(2), $\hat{\theta}_{\mathrm{ADJ1}(\psi)}(\alpha)$ is *third-order* equivalent to both $\hat{\theta}_{\mathrm{STUD}}(\alpha)$ and $\hat{\theta}_{\mathrm{ABC}}(\alpha)$ in the sense that their expansions match up to and including the term of order $n^{-3/2}$.

REFERENCES

ABRAMOVITCH, L. and SINGH, K. (1985). Edgeworth corrected pivotal statistics and the bootstrap. *Ann. Statist.* **13** 116–132.

HALL, P. (1983). Inverting an Edgeworth expansion. *Ann. Statist.* **11** 569–576.

HALL, P. (1986). On the bootstrap and confidence intervals. *Ann. Statist.* **14** 1431–1452.

LOH, W.-Y. (1984). Estimating an endpoint of a distribution with resampling methods. *Ann. Statist.* **12** 1543–1550.

LOH, W.-Y. (1987). Calibrating confidence coefficients. *J. Amer. Statist. Assoc.* **82** 155–162.

DEPARTMENT OF STATISTICS
UNIVERSITY OF WISCONSIN
1210 WEST DAYTON STREET
MADISON, WISCONSIN 53706

J. ROBINSON

University of Sydney

My major comment concerns the relative importance of the approximation of the critical points and of coverage error. It appears to me that much greater emphasis should be placed on the accuracy of the approximation of the bootstrap critical points to the theoretical points. The theoretical critical points based on θ should have been chosen as the best ones, in the nontechnical sense that they are thought to be better than any others available and the interval based on these has exact coverage α. Then, because we are not, in fact, able to find these critical points, we need an approximation; this can be provided by the bootstrap. Then we need to examine first the closeness of the approximating confidence interval to the theoretical one which we would use if we could. Finally, the coverage error for the approximation can be examined.

The point is more strongly made with reference to bootstrap simulations, which are not directly referred to here, the assumption being in this work that the number of simulations B, say, is very large. There is a discussion of them in Hall (1986), where it is shown that even for small B, say 19 for a 95% one-side interval, the coverage error is of the same order as if we simulated an infinite number of times. However, the accuracy of the approximation to the simulated critical point in the Studentized case is of order $n^{-1/2}B^{-1/2}$ compared to an accuracy for the infinitely resampled bootstrap of order $n^{-3/2}$. So B must be at least of size n^2 to make these approximations comparable. The reason for the accuracy of the coverage is that an averaging over all possible bootstrap simulations of size B has taken place in its calculation. Thus the particular approximation based on B simulations, compared to the real bootstrap approximation, may be gravely in error, although an average of these errors, taken over an inappropriate set, is small. I believe that in this case it is apparent that the accuracy