UNIVERSITY OF COPENHAGEN

Comment on: "Testing Hypotheses in Non-parametric Models of Production" by Kneip, Simar, and Wilson (2016, JBES)

Asmild, Mette; Henningsen, Arne; Kronborg, Dorte; Rønn-Nielsen, Anders

Publication date:
2019

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Asmild, M., Henningsen, A., Kronborg, D., & Rønn-Nielsen, A. (2019). *Comment on: "Testing Hypotheses in Non-parametric Models of Production" by Kneip, Simar, and Wilson (2016, JBES)*. Department of Food and Resource Economics, University of Copenhagen. IFRO Working Paper , No. 2019/07

DEPARTMENT OF FOOD AND RESOURCE ECONOMICS
UNIVERSITY OF COPENHAGEN

IFRO Working Paper

Comment on: "Testing Hypotheses in Nonparametric Models of Production" by Kneip, Simar, and Wilson (2016, JBES)

Mette Asmild
Arne Henningsen
Dorte Kronborg
Anders Rønn-Nielsen

2019 / 07

IFRO Working Paper 2019 / 07

Comment on: "Testing Hypotheses in Non-parametric Models of Production" by Kneip, Simar, and Wilson (2016, JBES)

Authors: Mette Asmild[1], Arne Henningsen[1], Dorte Kronborg[2], Anders Rønn-Nielsen[2]
[1] Department of Food and Resource Economics, University of Copenhagen
[2] Center for Statistics, Department of Finance, Copenhagen Business School

JEL-classification: C12, C14, C44, C61, D24

Published: October 2019

See the full series IFRO Working Paper here:
www.ifro.ku.dk/english/publications/ifro_series/working_papers/

Department of Food and Resource Economics (IFRO)
University of Copenhagen
Rolighedsvej 25
DK 1958 Frederiksberg DENMARK
www.ifro.ku.dk/english/

Comment on: "Testing Hypotheses in Nonparametric Models of Production" by Kneip, Simar, and Wilson (2016, JBES)

Mette Asmild[1], Arne Henningsen[1], Dorte Kronborg[2,*], and Anders Rønn-Nielsen[2]

[1] Department of Food and Resource Economics, University of Copenhagen, Denmark

[2] Center for Statistics, Department of Finance, Copenhagen Business School, Denmark

[*] Corresponding author: dk.mes@cbs.dk

Keywords: DEA, Returns to scale, Hypothesis test, Inference

JEL-codes: C12, C14, C44, C61, D24

In Kneip et al. (2016) a series of tests for important hypotheses when nonparametric efficiency studies are performed are suggested. As the authors state, it is essential to validate the assumptions made, e.g., assumptions on returns to scale and/or convexity, before concluding on real-world empirical studies.

To ensure that their proposed test statistic for returns to scale does not have a degenerate asymptotic distribution under the null hypothesis, the authors suggest to randomly split the sample into two samples. However, this initial random split introduces uncontrollable randomness, as the resulting p–values are highly dependent on the specific initial split. Thus, very different conclusions can be obtained for a give data set depending on this randomness.

Specifically, in an analysis carried out for the benchmark regulation of electricity distributors in Denmark, the test for returns to scale, i.e., the test statistic in equation 39 in Kneip et al. (2016), has been applied to a data set with 28 companies, 3 output variables and 1 input variable (total cost). We have performed the test 10 000 times each with a unique execution of choosing the random split. The resulting p–values range from 0 to 1 and have the empirical distribution illustrated in Figure 1.

This clearly highlights the randomness of the p–values due to the initial split in the given data set. This problem is not easily overcome. One solution could have been to combine results from more than one initial split, but the joint distribution of test statistics across different splits is not obtainable. Kneip et al. (2016) suggest to base inference on bootstrapped confidence limits but the effect of the randomness of the initial split is undoubtedly inherited in this method. We suggest to use the permutation based

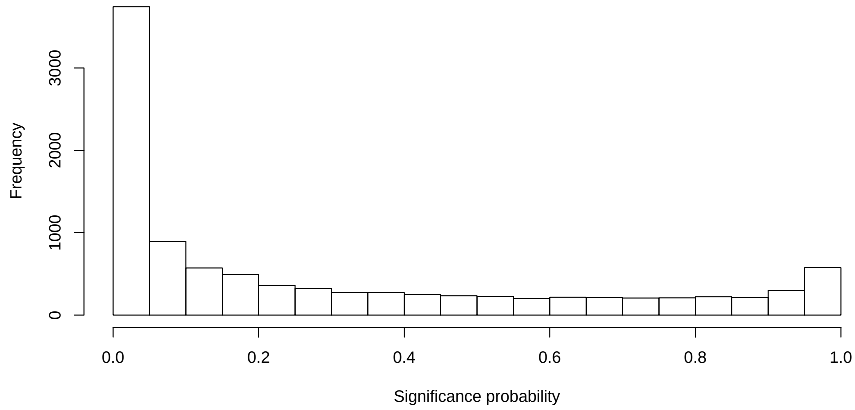

Figure 1: Distribution of p–values based on 10 000 random splits

test for returns to scale introduced in Rønn-Nielsen et al. (2019), which does not suffer from the described weakness.

To investigate the randomness of the p–values created by the initial random split in situations with more observations than in the electricity example, we have simulated a number of data sets following simulation procedure 1 in Rønn-Nielsen et al. (2019). In the simulations, the parameter γ measures the departure from the CRS hypothesis. When $\gamma = 1$, the hypothesis is true, while decreasing the value will give an increasing departure from the hypothesis. For each of the parameter values $\gamma = 0.96$ and 1 we have generated data sets following the algorithm with $n = 200$ and 1000 observations. For each of the four data sets we have performed the test 1000 times. The resulting p–values are illustrated in the histograms in Figure 2.

The upper left histogram in Figure 2 corresponds to a data set with 200 observations and with $\gamma = 1$, such that the CRS hypothesis is actually true. Here we see the same pattern as in the previous example: Given the data set, the p–value exhibits a very high degree of randomness due to the initial split. This is also seen in the lower left histogram that represents a data set with $n = 1000$, where the CRS hypothesis is true. Also here, all values of the p–value are possible outcomes.

The upper right histogram of Figure 2 shows the different p–values for a dataset with $n = 200$ observations and with $\gamma = 0.96$. Thus, here the CRS hypothesis is incorrect and we consequently see that the p–value is more often relatively small. However, again very different outcomes are possible. This pattern is to a large extent repeated when the number n of observations is increased to 1000, as depicted in the lower right histogram: Though the p–value is most often rather small, there is a non–negligible chance of observing a very large value.

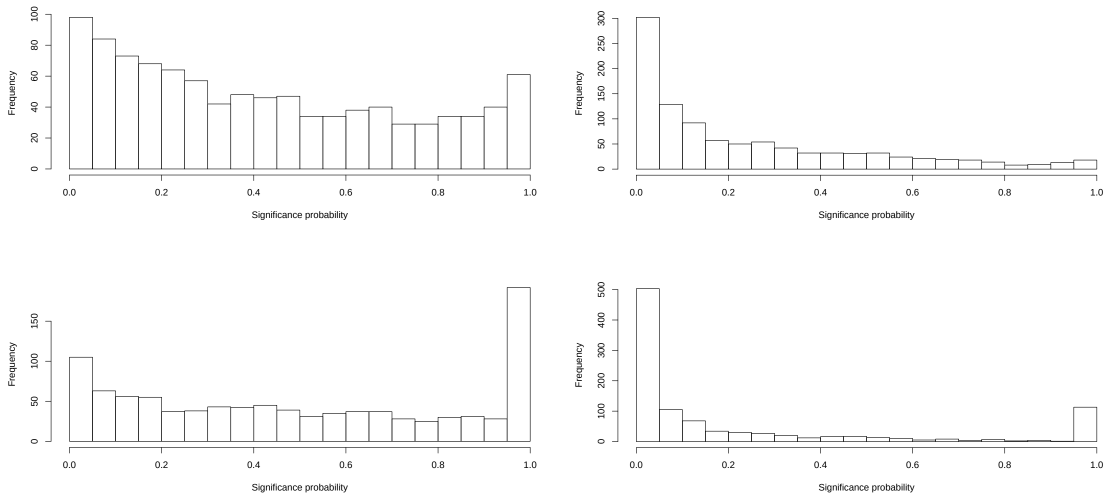

Figure 2: Simulated p–values based on 1000 random splits for all combinations of $\gamma = 0.96, 1$ and $n = 200, 1000$. The first row represents $n = 200$, while $n = 1000$ in the second. The first column represents $\gamma = 1$, and $\gamma = 0.96$ in the second column.

In conclusion, the test for CRS proposed by Kneip et al. (2016) produces p–values that for a given data set exhibit way too much randomness, and thereby leaves the researcher with a very unreliable basis to draw conclusions upon.

References

Kneip, A., Simar, L., and Wilson, P.W. (2016). Testing Hypotheses in Nonparametric Models of Production. *Journal of Business & Economic Statistics* 34, 435–456.

Rønn-Nielsen, A., Kronborg, D., and Asmild, M. (2019). Exact tests on return to scale and comparisons of production frontiers in nonparametric models. Department of Food and Resource Economics, University of Copenhagen, Frederiksberg, Denmark. IFRO Working Paper 2019/04.