

Comments on the Manuscript
“Various Versatile Variances: An Object-Oriented Implementation
of Clustered Covariances in R”
(JSS 3214)

The manuscript addresses a topic that is very relevant for many applied statisticians and econometricians—covariances for analyses of data with clustered observations—and describes a software package—the R package **sandwich**—that implements the various approaches that are described in the manuscript. The software package is implemented and documented very well, while the manuscript is mostly well written but lacks some clarity. In the following, I will give suggestions for improving the manuscript:

1. In equation (2), the definition of the term “ $\psi'(y, x, \theta)$ ” unclear. I guess that this term should indicate $\partial\psi(y, x, \theta)/\partial\theta = \partial^2\Psi(y, x, \theta)/\partial\theta^2$ and I recommend to replace “ $\psi'(y, x, \theta)$ ” by an expression that readers can easily understand (e.g., “ $\partial\psi(y, x, \theta)/\partial\theta$ ” or “ $\partial^2\Psi(y, x, \theta)/\partial\theta^2$ ”).
2. The matrix defined in equation (5) is slightly unclear. I guess that the terms “ $\psi(y_i, x_i, \hat{\theta})$ ” ($i \in \{1, n\}$) in equation (5) indicate *row*-vectors. As vectors are usually assumed to be column-vectors (unless stated otherwise) and $\psi(y_i, x_i, \hat{\theta})$ seem to be column-vectors in equation (7), I recommend to add transformation signs to the two vectors in equation (5) so that they become “ $\psi(y_i, x_i, \hat{\theta})^\top$ ”.
3. It seems to me that equation (12) is incorrect, because this equation is basically identical to equation (7) and, thus, does not take into account clustering (as it simply takes the sum over all observations). I guess that the correct equation is something like:

$$\hat{M}_{CL} = \frac{1}{n} \sum_{g=1}^G \left(\sum_{i=1}^{n_g} \psi(y_{ig}, x_{ig}, \hat{\theta}) \right) \left(\sum_{i=1}^{n_g} \psi(y_{ig}, x_{ig}, \hat{\theta}) \right)^\top \quad (1)$$

4. The “hat matrix H ” that is mentioned on page 6 is not defined in the manuscript. I recommend that the authors define this matrix in the manuscript.
5. On page 7, several parts in the paragraph that starts with “Petersen (2009)” are unclear to me. I recommend that the authors reformulate and perhaps extend this paragraph in order to make it easier understandable.
6. It is not clear to me which of the approaches that are described in sections 3.3 and 3.4 of the manuscript (and that are implemented in the **sandwich** package) can deal with

clustering of individuals in panel data sets, e.g., if the panel data set includes data from T time periods for n individuals that are ‘clustered’ in G groups of individuals. Please clarify this in the manuscript.

7. In equation (21), it is unclear to me how this approach takes into account the time dimension of the panel data set (as none of the terms in this equation has a subscript t). Furthermore, as subscripts i and j seem to indicate individuals (or observations?), it seems to be weird that the ordering of the individuals (or of the observations) affects the weights that account for ‘correlation’ between each pair of individuals (or observations). Please clarify this in the manuscript.
8. In equation (22), similarly to equation (21), it seems to be weird that the ordering of the individuals affects the weights that account for ‘correlation’ between each pair of individuals. Please clarify this in the manuscript.
9. In section 3.4, the use of subscripts for individuals, groups, and time periods is confusing and perhaps inconsistent. I recommend that the authors considerably revise and extend this section so that it is less confusing and easier to understand.
10. On page 11, the closing parenthesis behind “NW1994” needs to be removed.
11. Equation (25) seems to be inconsistent with equations (29) and (30), because β_1 and β_2 seem to be scalars in equation (25), while they seem to be vectors in equations (29) and (30). This needs either to be corrected or—if it is correct—more clearly explained.
12. In Figures 1 and 2, the symbols for “standard” and “basic” (red and pink triangles) as well as the symbols for “random” and “gee” (violet and red plus signs) look very similar and cannot be easily distinguished from each other. I recommend to use clearly different symbols (e.g., dark red triangle, orange up-side-down triangle, black plus sign, grey asterisk).
13. In the captions of Figures 1, 2, and 3, the meaning of the symbol “ ρ ” is unclear. Are there typos and the “ ρ ”s should be in fact “ ρ_x ”s or do the “ ρ ”s indicate something else? This needs to be correct or explained.
14. Several parts of the manuscript are not easily comprehensible for people who have limited prior knowledge about ‘sandwich’ methods and clustering in general. I highly recommend that the authors revise and extend their explanations so that the manuscript can be more easily understood by practitioners who want to apply the methods that are described in the manuscript and are implemented in the **sandwich** package.