\
& + \sum_{j\in J_i} \Delta_{2,j}(g_j\eta_{j-1} + g_{j-1}\eta_j + \eta_j\eta_{j-1})\\
& - \sum_{j\in J_i} \lambda_{ij}(f_{i+1}\eta_j + \varepsilon_{i+1}g_j + \varepsilon_{i+1}\eta_j)\\
& - \sum_{j\in J_i} \lambda_{ij}(f_i\eta_{j+1} + \varepsilon_i g_{j+1} + \varepsilon_i\eta_{j+1})
\end{aligned}
$$

and

$$J_i := \left\{ j\colon \frac{i}{\lambda} - 1 < j < \frac{i+1}{\lambda} \right\}, \qquad i = 2, \dots, m.$$

We find that $J_i \neq \varnothing$ $(i = 2, \dots, m)$,

$$\#(J_i \cap J_{i+1}) = \begin{cases} 1, & \dfrac{i+1}{\lambda} \notin \mathbb{N}, \\[2mm] 0, & \dfrac{i+1}{\lambda} \in \mathbb{N}, \qquad i = 2, \dots, m-1, \end{cases}$$

and that

$$J_i \cap J_{i+l} = \varnothing, \qquad l \geq 2.$$

Hence $(\Psi_i)_{i=1,\dots,m}$ constitutes an array of rowwise 3-dependent centered random variables. Now the assertion follows from Orey's (1958) central limit theorem for k-dependent random variables, provided the limiting variance of $(n+m)^{1/2}(\hat{M}^2 - M^2)$ exists. To show this existence we calculate in a first step the variances of expressions I–IV:

$$\begin{aligned}
V[\mathrm{I}] &= V\Bigl[\sum_i \Delta_{1,\,i}(\varepsilon_i\varepsilon_{i+1} + f_{i+1}\varepsilon_i + f_i\varepsilon_{i+1})\Bigr] + o(n^{-1})\\
&= V\Bigl[\sum_i L_{1,\,i}\Bigr] + o(n^{-1})\\
&= \sum_i V[L_{1,\,i}] + o(n^{-1})\\
&= \sum_i \Delta_{1,\,i}^2(\sigma_\varepsilon^4 + 4\sigma_\varepsilon^2 f_i^2) + o(n^{-1}),
\end{aligned}$$

where we used the notation $L_{1,\,i} = \Delta_{1,\,i}(\varepsilon_i\varepsilon_{i+1} + 2\varepsilon_i f_i)$ and the fact that $\mathrm{Cov}[L_{1,\,i}, L_{1,\,j}] = 0$, $i \neq j$. Similarly, we find that

$$V[\mathrm{II}] = \sum_j \Delta_{2,\,j}^2(\sigma_\eta^4 + 4\sigma_\eta^2 g_j^2) + o(n^{-1}).$$

Let $H_{ij} = \varepsilon_{i+1}g_j + \varepsilon_{i+1}\eta_j + f_{i+1}\eta_j$, then

$$\begin{aligned}
V[\mathrm{III}] &= V\Bigl[\sum_{i,\,j}\lambda_{ij}H_{ij}\Bigr]\\
&= V\Bigl[\sum_{i,\,j}\lambda_{ij}\varepsilon_{i+1}\eta_j\Bigr] + V\Bigl[\sum_{i,\,j}\lambda_{ij}\varepsilon_{i+1}g_j\Bigr] + V\Bigl[\sum_{i,\,j}\lambda_{ij}f_{i+1}\eta_j\Bigr] + o(n^{-1})
\end{aligned}$$

because $\mathrm{Cov}[\varepsilon_i\eta_j, \varepsilon_i g_j] = \mathrm{Cov}[\varepsilon_i\eta_j, f_i\eta_j] = \mathrm{Cov}[\varepsilon_i g_j, f_i\eta_j] = 0$. Finally, we obtain

$$V[\mathrm{III}] = \sigma_\varepsilon^2\sigma_\eta^2\sum_{i,j}\lambda_{ij}^2 + \sigma_\varepsilon^2\sum_i\Bigl(\sum_j\lambda_{ij}g_j\Bigr)^2 + \sigma_\eta^2\sum_j\Bigl(\sum_i\lambda_{ij}f_i\Bigr)^2.$$

Similarly,

$$V[\mathrm{IV}] = V[\mathrm{III}] + o(n^{-1}).$$

It remains to calculate the covariances of expressions I–IV. We have

$$\begin{aligned}
\mathrm{Cov}[\mathrm{I},\mathrm{II}] &= 0,\\
2\,\mathrm{Cov}[\mathrm{I},\mathrm{III}] &= -2\sum_i\sum_{l,\,k}\lambda_{lk}\,\mathrm{Cov}[L_i, H_{lk}]\\
&= -4\sigma_\varepsilon^2\sum_{i,\,k}\Delta_{1,\,i}\lambda_{ik}f_i g_k + o(n^{-1})
\end{aligned}$$

and

$$2\,\mathrm{Cov}[\mathrm{I}, \mathrm{IV}] = 2\,\mathrm{Cov}[\mathrm{I}, \mathrm{III}] + o(n^{-1}).$$

Similarly,

$$2\,\mathrm{Cov}[\mathrm{II}, \mathrm{III}] = -4\sigma_\eta^2 \sum_{k,\,j} \Delta_{2,\,j} \lambda_{kj} f_k g_j + o(n^{-1})$$

and

$$2\,\mathrm{Cov}[\mathrm{II}, \mathrm{IV}] = 2\,\mathrm{Cov}[\mathrm{II}, \mathrm{III}] + o(n^{-1}).$$

Finally,

$$\begin{aligned} 2\,\mathrm{Cov}[\mathrm{III}, \mathrm{IV}] &= 2 \sum_{i,\,j,\,r,\,s} \lambda_{ij}\lambda_{rs}\,\mathrm{Cov}\big[\varepsilon_{i+1} g_j + \varepsilon_{i+1}\eta_j + f_{i+1}\eta_j, \\ &\qquad\qquad \varepsilon_r g_{s+1} + \varepsilon_r \eta_{s+1} + f_r \eta_{s+1}\big] \\ &= 2\,\sigma_\varepsilon^2 \sum_{i,\,j,\,s} \lambda_{ij}\lambda_{i+1,\,s} g_j g_{s+1} + 2\,\sigma_\eta^2 \sum_{i,\,r,\,s} \lambda_{i,\,s+1}\lambda_{rs} f_{i+1} f_r \\ &\quad + 2\,\sigma_\varepsilon^2\sigma_\eta^2 \sum_{i,\,s} \lambda_{i,\,s+1}\lambda_{i+1,\,s}. \end{aligned} \tag{23}$$

The definition of λ_{ij} in (4) shows that

$$\lambda_{i,\,s+1}\lambda_{i+1,\,s} = 0, \qquad i = 1, \dots, m-1,\; s = 1, \dots, n-1,$$

and hence the last expression in (23) vanishes. We also obtain for $n, m \to \infty$,

$$(n+m)\,V[\mathrm{I}] \to \kappa^{-1}\left(\sigma_\varepsilon^4 + 4\sigma_\varepsilon^2 \int_0^1 f^2(t)\,dt\right),$$

$$(n+m)\,V[\mathrm{II}] \to (1-\kappa)^{-1}\left(\sigma_\eta^4 + 4\sigma_\eta^2 \int_0^1 g^2(t)\,dt\right),$$

$$(n+m)\,2\,(\mathrm{Cov}[\mathrm{I}, \mathrm{III}] + \mathrm{Cov}[\mathrm{I}, \mathrm{IV}]) \to -8\kappa^{-1}\sigma_\varepsilon^2 \int_0^1 f(t)g(t)\,dt,$$

$$(n+m)\,2\,(\mathrm{Cov}[\mathrm{II}, \mathrm{III}] + \mathrm{Cov}[\mathrm{II}, \mathrm{IV}]) \to -8(1-\kappa)^{-1}\sigma_\eta^2 \int_0^1 f(t)g(t)\,dt$$

and from Lemmas A1 and A2,

$$\begin{aligned} (n+m)(V[\mathrm{III}] + V[\mathrm{IV}]) \to\ & 2\sigma_\varepsilon^2\sigma_\eta^2 \iota(r,s) + 2\sigma_\varepsilon^2\kappa^{-1} \int_0^1 g^2(t)\,dt \\ & + 2\sigma_\eta^2(1-\kappa)^{-1} \int_0^1 f^2(t)\,dt. \end{aligned}$$

Finally, we have for $n, m \to \infty$,

$$\begin{aligned} (n+m)\,2\,\mathrm{Cov}[\mathrm{III}, \mathrm{IV}] \to\ & 2\sigma_\varepsilon^2\kappa^{-1} \int_0^1 g^2(t)\,dt \\ & + 2\sigma_\eta^2(1-\kappa)^{-1} \int_0^1 f^2(t)\,dt \end{aligned} \tag{24}$$

as $m, n \to \infty$, which can be seen as follows. We obtain from (23) that

$$2\,\mathrm{Cov}[\mathrm{III}, \mathrm{IV}] = 2\sigma_\varepsilon^2 \sum_i \int_{t_{1,i}}^{t_{1,i+1}} g_{\langle n\rangle}(t)\,dt \int_{t_{1,i+1}}^{t_{1,i+2}} \tilde{g}_{\langle n\rangle}(t)\,dt$$
$$+ 2\sigma_\eta^2 \sum_s \int_{t_{2,s}}^{t_{2,s+1}} f_{\langle m\rangle}(t)\,dt \int_{t_{2,s+1}}^{t_{2,s+2}} \tilde{f}_{\langle m\rangle}(t)\,dt,$$

where $f_{\langle m\rangle}, g_{\langle n\rangle}$ were defined in (22) and

$$\tilde{g}_{\langle n\rangle}(t) = \sum_{j=0}^{n} g(t_{2,j+1})\,\mathbf{I}_{[t_{2,j},\,t_{2,j+1})}(t), \qquad \tilde{f}_{\langle m\rangle}(t) = \sum_{i=0}^{m} f(t_{1,i+1})\,\mathbf{I}_{[t_{1,i},\,t_{1,i+1})}(t).$$

Now by assumptions (C1) and (C3),

$$\sup_{t\in[0,1]} |(\tilde{g}_{\langle n\rangle} - g_{\langle n\rangle})(t)| = O(n^{-\gamma})$$

and hence

$$\sup_{i=0,\dots,m-1} \left| \int_{t_{1,i}}^{t_{1,i+1}} (\tilde{g}_{\langle n\rangle} - g_{\langle n\rangle})(t)\,dt \right| = O(n^{-1-\gamma}).$$

This gives

$$2\,\mathrm{Cov}[\mathrm{III}, \mathrm{IV}] = 2\sigma_\varepsilon^2 \sum_i \int_{t_{1,i}}^{t_{1,i+1}} g_{\langle n\rangle}(t)\,dt \int_{t_{1,i+1}}^{t_{1,i+2}} g_{\langle n\rangle}(t)\,dt$$
$$+ 2\sigma_\eta^2 \sum_j \int_{t_{2,j}}^{t_{2,j+1}} f_{\langle m\rangle}(t)\,dt \int_{t_{2,j+1}}^{t_{2,j+2}} f_{\langle m\rangle}(t)\,dt + O(n^{-1-\gamma}).$$

Furthermore, for $i = 1, \dots, m-1$ there exist $\xi_i \in [t_{1,i}, t_{1,i+1}]$, such that

$$\left| \int_{t_{1,i+1}}^{t_{1,i+2}} g_{\langle n\rangle}(t)\,dt - \int_{t_{1,i}}^{t_{1,i+1}} g_{\langle n\rangle}(t)\,dt \right| = \left| g_{\langle n\rangle}(\xi_{i+1})\Delta_{1,i+2} - g_{\langle n\rangle}(\xi_i)\Delta_{1,i+1} \right|$$
$$= \frac{1}{m}\left| g(t_{1,i+1}) - g(t_{1,i}) \right| = o((m \vee n)^{-1}),$$

where we used the Hölder continuity of g for the last equality. A similar argument applies for the second term, which gives

$$2\,\mathrm{Cov}[\mathrm{III}, \mathrm{IV}] = 2\sigma_\varepsilon^2 \sum_{i=0}^{m-1} \left(\int_{t_{1,i}}^{t_{1,i+1}} g_{\langle n\rangle}(t)\,dt \right)^2$$
$$+ 2\sigma_\eta^2 \sum_{j=0}^{n-1} \left(\int_{t_{2,j}}^{t_{2,j+1}} f_{\langle m\rangle}(t) \right)^2 + o(n^{-1}).$$

An application of Lemma A2 yields now (24). Adding all variances and covariances gives the required variance and completes the proof of Theorem 1 for equidistant designs.

The proof for asymptotic uniform designs follows exactly the same pattern, where similar arguments as in Lemma A1 have to be used. In this case, the asymptotic variances of III and IV are a consequence of Lemma A2. □

LEMMA A3. *Assume the set-up of the k sample problem in Theorem* 3 *of Section* 3.3. *Under the assumptions* (C1)–(C3) *we have for* $j \neq k$, $i \neq j, k$,

$$\begin{aligned}\lim_{N\to\infty} N\,\mathrm{Cov}[\hat{M}^2(f_i, f_j), \hat{M}^2(f_i, f_k)] &= \kappa_i^{-1}\biggl\{\sigma_i^4 + 4\sigma_i^2\biggl[\int_0^1 f_i^2(t)\,dt - \int_0^1 f_i(t)f_j(t)\,dt \\ &\qquad - \int_0^1 f_i(t)f_k(t)\,dt + \int_0^1 f_j(t)f_k(t)\,dt\biggr]\biggr\} \\ &= \kappa_i^{-1}\biggl\{\sigma_i^4 + 4\sigma_i^2\int_0^1 (f_i(t)-f_j(t))(f_i(t)-f_k(t))\,dt\biggr\} \\ &= \kappa_i^{-1}\zeta_{j,k}^{(i)}.\end{aligned}$$

PROOF. If we denote the expressions in (I)–(III) at the beginning of the proof of Theorem 1 by $\int \hat{f}^2$, $\int \hat{g}^2$ and $\int \hat{f g}$, we obtain

$$\begin{aligned}
&\text{(a)} && N\,\mathrm{Cov}\bigl[\hat{M}^2(f_i, f_j), \hat{M}^2(f_i, f_k)\bigr] = N\,\mathrm{Cov}\biggl[\int \hat{f}_i^2, \int \hat{f}_i^2\biggr] \\
&\text{(b)} && \qquad + N\,\mathrm{Cov}\biggl[\int \hat{f}_i^2, -2\int \widehat{f_i f_j}\biggr] \\
&\text{(c)} && \qquad + N\,\mathrm{Cov}\biggl[\int \hat{f}_i^2, -2\int \widehat{f_i f_k}\biggr] \\
&\text{(d)} && \qquad + N\,\mathrm{Cov}\biggl[-2\int \widehat{f_i f_j}, -2\int \widehat{f_i f_k}\biggr].
\end{aligned}$$

The expressions (a)–(c) are treated similarly to the proof of Theorem 1 and give the corresponding terms in $\zeta_{j,k}^{(i)}$. The last term (d) is evaluated as

$$\begin{aligned}N\,\mathrm{Cov}\biggl[-2\int \widehat{f_i f_j}, -2\int \widehat{f_i f_k}\biggr] &= 4N\,\sigma_i^2 \sum_{l_i=1}^{n_i} \int_{t_{i,l_i}}^{t_{i,l_i+1}} f_{\langle j\rangle}(t)\,dt \int_{t_{i,l_i}}^{t_{i,l_i+1}} f_{\langle k\rangle}(t)\,dt \\ &= 4N\,\sigma_i^2 \sum_{l_i=1}^{n_i} \Delta_{i,l_i}^2 f_{\langle j\rangle}(\xi_{i,l_i}) f_{\langle k\rangle}(\xi_{i,l_i}),\end{aligned}$$

where $\xi_{i,l_i} \in [t_{i,l_i}, t_{i,l_i+1})$, $l_i = 1, \ldots, n_i$, $i = 1, \ldots, k$. By a generalization of Lemma A2 this converges to $4\kappa_i^{-1}\sigma_i^2 \int f_j(t)f_k(t)\,dt$. □

PROOF OF THEOREM 3. Taking into account Lemma A3 and the proof of Theorem 1, we obtain

$$
\begin{aligned}
\lim_{N\to\infty} V\Big[N^{1/2}\sum_{i<j}\hat{M}^2_{ij}\Big] &= \sum_{i<j}\xi^2_{ij} + \lim_{N\to\infty} N \sum_{\substack{i<j,\, s<t\\ (i,\,j)\neq(s,\,t)}} \operatorname{Cov}\big[\hat{M}^2_{ij},\ \hat{M}^2_{st}\big]\\
&= \sum_{i<j}\xi^2_{ij} + 2\sum_{i<j<l}\kappa_i^{-1}\zeta^{(i)}_{j,l} + 2\sum_{j<l<i}\kappa_i^{-1}\zeta^{(i)}_{j,l}\\
&\quad + 2\sum_{l<i<j}\kappa_i^{-1}\zeta^{(i)}_{j,l}\\
&= \sum_{i<j}\xi^2_{ij} + \sum_{i\neq j,\, i\neq l,\, j\neq l}\kappa_i^{-1}\zeta^{(i)}_{j,l},
\end{aligned}
\tag{25}
$$

where

$$
\zeta^{(i)}_{j,l} = \sigma_i^4 + 4\sigma_i^2\int_0^1 (f_i(t)-f_j(t))(f_i(t)-f_l(t))\,dt
$$

and

$$
\xi^2_{ij} = \kappa_i^{-1}(\sigma_i^4 + 4\sigma_i^2 M^2(f_i, f_j)) + \kappa_j^{-1}\big(\sigma_j^4 + 4\sigma_j^2 M^2(f_i, f_j)\big) + 2\sigma_i^2\sigma_j^2\iota(r_i, r_j).
$$

The representation of ξ_k^2 in Theorem 3 now follows by straightforward algebra. For example, we obtain for the factor appearing with $\sum_{i=1}^k \kappa_i^{-1}\sigma_i^2$ from (25),

$$
\begin{aligned}
&\sum_{i<j}\big(\kappa_i^{-1}\sigma_i^2 + \kappa_j^{-1}\sigma_j^2\big)M^2(f_i, f_j)\\
&\quad + \sum_{i=1}^k \kappa_i^{-1}\sigma_i^2 \sum_{j\neq i}\ \sum_{l\neq i,\, l\neq j}\int_0^1 (f_i(t)-f_j(t))(f_i(t)-f_l(t))\,dt\\
&\quad = \sum_{i=1}^k \kappa_i^{-1}\sigma_i^2 \sum_{j=1}^k\sum_{l=1}^k \int_0^1 (f_i(t)-f_j(t))(f_i(t)-f_l(t))\,dt\\
&\quad = k^2\sum_{i=1}^k \kappa_i^{-1}\sigma_i^2\int_0^1 (f_i(t)-\bar{f}(t))^2\,dt.
\end{aligned}
$$

The other terms are treated similarly. The proof of the asymptotic normality is essentially the same as in the two-sample case and therefore omitted. □

PROOF OF THEOREM 4. We will only give a sketch of the proof, because it follows exactly the same pattern as the proof of Theorem 1. Asymptotic normality follows again from the central limit theorem for m-dependent random variables. A simple calculation shows that

$$
E[\hat{M}^2] = M^2 + o(n^{-1/2}).
$$

Observe that the decomposition I–IV as in the proof of Theorem 1 still holds true with λ_{ij} as in (4). For example, we obtain, for the limiting variance of expression I,

$$\begin{aligned} V[\sqrt{2n}\,\mathrm{I}] &= 2n \sum_i \Delta_{1,i}^2 \big(\sigma_\varepsilon^4 + 4\sigma_\varepsilon^2 f_i^2\big) + o(1) \\ &= 2\int_0^1 \left(\frac{d}{ds} H^{-1}(s)\right)^2 \big(\sigma_\varepsilon^4 + 4\sigma_\varepsilon^2 f^2(s)\big)\, ds + o(1) \\ &= 2\left\{\sigma_\varepsilon^4 \int_0^1 \frac{1}{h(s)}\, ds + 4\sigma_\varepsilon^2 \int_0^1 \left(\frac{1}{h\circ H^{-1}(s)}\right)^2 f^2(s)\, ds\right\} + o(1). \end{aligned}$$

Expression II is treated similarly. Moreover, the first expressions of $V[\mathrm{III}]$ and $V[\mathrm{IV}]$, respectively, reduce to

$$2n \sum_{i,j} \lambda_{ij}^2 = 2n \sum_{i=1}^n \lambda_{ii}^2 = 2n \sum_{i=1}^n \Delta_{1,i}^2 = 2\int_0^1 \left(\frac{d}{ds} H^{-1}(s)\right)^2 ds + o(1),$$

which proves the assertion. □

Acknowledgments. We thank J. Hart for providing us with the rainfall data, M. Sahm for helpful discussions and I. Gottschlich for typing parts of this manuscript. Suggestions and comments of an Associate Editor and a referee are gratefully acknowledged. Parts of this paper were written when A. Munk was visiting the Statistics Department at the Wharton School of Business, and this author thanks L. D. Brown and the Wharton School for their hospitality.

REFERENCES

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

BUCKLEY, M. J., EAGLESON, G. K. and SILVERMAN, B. W. (1988). The estimation of residual variance in nonparametric regression. *Biometrika* **75** 189–199.

CARTER, C. K. and EAGLESON, G. K. (1992). A comparison of variance estimators in nonparametric regression. *J. Roy. Statist. Soc. Ser. B* **54** 773–780.

COX, D., KOH, G., WAHBA, G. and YANDELL, B. S. (1988). Testing the (parametric) null model hypothesis in (semiparametric) partial and generalized spline models. *Ann. Statist.* **16** 113–119.

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

DETTE, H., MUNK, A. and WAGNER, T. (1997). A review of variance estimators with applications to multivariate nonparametric regression models. In *Multivariate Design and Sampling* (S. Ghosh, ed.) Dekker, New York. To appear.

DETTE, H., MUNK, A. and WAGNER, T. (1998). Estimating the variance in nonparametric regression by quadratic forms—what is a reasonable choice? *J. Roy. Statist. Soc. Ser. B* **60** 751–764.

EUBANK, R. L. and SPIEGELMAN, C. H. (1990). Testing the goodness of fit of a linear model via nonparametric regression techniques. *J. Amer. Statist. Assoc.* **85** 387–392.

GASSER, T., SROKA, L. and JENNEN-STEINMETZ, C. (1986). Residual variance and residual pattern in nonlinear regression. *Biometrika* **73** 625–633.

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., KAY, J. W. and TITTERINGTON, D. M. (1990). Asymptotically optimal difference-based estimation of variance in nonparametric regression. *Biometrika* **77** 521–528.

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** 83–89.

HOCKING, R. R. (1985). *The Analysis of Linear Models*. Brooks/Cole, Monterey, CA.

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.

OREY, S. (1958). A central limit theorem for m-dependent random variables. *Duke Math. J.* **52** 543–546.

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

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

ULLAH, A. and ZINDE-WALSH, V. (1992). On the estimation of residual variance in nonparametric regression. *J. Nonparametr. Statist.* **1** 263–265.

FAKULTÄT UND INSTITUT FÜR MATHEMATIK
RUHR-UNIVERSITÄT BOCHUM
UNIVERSITÄTSSTR. 150, GEBÄUDE NA III
44780 BOCHUM
GERMANY
E-MAIL: holger.dette@ruhr-uni-bochum.de
axel.munk@ruhr-uni-bochum.de