Bootstrap estimate of standard errors

We use nonparametric bootstrap to estimate standard errors for subgroup estimates. Let $W_1, \ldots, W_n$ be the observed values in the data. We take a random sample of size n from this data, with replacement. Denote this bootstrap sample by $W_{b1}, \ldots, W_{bn}$. We now compute our estimates using the bootstrap sample:

$$\hat{\theta}_b^* = g(w_{b1}, \ldots, w_{bn}), \tag{1}$$

where $b = 1, \ldots, B$ denotes the bootstrap samples and $\hat{\theta}_b^*$ is the bth set of parameters for each subgroup. This leads to a collection of B bootstrap estimates, $\hat{\theta}_1^*, \ldots, \hat{\theta}_B^*$. The bootstrap covariance matrix is

$$S_B = \frac{\sum_b (\hat{\theta}_b^* - \hat{\bar{\theta}}^*)(\hat{\theta}_b^* - \hat{\bar{\theta}}^*)'}{B-1}, \tag{2}$$

where $\hat{\bar{\theta}}^* = \sum_b \hat{\theta}_b^* / n$. S_B is the bootstrap estimate of $Cov(\hat{\theta})$, the covariance matrix of $\hat{\theta}$.

Confidence intervals and p-values are obtained by using the so-called percentile method, which is a nonparametric. Reported p-values are based on the empirical distribution of bootstrap estimates, that is, the proportion of estimates that lie further in the tails than the actual estimate. This method yields the confidence interval

$$[\theta^*_{\alpha/2}, \theta^*_{1-\alpha/2}], \tag{3}$$

where θ^*_p it the pth quantile (i.e. the 100pth percentile) of the bootstrap distribution $(\hat{\theta}_1^*, \ldots, \hat{\theta}_B^*)$.