# Review of "cv: An R Package for Cross-Validating Regression Models"

This package provides generic methods for performing cross-validation across linear models, GLMs, mixed models, and more. It implements various CV schemes, utilising fast computations where possible, as demonstrated by the example in Section 2 of the paper. The paper is well written, informative and highly pedagogical. Clearly, the methods offered by the package are widely useful. Moreover, the detailed explanations and examples make the paper an excellent software manual—I particularly appreciated the section explaining how to perform CV for mixed models via standard or clustered CV. The only drawback is that the paper is rather lengthy, though I do not see this as a significant issue. My additional comments are as follows:

1. I was surprised to see references to alternative packages only at the end of the paper rather than in the introduction. It might be beneficial to include a sentence in the introduction to clarify this.

2. What is the purpose of using `insight::get_data` on page 5? The `insight` package is also mentioned on page 31. I believe it would be helpful to include a brief explanation of why it is being used.

3. Does the package allow for the addition of an uncertainty interval to the plots in Figure 2? I think that including a measure of uncertainty would be useful.

4. In Section 4.1, six pages of the text are devoted to transformations before the actual CV content is addressed. Would it be possible/appropriate to shorten the discussion on transformations?

5. On page 21, when you state "We proceed to transform the numeric predictors in the Auto regression towards multivariate normality", I assume that the transformations being applied are marginal rather than joint, as joint transformations would cause the covariates to lose their interpretation. If so, then perhaps it would be worth clarifying that `powerTransform` adjusts the margins to achieve multivariate Gaussianity.

6. On page 29, when computing `mse(Auto$mpg, exp(fitted(m.step)))`, I think that you could use the relationship between the normal and log-normal distributions to correct for bias.

7. Finally, the computational section in Section 7 might be better placed in an appendix. It seems somewhat odd to conclude the paper with these details.