Then by the Lipschitz-continuity of g in a neighborhood of β, there is a $\delta > 0$ such that on A_n^δ,

$$\|G(\zeta_s) - G(\hat{\beta})\| \le c\|\hat{\beta}_s - \hat{\beta}\|,$$

where c is a positive constant. Let $I_{A_n^\delta}$ be the indicator function of A_n^δ. Then

$$\begin{aligned} E|\hat{B}_{\mathrm{WR}} I_{A_n^\delta}| &\le \frac{r-k+1}{n-r}\Sigma_r W_s \left[E\left(\|G(\zeta_s) - G(\hat{\beta})\|^2 I_{A_n^\delta}\right) E\|\hat{\beta}_s - \hat{\beta}\|^2\right]^{1/2} \\ &\le c\frac{r-k+1}{n-r}\Sigma_r W_s E\|\hat{\beta}_s - \hat{\beta}\|^2 \\ &= c\,\mathrm{Tr}\left[Ev_{J,r}(\hat{\beta})\right] \\ &= O(n^{-1}), \end{aligned}$$

where the last equality follows from Theorem 1 of Shao and Wu (1985). Hence

$$\hat{B}_{\mathrm{WR}} I_{A_n^\delta} = O_p(n^{-1}).$$

From the lemma, $\mathrm{Prob}(A_n^\delta) \to 1$ as $n \to \infty$. Thus

$$\hat{B}_{\mathrm{WR}} = O_p(n^{-1}). \qquad \square$$

REFERENCES

Hinkley, D. V. (1977). Jackknifing in unbalanced situations. *Technometrics* **19** 285–292.

Miller, R. G. (1974). An unbalanced jackknife. *Ann. Statist.* **2** 880–891.

Shao, J. (1986). On resampling methods for variance and bias estimation in linear models. Technical report, Department of Statistics, Univ. of Wisconsin-Madison.

Shao, J. and Wu, C. F. J. (1985). Heteroscedasticity-robustness of jackknife variance estimators in linear models. Technical report 778, Department of Statistics, Univ. of Wisconsin-Madison.

Weber, N. C. and Welsh, A. H. (1983). Jackknifing the general linear model. *Austral. J. Statist.* **25** 425–436.

Department of Statistics
University of Wisconsin
Madison, Wisconsin 53706

Jeffrey S. Simonoff and Chih-Ling Tsai

New York University

We would like to congratulate the author on a very interesting paper, and discuss some issues arising from jackknifing nonlinear models (Section 8). Much of what is presented here is based on Simonoff and Tsai (1986); V is the $n \times p$ matrix of first partial derivatives of $f(\cdot)$ with respect to $\boldsymbol{\theta}$, while W is the $n \times p \times p$ array of second partial derivatives.

1. Alternative weighting schemes. The weighted jackknife originally suggested by Hinkley (1977) was applied to nonlinear models by Fox et al. (1980), and has the form

$$\hat{\theta}_{LQ} = \hat{\theta} + \sum_i (V'V)^{-1}\mathbf{V}_i r_i = \hat{\theta}.$$

This does not address the bias problem, naturally leading to the investigation of alternative weighting schemes. We have had success with a weighting that leads to

$$\hat{\theta}_{RLQ} = \hat{\theta} + \sum_i (1 - h_{ii})(V'V)^{-1}\mathbf{V}_i r_i,$$

where $h_{ii} = \mathbf{V}_i'(V'V)^{-1}\mathbf{V}_i$. Although this does not have the theoretical justification of the weights W_s of (4.1), it can be noted that the effect of any observation on $\hat{\theta}_{RLQ}$ is a decreasing function of h_{ii}, the leverage.

2. Observed versus expected Fisher information. Skovgaard (1985) showed that for general parametric families the observed Fisher information is an estimate of the inverse variance in the conditional distribution of maximum likelihood estimates, given an ancillary statistic. Thus, the observed information is more meaningful than the expected information in assessing the precision of MLE's. In the nonlinear regression framework, the expected information corresponds to $V'V$ (this is Wu's suggestion for nonlinear models), while the observed information is $V'V - [\mathbf{r}'][W]$ (the bracket multiplication is defined in Bates and Watts (1980); the term is a function of the intrinsic curvature array). An unweighted jackknife approach based on this observed information was termed a "modified" jackknife by Simonoff and Tsai (1986) and, when combined with the alternatively weighted parameter estimate mentioned earlier, was found to be robust with respect to outliers, unbalanced designs and curvature effects.

3. Outliers. Wu has extensively investigated violation of the constant variance assumption, but has not studied the effect of an outlier. Table 1 outlines the results of a small simulation study based on the quadratic regression model analyzed in Section 10, $y_i = \beta_0 - 2\theta\beta_2 x_i + \beta_2 x_i^2 + e_i$, $i = 1(1)12$. The errors $e_1, \dots, e_{11}$ are $N(0,1)$, with e_{12} being $N(0, \sigma^2)$, $\sigma = 1$ (no outlier) or 3 (outlier). There were 1000 simulation replications, and nominal coverage rates were all 95%. Note that (as Wu points out) $\theta = 8$ corresponds to a high curvature model. The entry LQ refers to the Fox et al. (1980) weighted jackknife, $J(1)$ to Wu's delete-one jackknife, $J(1)M$ to the delete-one jackknife with weights based on observed, rather than expected, information, and RLQM to the reweighted and modified jackknife proposed in Simonoff and Tsai (1986).

Several points are apparent from this table. It can be noted that (as the theoretical results reported by Wu for linear regression suggest) confidence regions based on $J(1)$ have uniformly better coverage than ones based on LQ. Further, with respect to both bias and coverage, $J(1)M$ is an improvement over $J(1)$ (and outperforms the MLE), indicating the importance of using observed rather than expected information (the entries * are due to the variance sometimes not being positive definite under severe curvature and an outlier). The

TABLE 1
Biases and confidence region coverage levels for quadratic regression model (nominal coverage 95%).

	Bias			Coverage			Coverage
	β_0	θ	β_2	β_0	θ	β_2	$(\beta_0, \theta, \beta_2)$
		(1) $\beta_0 = 0$, $\theta = 8$, $\beta_2 = -0.25$; no outlier					
MLE	−0.00266	0.07541	−0.00012	89.9	88.9	89.9	76.1
LQ	−0.00266	0.07541	−0.00012	85.3	77.4	83.5	55.2
$J(1)$	0.28099	0.16291	0.01394	88.1	85.6	87.7	61.4
$J(1)M$	0.05570	−0.17410	0.00441	89.6	86.6	88.3	56.4
RLQM	−0.00093	0.07661	0.00008	96.4	94.4	96.5	79.4
		(2) $\beta_0 = 0$, $\theta = 8$, $\beta_2 = -0.25$; outlier					
MLE	−0.03359	0.45568	−0.00166	82.9	65.8	65.2	44.5
LQ	−0.03359	0.45568	−0.00166	77.7	55.6	58.3	31.6
$J(1)$	0.52607	0.45261	0.03350	81.6	69.7	74.1	52.7
$J(1)M$	(*)	(*)	(*)	83.5	72.1	79.5	59.1
RLQM	−0.03037	−0.05202	−0.00154	92.1	80.0	85.1	54.2

most effective approach, however (being robust to both curvature and an outlier), is RLQM. The poor results for simultaneous confidence regions are due to severe nonlinearity.

REFERENCES

BATES, D. M. and WATTS, D. G. (1980). Relative curvature measures of nonlinearity. *J. Roy. Statist. Soc. Ser. B* **42** 1-25.

FOX, T., HINKLEY, D. and LARNTZ, K. (1980). Jackknifing in nonlinear regression. *Technometrics* **22** 29–33.

HINKLEY, D. V. (1977). Jackknifing in unbalanced situations. *Technometrics* **19** 285–292.

SIMONOFF, J. S. and TSAI, C. L. (1986). Jackknife-based estimators and confidence regions in nonlinear regression. *Technometrics* **28** 103–112.

SKOVGAARD, I. M. (1985). A second-order investigation of asymptotic ancillarity. *Ann. Statist.* **13** 534–551.

STATISTICS AND OPERATIONS RESEARCH AREA
GRADUATE SCHOOL OF BUSINESS ADMINISTRATION
NEW YORK UNIVERSITY
100 TRINITY PLACE
NEW YORK, NEW YORK 10006

DEPARTMENT OF MATHEMATICS
THE UNIVERSITY OF TEXAS AT AUSTIN
AUSTIN, TEXAS 78712

KESAR SINGH

Rutgers University

I congratulate Professor Wu for this important contribution on resampling procedures for regression analysis. The representations reported in Section 3 are