

## Comments on the p-Value Debate and Good Statistical Practice

Joachim Aurbacher <sup>1</sup> , Enno Bahrs <sup>2</sup> , Martin Banse <sup>3</sup> , Sebastian Hess\* <sup>2</sup> ,  
Stefan Hirsch <sup>2</sup> , Silke Hüttel <sup>4</sup> , Uwe Latacz-Lohmann <sup>5</sup> , Oliver Mußhoff <sup>4</sup> ,  
Martin Odening <sup>6</sup>  and Ramona Teuber <sup>1</sup> 

<sup>1</sup> Justus Liebig Universität Giessen

<sup>2</sup> University of Hohenheim, Stuttgart

<sup>3</sup> Thünen Institute, Braunschweig

<sup>4</sup> Georg-August-Universität Göttingen

<sup>5</sup> Christian-Albrechts-Universität zu Kiel

<sup>6</sup> Humboldt-Universität zu Berlin

\*Correspondence: Sebastian Hess, [s.hess@uni-hohenheim.de](mailto:s.hess@uni-hohenheim.de)

Aspects of common publication practice in scientific journals relating to articles that rely on statistical methods and inference have been the subject of critical debate. Key parts of this debate concern the reproducibility of published empirical results, their robustness and unbiasedness, and inferential conclusions concerning the targeted population (e.g. Heckeley et al., 2023).

While it ought be beyond question that scientific journals i) should minimise incentives to commit malpractice as a means to boost the chances of publication, and ii) should not tolerate insufficient transparency, biasedness or inaccuracies with regard to statistical inference, there is no consensus as to how this might best be achieved in practice.

Ultimately journal editors, authors, reviewers and the professional society share responsibility for the accuracy and quality of statistical analyses in published articles.

Possible measures and remedies that have been discussed in this context include:

- ensuring as much transparency as possible with regard to the data, programming code and software used
- distinguishing clearly between effect size and the precision of estimates, with clear reporting of the targeted population for inferential conclusions and potential biases from non-random sampling
- providing greater recognition and improved publication opportunities for scientific replication studies that review and verify empirical findings that have already been published in other studies
- undertaking unbiased assessments of empirical findings by reviewers and editors: even statistically insignificant results can provide relevant insights and should not be discarded per se

- giving consideration again to the widely accepted use of the concept of statistical significance testing (in particular, null hypothesis statistical testing (NHST)), according to some scholars.

Members of the Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaues e.V. (GeWiSoLa, Germany's society of agricultural economics and social sciences) and members of this journal's Editorial Board (e.g. Hirschauer et al., 2022) have in the past repeatedly emphasised the importance of the above issues for this and other agricultural economics journals. In addition to discussions by the Editorial Board, the following have also recently taken place:

- a session on the p-value debate at the 2018 annual GeWiSoLa conference
- the establishment of a working group on the use of p-values and statistical practice in agricultural economics
- a session at the 2020 annual GeWiSoLa conference
- a broad survey of the agricultural economics profession about the perception of statistical malpractice and potential remedies
- a pre-conference workshop at the 2021 annual GeWiSoLa conference
- participation in a (virtual) session at the 2021 meeting of the International Association of Agricultural Economists (IAAE)
- the dissemination of a working paper with an open review
- the publication of a paper by Heckeley et al. (2023), which was a key outcome of the working group
- integration of these topics into the training of masters and doctoral students by members of this editorial board and the professional society.

Within our Editorial Board, there is agreement on the following:

- researchers should check preconditions for the application of statistical inference, particularly whether data are random samples of a population, and should discuss the implications when some of these assumptions are violated in their analysis, e.g. due to non-random (convenience) sampling
- researchers should consider whether using p-values is meaningful or whether they are dispensable and should be substituted or complemented by other statistical tools, and in any event p-values should not be overinterpreted
- the presentation and interpretation of statistical results should also involve the economic relevance of variables rather than focusing solely on their statistical significance, and interpretation should clearly differentiate between these two approaches
- the GJAE shall ensure there is a review process for submissions with statistically insignificant results that is just as objective and open-ended as all other submissions; results that are not statistically significant are not per se a reason to reject a submission
- the GJAE is open to the submission of replication studies
- the GJAE strives for the greatest possible transparency with regard to the data, programming code and software used, and encourages authors to disclose these materials during the review process and to publish them as digital attachments.

As the Editorial Board of the GJAE, we would like to encourage authors and reviewers to engage in this debate. In relation to this, we recommend the article by Heckeley et al. (2023) for its overview of the topic and references to other literature.

However, we also acknowledge that there are likely to be counter-arguments and objections to our position and that the discussion will continue. We therefore invite our readers to publish their own views about the debate in this journal.

## References

- Heckelei, T., Hüttel, S., Odening, M. & Rommel, J. (2023): The p-Value Debate and Statistical (Mal) practice—Implications for the Agricultural and Food Economics Community. *German Journal of Agricultural Economics* 72 (1): 47-67. <https://doi.org/10.30430/gjae.2023.0231>.
- Hirschauer, N., Grüner, S. & Mußhoff, O. (2022): *Fundamentals of Statistical Inference: What is the Meaning of Random Error?* Springer, SpringerBriefs in Applied Statistics and Econometrics, Cham. <https://doi.org/10.1007/978-3-030-99091-6>.

## Contact author

Prof. Dr. Sebastian Hess  
Department of Agricultural Policy and Market Analysis, University of Hohenheim  
Schwerzstrasse 46, 70593 Stuttgart, Germany  
email: [s.hess@uni-hohenheim.de](mailto:s.hess@uni-hohenheim.de)