The Annals of Statistics
1986, Vol. 14, No. 3, 1152–1170

ASYMPTOTIC BEHAVIOR OF THE EMPIRIC DISTRIBUTION OF M-ESTIMATED RESIDUALS FROM A REGRESSION MODEL WITH MANY PARAMETERS[1]

BY STEPHEN PORTNOY

University of Illinois

Consider a regression model $Y_i = x_i'\beta + R_i$, $i = 1, \dots, n$, where $\{R_i\}$ are i.i.d. with c.d.f., F; $x_i \in R^p$ and $\beta \in R^p$. Let $\hat{\beta}$ be a M-estimator defined using kernel, ψ; let $\hat{F}_n(x)$ denote the empiric distribution of the residuals, $Y_i - x_i'\hat{\beta}$, and let $\hat{F}_n^*$ be the empiric c.d.f. of the errors, $\{R_i\}$. Under suitable smoothness conditions on ψ, F, and the density $F' = f$ and conditions requiring essentially that $\{x_i\}$ behave like a random sample from some distribution in R^p, it is shown that, for fixed x,

$$\sqrt{n}\left(\hat{F}_n(x) - \hat{F}_n^*(x) - H_n(x)\right) - \frac{p}{\sqrt{n}} g(x) \to_p 0,$$

where $g(x) = af(x)\psi(x) + bf'(x)$ and $H_n(x) = (1/nd)f(x)\sum_{i=1}^n \psi(R_i)$ if the design has a constant term [and $H_n(x)$ vanishes otherwise]. A tightness result shows that if $p/\sqrt{n} \to c$, $\sqrt{n}(\hat{F}_n(x) - F(x))$ converges weakly to a Gaussian process with drift given by the bias term $cg(x)$, and covariance function strongly affected by $H_n(x)$ and different from that for the usual Brownian bridge. In the course of the proof, an expansion for the fitted values, $x_i'\hat{\beta}$, is obtained, with error $O_p(p^{11/4} \ln^2 n/n^2) = o_p(1/\sqrt{n})$ if p^2/n is bounded.

1. Introduction. The use of residuals in analyzing linear (and nonlinear) models has become extremely widespread. The basic results here concern the asymptotic behavior of residuals from regression models when M-estimators are used and the number of parameters is permitted to grow with the sample size. To be precise, consider the general linear model

$$Y_i = x_i'\beta + R_i, \qquad i = 1, \dots, n, \tag{1.1}$$

where $\{R_1, \dots, R_n\}$ are i.i.d. with c.d.f. F, $\{x_1, \dots, x_n\}$ are (fixed) vectors in R^p and $\beta \in R^p$. Let ψ be a given kernel and define the M-estimator, $\hat{\beta}$, to be any solution of the vector equation

$$0 = \sum_{i=1}^n x_i \psi(Y_i - x_i'\beta). \tag{1.2}$$

For the results here, we assume $\beta = 0$ without loss of generality. Define $\hat{F}_n(x)$ to

Received October 1984; revised December 1985.
[1]Research partially supported by NSF Grants MCS 83-01834 and DMS 85-03785.
AMS 1980 *subject classifications*. Primary 62G35; secondary 62E20, 62J05.
Key words and phrases. Regression, residuals, M-estimators, empiric c.d.f., asymptotics.

be the empiric c.d.f. of the residuals:

$$\hat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} I(Y_i - x_i'\hat{\beta} \le x) = \frac{1}{n}\sum_{i=1}^{n} I(R_i - x_i'\hat{\beta} \le x) \tag{1.3}$$

(since $\beta = 0$), where $I(\cdot)$ is the indicator function of its argument. Let $\hat{F}_n^*(x)$ be the empiric c.d.f. of the errors, $\{R_i\}$.

If F is normal and $\psi(u) = u$ (least squares), the joint distribution of residuals is well known. In other cases, asymptotic computations will generally be necessary. If p is fixed as $n \to \infty$, classical methods can be used. However, in most applications, if n is large, models with large p will be considered. For example, in a regression model with five independent variables, $n = 100$ might be considered adequate for reasonable asymptotic approximation. However, in such cases quadratic models are often considered, providing a model with $p = 21$ parameters; and, thus p^2/n is moderately large. The basic result presented here depends on whether or not the design has a constant term; that is, the first coordinate of x_i satisfies

$$x_{i1} = 1, \quad \text{for } i = 1, 2, \ldots, n. \tag{1.4}$$

The result (Theorem 3.1) is the following: let $f(x)$ be the density of R, $d = E\psi'(R)$ and $\sigma^2 = \operatorname{Var}\psi(R)$. Then

$$\sqrt{n}\,(\hat{F}_n(x) - \hat{F}_n^*(x) - H_n(x)) - \frac{p}{\sqrt{n}}g(x) \to_p 0,$$

where

$$g(x) = \frac{\sigma^2}{2d^2}f'(x) + \frac{1}{2d}f(x)\psi(x)$$

and

$$H_n(x) = \frac{1}{nd}f(x)\sum_{i=1}^{n}\psi(R_i),$$

if (1.4) holds and $H_n(x)$ vanishes otherwise. The $H_n(x)$ term arises from the estimation of the coefficient of the constant term. When p is fixed, it was considered in a regression setting by Koul (1969) and Pierce and Kopecky (1979), and in more general situations by Burke, Csörgő, Csörgő, and Révész (1979), Loynes (1980), and Shorack (1985) where its strong effect on the asymptotic distribution is discussed. If p is fixed and $f(x)$ is known (e.g., in testing a simple null hypothesis), $H_n(x)$ can be appropriately estimated since $d\hat{\beta}_1 = (1/n)\Sigma\psi(R_i) + o(1/\sqrt{n})$. Thus, by adjusting for $H_n(x)$ it is possible to construct a process $\sqrt{n}\,(\hat{F}_n - F - \hat{H}_n)$, which converges to the usual transformed Brownian bridge. It appears that this adjustment may be possible even if $p \to \infty$. If $f(x)$ cannot be assumed, it too must be estimated [to order $o(1/\sqrt{n})$], and this is a very difficult problem, particularly if $p \to \infty$. It should be noted that it is relatively easy to adjust for the bias term $g(x)$ since all that is required is consistent estimators of $f(x)$ and $f'(x)$ (which can be obtained under the conditions used here).

These results depend upon the asymptotic behavior of the M-estimators, $\hat{\beta}$, which has been considered by Huber (1973, 1981), Yohai and Maronna (1979), and more recently by Portnoy (1984, 1985a). The first two references use more or less classical results, but require a stronger condition than $p^2/n \to 0$. The last references show that in the regression case (where $\{x_i\}$ "act" like a sample from some distribution in R^p), $\|\hat{\beta}\|^2 = O_p(p/n)$, and $\max_i |x_i'\hat{\beta}| \to_P 0$ if p is sufficiently small compared to n. The author (1985a) shows that $p^{3/2} \ln n/n \to 0$ is sufficient but conjectures that $p^{1+\varepsilon}/n \to 0$ should work. In fact Theorem 2.1 here extends the earlier result slightly obtaining a higher-order expansion of $x_i'\hat{\beta}$ (with four additional terms) with error of order $(p^{11/8}(\ln n)/n)^2$. This result clearly indicates the difficulty of using such expansions to try to verify the conjecture.

Consistency results for $\hat{F}_n$ were considered by Freedman (1981) and Bickel and Freedman (1983) in the case of least-squares estimators, and by Shorack (1982) in the case of more general M-estimators. These results were presented in the context of showing that the Bootstrap method based on residuals is consistent. Bickel and Freedman show that if $p/n \to 0$ then the Mallows distance between $\hat{F}_n$ and F tends to zero (in the "least-squares" case). They use this result to prove consistency of the Bootstrap distribution of a fixed contrast if $p/n \to 0$ and of the Bootstrap distribution of $\hat{\beta}^*$ (in R^p) if $p^2/n \to 0$. Shorack (1982) shows consistency of the Bootstrap distribution of a contrast in the M-estimator case if $p^2/n \to 0$.

2. An expansion for $x_i'\hat{\beta}$. The main result, Theorem 3.1, requires an expansion of $x_i'\hat{\beta}$ in terms of sums of functions $\{R_i\}$ with an error $o_p(1/\sqrt{n})$ if $p^2/n = O(1)$. Theorem 3.1 of Portnoy (1985a) presents such an expansion with error terms involving $x_i'\hat{\beta}$ and which are shown to be $O_p(p^{3/2}(\ln n)^{5/4}/n)$ (which clearly is not sufficient). Theorem 2.1 provides an adequate expansion with appropriate error terms.

Since results from Portnoy (1985a) will be used, some of the conditions of that paper will be required and some further notation is needed. The conditions in Portnoy (1985a) relating p and n are weaker than the condition $p^2/n = O(1)$ used here. The conditions on ψ include symmetry conditions (also required on the distribution of R) and three bounded continuous derivatives—a fourth bounded derivative is needed here. As noted earlier, conditions on $\{x_i\}$ are somewhat artificial since they are designed to hold only in typical regression cases where $\{x_i\}$ can be considered as a sample from some distribution in R^p. They are generally stated in terms of the vectors

$$y_i = (X'X)^{-1} x_i, \qquad i = 1, \dots, n,$$

and include equations (2.24) and (2.31) here. It should be noted that if (1.4) holds then by subtracting the column mean $\bar{x}_j$ from the jth column we can construct an equivalent design with $y_i'y_l = (1/n) + z_i'z_l$, where $\{z_i\}$ can be expected to satisfy the conditions stated for y_i (as defined originally). Using this fact, it is not difficult to show that the conditions (2.31) and those of Portnoy (1985a) hold if (1.4) holds (and the conditions hold for $\{z_i\}$). Only condition (2.24) must be treated differently depending on whether or not (1.4) holds. In Portnoy (1985b), the conditions are shown to hold in probability if the distribution of $\{x_\iota\}$ is a scale mixture of a multivariate normal, although it should be possible to generalize this distribution.

Theorem 2.1. *Assume the conditions for Theorem* 3.1 *of Portnoy* (1985a). *Assume further that* ψ *has a uniformly bounded fourth derivative, and that*

$$\limsup(p^2/n) \le B_0 < \infty, \qquad \liminf(p^2/n) \ge B_1 > 0. \tag{2.1}$$

Then, uniformly in $i = 1, \dots, n$,

$$(x_i'\hat{\beta}) = A_i + B_i + C_i + D_i + E_i + o_p(1/\sqrt{n}), \tag{2.2}$$

where, for some constants c_1, c_2, c_3 *and* $d = E\psi'(R)$,

$$\begin{aligned} A_i &= \frac{1}{d}\sum (y_i'y_l)\psi(R_l),\\ B_i &= \frac{1}{d}\sum\sum (y_i'y_{l_1})(y_{l_1}'y_{l_2})\psi(R_{l_2})\big(\psi'(R_{l_1}) - d\big),\\ C_\iota &= c_2\sum\sum\sum (y_i'y_{l_1})(y_{l_1}'y_{l_3})(y_{l_2}'y_{l_3})\psi(R_{l_2})\psi''(R_{l_1})\psi(R_{l_3}),\\ D_\iota &= c_3\sum\sum\sum\sum (y_i'y_{l_1})(y_{l_1}'y_{l_2})(y_{l_1}'y_{l_3})(y_{l_3}'y_{l_4})\psi''(R_{l_1})\psi(R_{l_2})\\ &\quad\times\big(\psi'(R_{l_3}) - d\big)\psi(R_{l_4}),\\ E_i &= c_1\sum\sum\sum\sum (y_i'y_{l_1})(y_{l_1}'y_{l_2})(y_{l_1}'y_{l_3})(y_{l_1}'y_{l_4})\psi'''(R_{l_1})\\ &\quad\times\psi(R_{l_2})\psi(R_{l_3})\psi(R_{l_4}). \end{aligned} \tag{2.3}$$

Remark. It is possible to prove Theorem 2.1 without (2.1). In fact, the lower-bound condition in (2.1) is only used following (2.9) to obtain the appropriate order for term B_i. If $p^2/n \to 0$, then B_i must be included in additional error terms. Nonetheless, it can be shown that these error terms are of sufficiently small order; and the lower bound is not needed at all. Furthermore, the upper bound can be replaced by $P^{11/4}\ln^2 n/n^2 \to 0$ [see (2.14)].

Proof. The results of Portnoy (1985a) are first extended to show that $x_i'\hat{\beta} = A_i + O_p(p^2/n^{3/2-\varepsilon})$ [see (2.17)]. This requires treating each of the terms in (2.4), generally by computing moments and using forms of Chebyshev's inequality. However, one term, $(y_i'V)$, requires a more accurate expansion for $(x_i'\hat{\beta})$; and thus (2.17) must be inserted in a further expansion of $y_i'V$ to obtain (2.2). The remainder of this section presents some details of this argument and also gives two technical lemmas. The casual reader may want to proceed directly to Section 3.

(i) From Portnoy (1985a), Equation (3.16),

$$\begin{aligned}(x_i'\hat{\beta}) = (y_i'\hat{\theta}) &= y_i'W + y_i'U + c_1 y_i'V + c_2 y_i'SW + c_3 y_i'SU + c_4 y_i'SV\\ &\quad + c_5 y_i'Se_1 + \sum_{k=2}^{\infty} y_i'S^k(W + U + V + e_1),\end{aligned} \tag{2.4}$$

where using results from Section 3 of Portnoy (1985a) (uniformly in $i = 1, \dots, n$),

$$y_i'W = A_i = O_p\left(\frac{p \ln n}{n}\right)^{1/2}, \qquad \|W\|^2 = O_p(p),$$

$$y_i'U = B_i, \qquad \|U\|^2 = O_p\left(\frac{p^2 \ln^2 n}{n}\right),$$

$$y_i'V = \sum\sum (y_i'y_{l_1})(y_{l_1}'y_{l_2})(x_{l_1}'\hat{\beta})^2 \psi(R_{l_2})\psi'''(R_{l_1}^*), \qquad \text{for some } \{R_{l_1}^*\},$$

$$\|V\|^2 = O_p\left(\frac{p^{7/2} \ln^2 n}{n^2}\right), \tag{2.5}$$

$$S = \sum\sum y_{l_1} y_{l_1}'(y_{l_1}'y_{l_2})\psi(R_{l_2})\psi''(R_{l_1}),$$

$$\sup\{u'Su\colon \|u\| = 1\} = O_p\left(\frac{p \ln n}{n}\right)^{1/2},$$

$$e_1 = O_p\left(\frac{p^{3/2}(\ln n)^{5/4}}{n}\right).$$

Also from Section 3 of Portnoy (1985a),

$$\max_i (x_i'\hat{\beta})^2 = O_p\left(\frac{p^3 \ln n}{n^2}\right), \qquad \sum_{i=1}^n (x_i'\hat{\beta})^2 = O_p(p), \tag{2.6}$$

$$\sum_{l=1}^n (y_i'y_l)\psi(R_l) = O_p\left(\frac{p \ln n}{n}\right)^{1/2} \quad \text{uniformly in } i = 1, 2, \dots, n. \tag{2.7}$$

Lastly, the following conditions will also be used here:

$$(y_i'y_l)^2 = O\left(\frac{p \ln n}{n^2}\right) \quad \text{uniformly in } i \neq l \tag{2.8}$$

and

$$\|y_i\|^2 = O\left(\frac{p}{n}\right) \quad \text{uniformly in } i.$$

Consider $y_i'U = B_i$: EB_i^{2k} is a $4k$-fold sum of terms of the form

$$(y_i'y_{l_{\nu_1}}) \cdots (y_i'y_{l_{\nu_{2k}}})(y_{l_{\nu_1}}'y_{l_{\mu_1}}) \cdots (y_{l_{\nu_{2k}}}'y_{l_{\mu_{2k}}})$$
$$\times E\left[\psi(R_{l_{\nu_1}})(\psi'(R_{l_{\mu_1}}) - d) \cdots \psi(R_{l_{\nu_{2k}}})(\psi'(R_{l_{\mu_{2k}}}) - d)\right].$$

Since $E\psi(R) = E(\psi'(R) - d) = E\psi(R)(\psi'(R) - d) = 0$, pairs of l_ν and pairs of l_μ subscripts must be equal. If $l_{\mu_j} \neq l_{\nu_j}$ then (using 2.8), the contribution to EB_i^{2k} is less than $Bn^{2k}(p \ln n/n^2)^{2k}$. If some $l_{\mu_j} = l_{\nu_j}$, $(p \ln n)^{1/2}/n$ is replaced by p/n, but there is one less sum—thus reducing the contribution by a factor of n

and making it smaller. Hence, $EB_i^{2k} = O(p \ln n/n)^{2k}$, and, given $\varepsilon > 0$,

$$P\left\{|B_i| \geq \frac{p}{n^{1-\varepsilon}}, \text{ for some } i = 1, \dots, n\right\} \leq n \frac{B\left(\dfrac{p \ln n}{n}\right)^{2k}}{\left(\dfrac{p}{n^{1-\varepsilon}}\right)^{2k}} = \frac{B(\ln n)^{2k}}{n^{2k\varepsilon - 1}} \to 0,$$

for k such that $2k\varepsilon > 1$. Therefore

$$B_i = O_p\left(\frac{p}{n^{1-\varepsilon}}\right) \quad \text{uniformly in } i = 1, \dots, n. \tag{2.9}$$

From (2.1), (2.9) implies $B_i = O_p(p^2/n^{3/2-\varepsilon})$.

Now consider $y_i'V$: using (2.6), (2.7), and (2.8),

$$\begin{aligned} |y_i'V| &\leq \sum_{l_1} |y_i'y_{l_1}|(x_{l_1}'\hat{\beta})^2 |\psi'''(R_{l_1}^*)| \left|\sum_{l_2}(y_{l_1}'y_{l_2})\psi(R_{l_2})\right| \\ &= O_p\left(\left(\frac{p \ln n}{n^2}\right)^{1/2} p \left(\frac{p \ln n}{n}\right)^{1/2}\right) = O_p\left(\frac{p^2 \ln n}{n^{3/2}}\right). \end{aligned} \tag{2.10}$$

Now define

$$C_i = y_i'SW = \sum\sum\sum(y_i'y_{l_1})(y_{l_1}'y_{l_3})(y_{l_2}'y_{l_3})\psi(R_{l_2})\psi''(R_{l_1})\psi(R_{l_3}). \tag{2.11}$$

Following the argument leading to (2.9), the main contribution to EC_i^{2k} arises when pairs of subscripts are equal. However, for each set of factors of the form $(y_i'y_{l_1})(y_{l_1}'y_{l_3})(y_{l_2}'y_{l_3})$, at most one pair of subscripts can be equal (without reducing the number of sums). Hence,

$$EC_i^{2k} = O\left(n^{3k}\left(\frac{p \ln n}{n^2}\right)^{2k}\left(\frac{p}{n}\right)^{2k}\right) = O\left(\frac{p^2 \ln n}{n^{3/2}}\right)^{2k}.$$

Therefore [as in (2.9)], for any $\varepsilon > 0$,

$$C_i = O_p\left(\frac{p^2}{n^{3/2-\varepsilon}}\right). \tag{2.12}$$

Now define $D_i = y_i'SU$. Using the bounds in (2.5),

$$D_i = O_p\left(\sqrt{\frac{p}{n}}\sqrt{\frac{p \ln n}{n}}\frac{p \ln n}{\sqrt{n}}\right) = O_p\left(\frac{p^2(\ln n)^{3/2}}{n^{3/2}}\right). \tag{2.13}$$

Similarly [using (2.1) also],

$$\begin{aligned} y_i'SV &= O_p\left(\sqrt{\frac{p}{n}}\sqrt{\frac{p \ln n}{n}}\frac{p^{7/4} \ln n}{n}\right) \\ &= O_p\left(\frac{p^{11/4}(\ln n)^{3/2}}{n^2}\right) = O_p\left(\frac{p^2}{n^{3/2}}\right), \end{aligned} \tag{2.14}$$

$$y_i'Se_1 = O_p\left(\sqrt{\frac{p}{n}}\sqrt{\frac{p \ln n}{n}}\frac{p^{3/2}(\ln n)^{5/4}}{n}\right) = O_p\left(\frac{p^{5/2}(\ln n)^2}{n^2}\right). \tag{2.15}$$

Lastly using the bound on $\|S\|$ and the geometric series, the final summation in (2.4), say e^*, satisfies

$$e^* = O_p\left\{\sqrt{\frac{p}{n}}\left(\frac{p\ln n}{n}\left(\frac{1}{1-\sqrt{\frac{p\ln n}{n}}}\right)\right)\sqrt{p}\right\} = O_p\left(\frac{p^2\ln n}{n^{3/2}}\right). \tag{2.16}$$

Therefore from (2.4) and (2.9) through (2.16), for $\varepsilon > 0$,

$$(x_i'\hat{\beta}) = A_i + O_p\left(\frac{p^2}{n^{3/2-\varepsilon}}\right). \tag{2.17}$$

(ii) Now reconsider $y_i'V$ (2.5) and continue the Taylor series expansion of ψ:

$$\begin{aligned} y_i'V &= \sum\sum(y_i'y_{l_1})(y_{l_1}'y_{l_2})\psi(R_{l_2})\left[\psi'''(R_{l_1}) + (x_{l_1}'\hat{\beta})\psi^{(4)}(\tilde{R}_{l_1})\right](x_{l_1}'\hat{\beta})^2 \\ &\equiv T_1 + T_2, \end{aligned} \tag{2.18}$$

where T_1 uses the $\psi'''(R_{l_1})$ term and T_2 uses the $(x_{l_1}'\hat{\beta})\psi^{(4)}(\tilde{R}_{l_1})$ term. Inserting (2.17) in T_2 and using (2.8), (2.7), (2.5) [for A_i in (2.17)] and (2.6), the sum over $l_2 \neq l_1 \neq i$ in T_2 contributes

$$\begin{aligned} T_2 &= O_p\left\{\frac{\sqrt{p\ln n}}{n}\sqrt{\frac{p\ln n}{n}}\left(\sqrt{\frac{p\ln n}{n}} + \frac{p^2}{n^{3/2-\varepsilon}}\right)p\right\} \\ &= O_p\left(\frac{p^{5/2}(\ln n)^{3/2}}{n^2}\right) \end{aligned} \tag{2.19}$$

[using (2.1)]. It is easy to see that the sum for $l_2 = l_1$ or $l_1 = i$ contributes a smaller-order term. Squaring (2.17) [and using (2.5)] and inserting in T_1 yields

$$\begin{aligned} T_1 &= \sum\sum(y_i'y_{l_1})(y_{l_1}'y_{l_2})\psi(R_{l_2})\psi'''(R_{l_1})\left(A_{l_1}^2 + O_p\left(\frac{p^{5/2}}{n^{2-\varepsilon}}\right)\right) \\ &= E_i + O_p\left\{\left(n\frac{\sqrt{p\ln n}}{n} + \frac{p}{n}\right)\sqrt{\frac{p\ln n}{n}}\left(\frac{p^{5/2}}{n^{2-\varepsilon}}\right)\right\} \\ &= E_i + O_p\left(\frac{p^{5/2}}{n^{2-\varepsilon^*}}\right), \quad \text{for some } \varepsilon^*, \end{aligned} \tag{2.20}$$

where E_i is defined in (2.3).

Lastly reconsider the error sum which can be rewritten

$$y_i'S^2W + y_i'S^2(U + V + e_1) + \sum_{k=3}^{\infty} y_i'S^k(W + U + V + e_1). \tag{2.21}$$

Computing $E(y_i'S^2W)^{2k}$ using the argument for EC_i^{2k}, it is possible (though tedious) to show that $E(y_i'S^2W)^{2k} = O(p^7(\ln n)^3/n^5)^k$. Hence [using (2.1)]

$$y_i'S^2W = O_p\left(\frac{p^{7/2}}{n^{5/2-\varepsilon}}\right) = O_p\left(\frac{p^{5/2}}{n^{2-\varepsilon}}\right)$$

uniformly in i. Using (2.5),

$$y_i'S^2(U + V + e_1) = O_p\left\{\sqrt{\frac{p}{n}}\frac{p\ln n}{n}\left(\frac{p\ln n}{\sqrt{n}} + \frac{p^{7/4}\ln n}{n} + \frac{p^{3/2}\ln n}{n}\right)\right\}$$
$$= O_p\left(\frac{p^{5/2}\ln^2 n}{n^2}\right),$$

$$\sum_{k=3}^{\infty} y_i'S^k(W + U + V + e_1)$$
$$= O_p\left\{\sqrt{\frac{p}{n}}\left(\frac{p\ln n}{n}\right)^{3/2}\left(\frac{1}{1-\sqrt{\frac{p\ln n}{n}}}\right)\left(\sqrt{p} + o_p(\sqrt{p})\right)\right\}$$
$$= O_p\left(\frac{p^{5/2}(\ln n)^{3/2}}{n^2}\right).$$

Therefore, the error sum is $O_p(p^{5/2}/n^{2-\varepsilon})$. Therefore, using (2.4) and combining (2.19), (2.20) (for $y_i'V$), (2.14), (2.15), and the error sum,

$$(x_i'\hat{\beta}) = A_i + B_i + C_i + D_i + E_i + O_p\left(\frac{p^{11/4}(\ln n)^2}{n^2} + \frac{p^{5/2}}{n^{2-\varepsilon}}\right),$$

from which the result (2.2) follows. □

To prove the main result, Theorem 3.1, certain results concerning random variables related to $A_i, \dots, E_i$ in (2.2) are needed. In particular, define for each $i = 1, \dots, n$, and for $l \neq i$,

$$\tilde{A}_i = \frac{1}{d}\sum_{j\neq i}(y_i'y_j)\psi(R_j), \qquad \tilde{A}_{il} = \frac{1}{d}\sum_{j\neq i,l}(y_i'y_j)\psi(R_j) \tag{2.22}$$

and define $\tilde{B}_i, \tilde{B}_{il}, \dots, \tilde{E}_i, \tilde{E}_{il}$ from (2.3) analogously. That is, a single subscript, i, involves y_i' and all sums avoid index i; a double subscript, il, involves y_i' and all sums avoid both indices i and l.

The results below will be stated using an error term slightly stronger than "o_p". Given a sequence $\{\gamma_n\}$ such that $\gamma_n \to +\infty$, define $\Delta(\gamma_n)$ to be a random variable satisfying

$$E\left(\Delta^2(\gamma_n)\right) = o\left(1/\gamma_n^2\right). \tag{2.23}$$

Note: In Lemma 2.3, Δ will also be a uniformly bounded function on R^2 and the argument γ_n will become a subscript, viz., $\Delta_{\gamma_n}(u, v)$.

LEMMA 2.2. *Assume the conditions of Theorem* 2.1 *and assume further that either*

$$(2.24)\qquad \left|\sum_{l\neq i} y_i' y_l\right| = O\left(\frac{p \ln n}{n}\right)^{1/2}$$

or, if the design has a constant term [(1.4) *holds*], $y_i'y_l = (1/n) + z_i'z_l$ *with* $\{z_i\}$ *satisfying the conditions for* $\{y_i\}$ [*including* (2.24)].

(i) *Then, with* $d = E\psi'(R)$ *and* $\sigma^2 = \operatorname{Var}\psi(R)$,

$$(2.25)\qquad \begin{cases} E\left(\dfrac{1}{\sqrt{n}}\displaystyle\sum_{i=1}^{n}\tilde{A}_i - \dfrac{f(x)}{d\sqrt{n}}\sum_{i=1}^{n}\psi(R_i)\right)^2 = o(1), & \textit{if } (1.4) \textit{ holds}, \\ E\left(\dfrac{1}{\sqrt{n}}\displaystyle\sum_{i=1}^{n}\tilde{A}_i\right)^2 = o(1), & \textit{otherwise}, \end{cases}$$

$$(2.26)\qquad E\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{A}_i^2 - \frac{\sigma^2}{d^2}\frac{p}{\sqrt{n}}\right)^2 = o(1).$$

(ii) *Let X denote B, C, D, or E and let* $\tilde{X}$ *denote* $\tilde{B}$, $\tilde{C}$, $\tilde{D}$, *or* $\tilde{E}$ *as modified by* (2.22). *Then*

$$(2.27)\qquad E\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}X_i\right)^2 = o(1) \quad \textit{and} \quad |X_i - \tilde{X}_i| = \Delta(\sqrt{n}).$$

Note also that $A_i = \tilde{A}_i + (1/d)\|y_i\|^2\psi(R_i)$.

(iii) *Let* $V_i = A_i + B_i + C_i + D_i + E_i$ *with* $\tilde{V}_i$ *and* $\tilde{V}_{il}$ *defined analogously to* (2.22). *Then*

$$(2.28)\qquad \tilde{V}_i = \Delta\left(\sqrt{n/p}\right), \qquad \tilde{V}_{il} = \Delta\left(\sqrt{n/p}\right), \quad \textit{and} \quad \tilde{V}_{il}^2 = \tilde{A}_{il}^2 + \Delta(\sqrt{n}).$$

REMARK. Note that condition (2.24) is similar to conditions in Portnoy (1985a); particularly, condition X4. It is easy to see that (2.24) will hold in probability using the arguments of Portnoy (1985b). Also note that (2.24) is the only condition affected by the presence of a constant term [the conditions in Portnoy (1985a) hold with or without a constant term in the design].

PROOF. The proof involves straightforward but very tedious computations. Since the calculations are all rather similar, only the case for D_i (one of the more complicated cases) will be sketched. So, to obtain the first part of (2.27), consider $E(\Sigma D_i/\sqrt{n})^2$. If (2.24) holds (i.e., the design lacks a constant), this involves a 10-fold sum over i_1, i_2 and eight subscripts l_j. As in the proof of Theorem 2.1, these subscripts must be equal at least in pairs. Thus, using the fact that $l_j \neq i_1$ or i_2 and $|y_{l_j}' y_{l_k}| \leq \|y_{l_j}\|\,\|y_{l_k}\| \leq Bp/n$, and using (2.24),

$$(2.29)\qquad \begin{aligned} E\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{D}_i\right)^2 &\leq \frac{B}{n}\sum_{l_1}\cdots\sum_{l_4}\left|\sum_{i_1}\sum_{i_2}(y_{i_1}'y_{l_1})(y_{i_2}'y_{l_2})\right|\frac{p^6}{n^6} \\ &= O\left(\frac{1}{n}\frac{p\ln n}{n}n^4\frac{p^6}{n^6}\right) = O\left(\frac{p^7\ln n}{n^4}\right) = o(1). \end{aligned}$$

If the design has a constant term, each factor $(y_i'y_l)$ can be written $(1/n) + (z_i'z_l)$, and the product can be expanded to yield

$$D_i = c_3 \sum\sum\sum\sum \Big\{ \frac{1}{n^4} + \frac{1}{n^3}\big[(z_i'z_{l_1}) + (z_{l_1}'z_{l_3}) + (z_{l_2}'z_{l_3}) + (z_{l_3}'z_{l_4})\big] + \frac{1}{n^2}\big[(z_i'z_{l_1})(z_{l_1}'z_{l_2}) + \cdots + (z_{l_1}'z_{l_3})(z_{l_3}'z_{l_4})\big] + \frac{1}{n}\big[(z_i'z_{l_1})(z_{l_1}'z_{l_2})(z_{l_1}'z_{l_3}) + \cdots + (z_{l_1}'z_{l_2})(z_{l_1}'z_{l_3})(z_{l_3}'z_{l_4})\big]\Big\} \times \psi''(R_{l_1})\psi(R_{l_2})\big(\psi'(R_{l_3}) - d\big)\psi(R_{l_4}).$$

Consider $E((1/\sqrt{n})\Sigma D_\iota)^2$. The $1/n^4$ term contributes

$$\frac{1}{n^8}\frac{1}{n}\sum_{i_1}\sum_{i_2} E(\text{8-fold sum}) = \frac{1}{n^8}O(n^4),$$

since subscripts in the 8-fold sum must be equal in pairs. Similarly, since $|z_i'z_l| \le \|z_i\|\,\|z_l\| = O(p/n)$, the $1/n^3$ term contributes $(1/n^6)nO(p^2/n^2)O(n^4) = O(p^2/n^3)$; and the $1/n^2$ term contributes $(1/n^4)nO(p^4/n^4)O(n^4) = O(p^4/n^3) \to 0$. For the $1/n$ term, first consider the last term [without $(z_i'z_{l_1})$]. In each term in the expectation of the square of the sum, at least two of the six factors must have unequal subscripts so that $(z_j'z_k)^2 = O(p \ln n/n^2)$. Thus, the contribution is

$$\frac{1}{n^2}nO\left(\frac{p \ln n}{n^2}\right)O\left(\frac{p^4}{n^4}\right)O(n^4) = O\left(\frac{p^5 \ln n}{n^3}\right) \to 0.$$

Lastly, for the term involving $(z_i'z_{l_1})$, either $i = l_1$ (which eliminates the sum over i and gives a smaller contribution), or $i \ne l_1$ (which allows the above argument to apply); thus proving the first part of (2.27) for D_i.

Equation (2.24) is not needed to obtain the second part of (2.27). Note that $D_i - \tilde{D}_\iota$ involves sums over no more than three subscripts (since at least one subscript l_1, l_2, l_3, or l_4 must equal i). Thus, $E(D_i - \tilde{D}_i)^2$ is a six-fold sum of products of eight factors of the form $(y_{l_j}'y_{l_k})$. From definition (2.3), at least two pairs of subscripts (l_j, l_k) must be unequal (in each term). Hence $E(D_i - \tilde{D}_i)^2$ can be bounded by

$$Bn^3\frac{p \ln n}{n^2}\frac{p^6}{n^6} = O\left(\frac{p^7 \ln n}{n^5}\right) = o\left(\frac{1}{n}\right).$$

The remaining results in Lemma 2.2 follow using similar expectation calculations. □

LEMMA 2.3. *Assume the hypotheses of Lemma* 2.2, *and define* (*for* $i \ne l$)

$$G_{il}(u) = \frac{1}{d}\Big\{(y_i'y_l)\psi(u) + (y_i'y_l)\tilde{A}_{il}(\psi'(u) - d) + \psi(u)\sum_{l_1 \ne i,l}(y_i'y_{l_1})(y_l'y_{l_1})\big(\psi'(R_{l_1}) - d\big)\Big\} \tag{2.30}$$

and similarly for $G_{li}(u)$. *Assume*

(2.31) $$\sum_{l_1 \neq i, l} (y_i' y_{l_1}) \|y_{l_1}\|^2 = O\left(\frac{p(\ln n)^{3/2}}{n}\right) \quad \text{uniformly in } i, l.$$

Then, with $\Delta(n)$ *defined in* (2.23) *and* $\Delta_{\gamma_n}(u)$ *uniformly bounded in* u *and of order* $\Delta(\gamma_n)$,

(2.32) $$\tilde{V}_i = \tilde{V}_{il} + G_{il}(R_l) + \Delta(n)$$

and

(2.33) $$G_{il}(u) = \Delta_{\sqrt{n}}(u).$$

Furthermore,

(2.34) $$\frac{1}{n} \sum\sum_{i \neq l} \left\{ F(x + \tilde{V}_{il}) \int_0^{x + \tilde{V}_{il}} G_{il}(u) f(u)\, du \right\} = \Delta(1)$$

and similarly when i *and* l *are interchanged.*

PROOF. First note that $G_{il}(R_l) = \tilde{A}_i - \tilde{A}_{il} + \tilde{B}_i - \tilde{B}_{il}$ exactly. Consider $\tilde{C}_i - \tilde{C}_{il}$:

$$\begin{aligned} \tilde{C}_i - \tilde{C}_{il} = {} & c_2(y_i' y_l) \psi''(R_l) \sum\sum_{l_1, l_2 \neq i, l} (y_i' y_{l_2})(y_{l_1}' y_{l_2}) \psi(R_{l_1}) \psi(R_{l_2}) \\ & + c_2 \psi(R_l) \sum\sum_{l_1, l_2 \neq i, l} (y_i' y_{l_1})(y_{l_1}' y_{l_2})(y_l' y_{l_2}) \psi''(R_{l_1}) \psi(R_{l_2}) \\ & + c_2 \psi(R_l) \sum\sum_{l_1, l_2 \neq i, l} (y_i' y_{l_1})(y_l' y_{l_1})(y_{l_1}' y_{l_2}) \psi''(R_{l_1}) \psi(R_{l_2}). \end{aligned}$$

As before, using (2.31),

$$\begin{aligned} E(\text{first term above})^2 & \leq B(y_i' y_l)^2 \sum\sum_{l_1, l_2 \neq i, l} \{ (y_i' y_{l_1})(y_i' y_{l_2}) \|y_{l_1}\|^2 \|y_{l_2}\|^2 \\ & \qquad + (y_i' y_{l_1})^2 (y_{l_1}' y_{l_2})^2 + (y_i' y_{l_1})(y_i' y_{l_2})(y_{l_1}' y_{l_2})^2 \} \\ & \leq B(y_i' y_l)^2 \left(\sum_{l_1 \neq i, l} (y_i' y_{l_1}) \|y_{l_1}\|^2 \right)^2 + O\left(n^2 \frac{p^3 \ln^3 n}{n^6} \right) \\ & = O\left(\frac{p \ln n}{n^2} \frac{p^2 (\ln n)^3}{n^2} + \frac{p^3 \ln^3 n}{n^4} \right) \\ & = \frac{1}{n^2} O\left(\frac{p^3 \ln^4 n}{n^2} \right) = o\left(\frac{1}{n^2} \right). \end{aligned}$$

The terms $\tilde{D}_i - \tilde{D}_{il}$ and $\tilde{E}_i - \tilde{E}_{il}$ can be similarly bounded (with even smaller bounds); and, hence, (2.32) holds. Similar (even easier) computations yield (2.33).

Lastly, for (2.34) note that [by Lemma 2.2 and (2.33)],

$$F(x+\tilde{V}_{il})\int_0^{x+\tilde{V}_{il}} G_{il}(u)f(u)\,du = \{F(x)+\tilde{V}_{il}f(x)+\tfrac{1}{2}\tilde{A}_{il}^2 f'(x)\}$$
$$\times\left\{\int_0^x G_{il}(u)f(u)\,du + \tilde{V}_{il}G_{il}(x)f(x) + \tfrac{1}{2}\tilde{A}_{il}^2\frac{d}{dx}(G_{iu}(x)f(x))\right\} + \Delta(n).$$

The computations are quite tedious, but expectations of each of the terms above can be computed (as in the earlier proofs), and it can be shown that $1/n$ times the double sum of each term tends to zero. □

3. The basic results.

THEOREM 3.1. *Assume Theorem* 2.1 *holds and assume the conditions for Lemmas* 2.2, 2.3, *and* 3.3. *Assume that* $p/\sqrt{n} \to c$. *Then for each fixed* x, *the empirical distribution of residuals* [*see* (1.3)] *satisfies*

$$\text{(3.1)}\quad \sqrt{n}\,(\hat{F}_n(x) - \hat{F}_n^*(x) - H_n(x)) - \frac{c}{2}\left(\frac{\sigma^2}{d^2}f'(x) + \frac{1}{d}f(x)\psi(x)\right) \to_p 0,$$

where

$$\text{(3.2)}\qquad \hat{F}_n^*(x) = \frac{1}{n}\sum_{i=1}^n I(R_i \le x), \qquad \sigma^2 = E\psi^2(R), \qquad d = E\psi'(R),$$

and

$$H_n(x) = \frac{1}{nd}f(x)\sum_{i=1}^n \psi(R_i),$$

if (1.4) *holds, and vanishes otherwise.*

PROOF. First, as in Lemma 2.2, define

$$\text{(3.3)}\qquad V_i = A_i + B_i + C_i + D_i + E_i$$

and define $\tilde{V}_i$ (which omits functions of R_i) and $\tilde{V}_{il}$ and $\tilde{V}_{li}$ analogously as in (2.22). Let $\delta_n = o(1/\sqrt{n})$. By Theorem 2.1 and Lemma 2.2, with probability tending to one,

$$\text{(3.4)}\qquad I(R_i - x_i'\hat{\beta} \le x) \le I(R_i \le V_i + x + \tfrac{1}{2}\delta_n) \le I_i,$$

where

$$\text{(3.5)}\qquad I_i = I\left(R_i - \frac{1}{d}\|y_i\|^2\psi(R_i) \le x + \tilde{V}_i + \delta_n\right).$$

Since the reverse inequality holds with δ_n replaced by $-\delta_n$, it suffices to consider $\hat{F}_n$ defined using I_i instead of $I(R_i - x_i'\hat{\beta} \le x)$.

The remainder of the proof is concerned with establishing the result

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (I_i - I(R_i \le x) - \tilde{U}_i) \to_P 0, \tag{3.6}$$

where

$$\tilde{U}_i = \tilde{V}_i f(x) + \tfrac{1}{2}\tilde{A}_i^2 f'(x) + \frac{1}{2d}\|y_i\|^2 f(x)\psi(x). \tag{3.7}$$

This result will be established by computing the second moment of (3.6). Notice that Lemma 2.2 [(2.25) and (2.27)] shows that $f(x)\Sigma\tilde{V}_i/\sqrt{n} \to_P H_n(x)$ (if the design has a constant) and $\Sigma(A_i^2 - (\sigma^2 p/d^2\sqrt{n}))/\sqrt{n} \to_p 0$ uniformly in $i = 1, \dots, n$. Hence, (3.6) immediately yields the result (3.1) (since $\Sigma\|y_i\|^2 = p$).

Now write

$$\begin{aligned} E\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (I_i - I(R_i \le x) - \tilde{U}_i)\right)^2 &= \frac{1}{n}\sum_{i=1}^{n} E(I_i - I(R_i \le x) - \tilde{U}_i)^2 \\ &\quad + \frac{1}{n}\sum\sum_{i \ne l} E(I_i - I(R_i \le x) - \tilde{U}_i)(I_l - I(R_l \le x) - \tilde{U}_l). \end{aligned} \tag{3.8}$$

Consider the second double sum term. By Lemma 2.3, note that with G_{il} defined in (2.30),

$$I_i = I\left(R_i - \frac{1}{d}\|y_i\|^2\psi(R_i) - G_{il}(R_l) + \Delta(n) \le x + \tilde{V}_{il} + \delta_n\right). \tag{3.9}$$

Now let $\mathscr{S}_{il}$ denote the σ-field generated by all R_j except R_i and R_l. Expanding the product in the double sum in (3.8) and conditioning on $\mathscr{S}_{il}$ yields nine terms, the first of which can be computed using Lemma 3.3 [see (3.21)] as follows (with $Z_{il} = x + \tilde{V}_{il} + \delta_n$):

$$\begin{aligned} E[I_i I_l | \mathscr{S}_{il}] &= F(Z_{il})F(Z_{li}) + \frac{1}{d}\|y_i\|^2\psi(Z_{il})f(Z_{il})F(Z_{li}) \\ &\quad + \frac{1}{d}\|y_l\|^2\psi(Z_{li})f(Z_{li})F(Z_{il}) \\ &\quad + \frac{1}{2d^2}\{\|y_i\|^4 K_1(Z_{il})F(Z_{li}) \\ &\qquad + \|y_l\|^4 K_1(Z_{li})F(Z_{il}) + 2\|y_i\|^2\|y_l\|^2 K_2(Z_{il})K_2(Z_{li})\} \\ &\quad + f(Z_{il})\int_{-\infty}^{Z_{li}} G_{il}(v)f(v)\,dv + f(Z_{li})\int_{-\infty}^{Z_{il}} G_{li}(u)f(u)\,du + \Delta(n), \end{aligned} \tag{3.10}$$

where (using notation from Lemma 3.3)

$$\begin{aligned} K_1(w) &= \int_{-\infty}^{w} (k_2(u) + k_1^2(u))f(u)\,du, \\ K_2(w) &= \int_{-\infty}^{w} k_1(u)f(u)\,du = \psi(w)f(w). \end{aligned} \tag{3.11}$$

Expanding f, and ψ about $x + \tilde{V}_{\iota l}$ and about x [using the fact that $\tilde{V}_{\iota l}$ is of form $\Delta(\sqrt{n/p})$ by Lemma 2.2],

$$\begin{aligned} E[I_\iota I_l|\mathscr{S}_{\iota l}] &= F(Z_{\iota l})F(Z_{l\iota}) + \frac{1}{d}\|y_\iota\|^2 f(x+\tilde{V}_{\iota l})\psi(x+\tilde{V}_{\iota l})F(Z_{l\iota}) \\ &\quad + \frac{1}{d}\|y_l\|^2 f(x+\tilde{V}_{l\iota})\psi(x+\tilde{V}_{l\iota})F(Z_{\iota l}) \\ &\quad + \frac{1}{2d^2}\{(\|y_\iota\|^4 + \|y_l\|^4)K_1(x)F(x) \\ &\qquad + 2\|y_\iota\|^2\|y_l\|^2 f^2(x)\psi^2(x)\} + \Delta(n), \end{aligned} \tag{3.12}$$

where $G_{\iota l}$ and $G_{l\iota}$ terms have been omitted since their double sum tends to zero by Lemma 2.3. Similarly, the other terms of the following forms can be computed [using Lemma 3.3, (3.22)]:

$$\begin{aligned} E[I_\iota I(R_l < x)|\mathscr{S}_{\iota l}] &= F(x + \tilde{V}_{\iota l} + \delta_n)F(x) + \frac{1}{2d^2}\|y_\iota\|^4 K_1(x)F(x) \\ &\quad + \frac{1}{d}\|y_\iota\|^2\psi(x+\tilde{V}_{\iota l})f(x+\tilde{V}_{\iota l})F(x) + \Delta(n), \end{aligned} \tag{3.13}$$

$$\begin{aligned} E[\tilde{U}_\iota I_l|\mathscr{S}_{\iota l}] &= \int_{-\infty}^{\infty}\int_{-\infty}^{x+\tilde{V}_{l\iota}+\delta_n}\Big\{\tilde{U}_{\iota l} + G_{\iota l}(v)f(x) \\ &\quad + \frac{1}{d}(y_\iota' y_l)\psi(v)\tilde{A}_{\iota l}\Big\} f^*(u,v)\,du\,dv + \Delta(n) \\ &= \tilde{U}_{\iota l}F(x + \tilde{V}_{l i} + \delta_n) + \frac{1}{d}\|y_\iota\|^2\tilde{U}_{\iota l}\psi(x+\tilde{V}_{l i})f(x+\tilde{V}_{l i}) + \Delta(n), \end{aligned} \tag{3.14}$$

$$E[\tilde{U}_i I(R_l \le x)|\mathscr{S}_{\iota l}] = \tilde{U}_{\iota l}F(x) + \Delta(n), \tag{3.15}$$

where, again, $G_{\iota l}$ and related terms are omitted (using Lemma 2.3).

With some tedious computation, terms of the form (3.12) through (3.15) can be summed to obtain (ignoring terms whose double sum tends to zero when divided by n)

$$\begin{aligned} &E[(I_\iota - I(R_\iota \le x) - \tilde{U}_i)(I_l - I(R_l \le x) - \tilde{U}_l)|\mathscr{S}_{\iota l}] \\ &\quad = (F(x+\tilde{V}_{il}+\delta_n) - F(x) - \tilde{U}_{il})(F(x+\tilde{V}_{li}+\delta_n) - F(x) - \tilde{U}_{l\iota}) \\ &\qquad + \frac{1}{d}\|y_\iota\|^2\psi(x+\tilde{V}_{\iota l})f(x+\tilde{V}_{il})\{F(x+\tilde{V}_{li}+\delta_n) - F(x) - \tilde{U}_{l\iota}\} \\ &\qquad + \frac{1}{d}\|y_l\|^2\psi(x+\tilde{V}_{l\iota})f(x+\tilde{V}_{l\iota})\{F(x+\tilde{V}_{\iota l}+\delta_n) - F(x) - \tilde{U}_{\iota l}\} \\ &\qquad + \frac{1}{d^2}\|y_i\|^2\|y_l\|^2 f^2(x)\psi^2(x) + \Delta(n). \end{aligned} \tag{3.16}$$

Now expanding $F(x + \tilde{V}_{il} + \delta_n)$ using Lemma 2.2 [see (2.29)],

$$F(x + \tilde{V}_{\iota l} + \delta_n) - F(x) - \tilde{U}_{\iota l} = -\frac{1}{d}\|y_i\|^2 f(x)\psi(x) + \Delta(\sqrt{n})$$

(and similarly for i and l reversed). Thus, the first (product) term in (3.16) contributes

$$\frac{1}{d^2}\|y_i\|^2\|y_l\|^2 f^2(x)\psi^2(x) + \Delta(n).$$

Similarly, the second and third terms in (3.16) each contribute the negative of this term; and these exactly cancel the last term in (3.16). Therefore,

$$\begin{aligned}\frac{1}{n}\sum\sum_{i\neq l} E(I_i - I(R_i \le x) - \tilde{U}_i)(I_l - I(R_l \le x) - \tilde{U}_l) \\ = \frac{1}{n}\sum\sum_{i\neq l} E\Delta(n) \\ = o\left(\frac{1}{n^2}n^2\right) \to 0.\end{aligned}$$

Lastly consider the first square term in (3.8). Using Lemma 2.2 and computation similar to the above, it can be shown that $E(I_i - I(R_i \le x) - \tilde{U}_i)^2 \to 0$ uniformly in i. Hence, both terms in (3.8) tend to zero, and the proof is complete. □

COROLLARY 3.2. *Following Theorem* 3.1, *define* $H_n(x)$ *as in* (3.2) *and*

$$g(x) = \frac{c}{2}\left(\frac{\sigma^2}{d^2}f'(x) + \frac{1}{d}f(x)\psi(x)\right). \tag{3.17}$$

If Theorem 3.1 *holds and* $\{x_1,\dots,x_k\}$ *are fixed, the joint distribution of the random variables* $\{\sqrt{n}(\hat{F}_n(x_i) - F(x_i) - H_n(x_i)) - g(x_i)\}_{i=1}^k$ *is the same as that of* $\{\sqrt{n}(\hat{F}_n^*(x_i) - F(x_i))\}_{i=1}^k$. *If Theorem* 2.1 *holds, the processes* $\{\sqrt{n}(\hat{F}_n(x) - F(x) - H_n(x)) - g(x)\}_{n=1}^\infty$ *are tight; and, hence, they converge weakly to the transformed Brownian bridge process to which* $\{\sqrt{n}(\hat{F}_n^*(x) - F(x))\}_{n=1}^\infty$ *converges, if* $p^2/n \to c$.

PROOF. The joint distribution result is an immediate consequence of Theorem 3.1. So it remains to obtain the tightness result. By Theorem 2.1, there is $\varepsilon_n \to 0$ such that with probability tending to one,

$$x_i'\beta - \varepsilon_n \le x_i'\hat{\beta} \le x_i'\beta + \varepsilon_n \quad \text{uniformly in } i = 1,\dots,n.$$

Thus, with probability tending to one,

$$\begin{aligned}\sup_x\left|\frac{1}{\sqrt{n}}\sum_{i=1}^n\{I(R_i - x_i'\hat{\beta} \le x + \delta) - I(R_i - x_i'\hat{\beta} \le x)\}\right| \\ \le \tfrac{1}{2}\sup_x\left|\frac{1}{\sqrt{n}}\sum_{i=1}^n\{I(R_i \le x + \delta + \varepsilon_n) - I(R_i \le x - \varepsilon_n)\}\right| \\ + \tfrac{1}{2}\sup_x\left|\frac{1}{\sqrt{n}}\sum_{i=1}^n\{I(R_i \le x + \delta - \varepsilon_n) - I(R_i \le x + \varepsilon_n)\}\right|.\end{aligned} \tag{3.18}$$

Therefore, tightness follows from tightness of the usual empiric c.d.f. [see, for example, Billingsley (1968), Section 15], uniform continuity of $g(x)$, and tightness of $\{\sqrt{n}H_n(x)\}$ [which follows from uniform continuity of $f(x)$]. □

REMARK. (1) The corollary implies that standard goodness-of-fit tests based on F_n should work if $p^2/n \to 0$ and the design lacks a constant, but not otherwise, unless $g(x)$ vanishes. If $\psi(x) = -f'(x)/f(x)$, then $g(x)$ will indeed vanish. However, $g(x)$ will generally not vanish.

(2) Whenever the design has a constant, the asymptotic distribution of $\sqrt{n}(\hat{F}_n(x) - F(x))$ will be affected [as in Pierce and Kopecky (1979)]. Without adjustment, these processes will converge weakly to the limiting process for $\sqrt{n}(\hat{F}_n^*(x) - F(x) + H_n(x)) + g(x)$, which is a Gaussian process with mean $g(x)$ and covariance function (for $x_1 < x_2$)

$$\begin{aligned}\gamma(x_1, x_2) &= \operatorname{Cov}\{I(R_i \le x_1) + f(x_1)\psi(R_i)/d,\ I(R_i \le x_2) + f(x_2)\psi(R_i)/d\}\\ &= F(x_1 \vee x_2) - F(x_1)F(x_2) + \frac{1}{d}f(x_1)\int_{-\infty}^{x_2}\psi(r)f(r)\,dr\\ &\quad + \frac{1}{d}f(x_2)\int_{-\infty}^{x_1}\psi(r)f(r)\,dr + \frac{\sigma^2}{d^2}f(x_1)f(x_2).\end{aligned}$$

Lastly, the following technical result is required in Theorem 3.1:

LEMMA 3.3. *Assume Lemma 2.2 holds. For fixed values of i and l, fix $\{R_j: j \neq i, l\}$ and [using (3.3) and (2.30)], define*

$$\begin{aligned}h(r, s) &= r - \frac{1}{d}\|y_i\|^2\psi(r) - (\tilde{V}_i - \tilde{V}_{il})\\ &= r - \frac{1}{d}\|y_i\|^2\psi(r) - G_{il}(s) + \Delta_n(r, s),\end{aligned} \tag{3.19}$$

where (from Lemma 2.3), Δ_n is uniformly bounded in (r, s) and $E\Delta_n^2 = o(1/n^2)$. From now on, let Δ_n denote a generic function satisfying these properties. [Note: $E\Delta^2_{\sqrt{n}} = o(1/n)$].

Consider the transformation

$$U = h(R_i, R_l), \qquad V = h(R_l, R_i). \tag{3.20}$$

Let $f(\cdot)$ be the density of R_j and assume that $l(u) \equiv \log f(u)$ has three bounded, continuous derivatives. Then the joint density of (U, V) satisfies

$$\begin{aligned}f_{U,V}(u, v) = f(u)f(v)\Big\{1 &+ \frac{1}{d}\|y_i\|^2k_1(u) + \frac{1}{d}\|y_l\|^2k_1(v)\\ &+ G_{il}(v)l'(u) + G_{li}(u)l'(v)\\ &+ \frac{1}{2d^2}\|y_i\|^4(k_2(u) + k_1^2(u))\\ &+ \frac{1}{2d^2}\|y_l\|^4(k_2(v) + k_1^2(v))\\ &+ \frac{1}{d^2}\|y_i\|^2\|y_l\|^2k_1(u)k_1(v) + \Delta_n(u, v)\Big\},\end{aligned} \tag{3.21}$$

where

$$(3.22)\qquad \begin{aligned} k_1(u) &= \psi(u)l'(u) + \psi'(u),\\ k_2(u) &= \psi^2(u)l''(u) + (\psi'(u))^2 + 2\psi(u)\psi'(u)l'(u). \end{aligned}$$

Similarly, consider the transformation

$$(3.23)\qquad \begin{aligned} U^* &= h(R_i, R_l),\\ V^* &= R_l. \end{aligned}$$

Then

$$(3.24)\qquad \begin{aligned} f_{U^*,V^*}(u,v) = f(u)f(v)\biggl\{1 &+ \frac{1}{d}\|y_i\|^2 k_1(u) + G_{il}(v)l'(u)\\ &+ \frac{1}{2d^2}\|y_i\|^4\bigl(k_2(u) + k_1^2(u)\bigr) + \Delta_n(u,v)\biggr\}. \end{aligned}$$

PROOF. First compute the matrix of partials,

$$(3.25)\qquad \begin{aligned} J &= \begin{pmatrix} \dfrac{\partial h(u,v)}{\partial u} & \dfrac{\partial h(u,v)}{\partial v}\\ \dfrac{\partial h(v,u)}{\partial u} & \dfrac{\partial h(v,u)}{\partial v} \end{pmatrix}\\ &= \begin{pmatrix} 1 - \dfrac{1}{d}\|y_i\|^2\psi'(u) & G'_{il}(v)\\ G'_{li}(u) & 1 - \dfrac{1}{d}\|y_l\|^2\psi'(v) \end{pmatrix}. \end{aligned}$$

Then, using Lemma 2.3 [which shows that $G_{il}(u)$ is of the form $\Delta_{\sqrt{n}}(u,v)$] and the fact that $\|y_i\|^2 \le Bp/n$, it is not difficult to compute

$$(3.26)\qquad J^{-1} = \begin{pmatrix} 1 + \dfrac{1}{d}\|y_l\|^2\psi'(v) + \Delta_{\sqrt{n}}(u,v) & \Delta_{\sqrt{n}}(u,v)\\ \Delta_{\sqrt{n}}(u,v) & 1 + \dfrac{1}{d}\|y_i\|^2\psi'(u) + \Delta_{\sqrt{n}}(u,v) \end{pmatrix}.$$

Note that all second partials of $h(u,v)$ are bounded by p/n times a bounded function of (u,v) with finite second moment. Since $u - h(u,v)$ also has this property, the inverse function can be expanded as follows [since $p^3/n^3 = o(1/n)$],

$$(3.27)\qquad \begin{aligned} \begin{pmatrix} r\\ s \end{pmatrix} &= \begin{pmatrix} u\\ v \end{pmatrix} + J^{-1}\begin{pmatrix} u - h(u,v)\\ v - h(v,u) \end{pmatrix} + \begin{pmatrix} \Delta_n(u,v)\\ \Delta_n(u,v) \end{pmatrix}\\ &= \begin{pmatrix} u + \dfrac{1}{d}\|y_i\|^2\psi(u) + G_{il}(v) + \Delta_n(u,v)\\ v + \dfrac{1}{d}\|y_l\|^2\psi(v) + G_{li}(u) + \Delta_n(u,v) \end{pmatrix}. \end{aligned}$$

Similarly,

$$\log\det J = \log\left(1 - \frac{1}{d}\|y_i\|^2\psi'(u)\right) + \log\left(1 - \frac{1}{d}\|y_l\|^2\psi'(v)\right) + \Delta_n(u,v). \tag{3.28}$$

Now, since R_i and R_l are independent,

$$\log f_{U,V}(u,v) = \log(f(r)f(s)/\det J) = l(r) + l(s) - \log\det J. \tag{3.29}$$

So using (3.27) and expanding $l(r)$ in a Taylor series,

$$l(r) = l(u) + \left(\frac{1}{d}\|y_i\|^2\psi(u) + G_{il}(v)\right)l'(u) + \frac{1}{2d^2}\|y_i\|^4\psi^2(u)l''(u) + \Delta_n(u,v). \tag{3.30}$$

Thus, expanding the log in (3.28) and inserting in (3.29),

$$\begin{aligned}\log f(u,v) = {} & l(u) + \frac{1}{d}\|y_i\|^2 k_1(u) + G_{il}(v)l'(u) \\ & + \frac{1}{2d^2}\|y_i\|^4\left(\psi^2(u)l''(u) + (\psi'(u))^2\right) \\ & + l(v) + \frac{1}{d}\|y_l\|^2 k_1(v) + G_{li}(u)l'(v) \\ & + \frac{1}{2d^2}\|y_l\|^4\left(\psi^2(v)l''(v) + (\psi'(v))^2\right) + \Delta_n(u,v).\end{aligned}$$

Therefore, the result (3.21) follows from exponentiating; and (3.22) follows similarly. □

REFERENCES

BICKEL, P. and FREEDMAN, D. (1983). Bootstrapping regression models with many parameters. In *Festschrift for Erich L. Lehmann* (P. Bickel, K. Doksum, and J. Hodges, Jr., eds.) 28–48. Wadsworth, Belmont, Calif.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

BURKE, M. D., CSÖRGŐ, M., CSÖRGŐ, S., and RÉVÉSZ, P. (1979). Approximation of the empirical process when parameters are estimated. *Ann. Probab.* **7** 790–810.

FREEDMAN, D. A. (1981). Bootstrapping regression models. *Ann. Statist.* **9** 1218–1238.

HUBER, P. (1973). Robust regression: Asymptotics, conjectures, and Monte Carlo. *Ann. Statist.* **1** 799–821.

HUBER, P. (1981). *Robust Statistics.* Wiley, New York.

KOUL, H. (1969). Asymptotic behavior of Wilcoxon type confidence procedures in multiple linear regression. *Ann. Math. Statist.* **40** 1950–1979.

LOYNES, R. M. (1980). The empirical distribution function of residuals from generalized regression. *Ann. Statist.* **8** 284–298.

PIERCE, D. A. and KOPECKY, K. J. (1979). Testing goodness of fit for the distribution of errors in regression models. *Biometrika* **66** 1–5.

PORTNOY, S. (1984). Asymptotic behavior of M-estimators of p regression parameters when p^2/n is large, I: Consistency. *Ann. Statist.* **12** 1298–1309.

PORTNOY, S. (1985a). Asymptotic behavior of M-estimators of p regression parameters when p^2/n is large, II: Asymptotic normality. *Ann. Statist.* **13** 1403–1417.

PORTNOY, S. (1985b). A central limit theorem applicable to robust regression estimators. *J. Multivariate Anal.* To appear.

SHORACK, G. (1982). Bootstrapping robust regression. *Comm. Statist. A—Theory Methods* **11** 961–972.

SHORACK, G. (1985). Empirical and rank processes of observations and residuals. *Canad. J. Statist.* **12** 319–322.

YOHAI, V. and MARONNA, R. (1979). Asymptotic behavior of M-estimators for the linear model. *Ann. Statist.* **7** 258–268.

DEPARTMENT OF STATISTICS
UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS 61801