

Winker, Peter

Article

Comment to "Bielefeld may in fact not exist - Empirical evidence from official population data" by Patrick Winter

Journal of Economics and Statistics

Provided in Cooperation with:

De Gruyter Brill

Suggested Citation: Winker, Peter (2023) : Comment to "Bielefeld may in fact not exist - Empirical evidence from official population data" by Patrick Winter, Journal of Economics and Statistics, ISSN 2366-049X, De Gruyter Oldenbourg, Berlin, Vol. 243, Iss. 1, pp. 39-41, <https://doi.org/10.1515/jbnst-2023-2001>

This Version is available at:

<https://hdl.handle.net/10419/333267>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>

Comment

Peter Winker*

Comment to “Bielefeld May In Fact Not Exist – Empirical Evidence From Official Population Data” (DOI <https://doi.org/10.1515/jbnst-2022-0038>) by Patrick Winter

<https://doi.org/10.1515/jbnst-2023-2001>

At first sight, the paper uses a classical approach from empirical economic research. It starts with a theory based null hypothesis, which is the complement of what should be demonstrated and which is challenged by the data. Given that this null hypothesis is clearly rejected, the author concludes that the alternative is the most likely explanation of this finding.

Unfortunately, the specific application combines pitfalls regarding different parts of the analysis in a way, which render the final conclusion unsustainable. I just mention three major objections:

1. The author states that the selection of a specific city is not at random, but derived from the theory of the so called “Bielefeld conspiracy”. This has the major advantage that – in contrast to a testing of many random cities – the result cannot be easily wiped aside as a realization of the error of first kind to which any statistical analysis is subject. However, for this argument to hold true, one would need first, a theory, i.e. at least a reference to this theory, and second, the conviction that no inverse causality might be present, i.e. that doubts about the existence of Bielefeld did not result from a special distribution of population numbers.
2. Let us assume for the moment being that this first issue can be addressed in a meaningful way. Then, in a next step, some conclusions are drawn from the null hypothesis adding two further central assumptions, namely that an existing city would report population numbers without bias and that such population numbers at the statistical district level have to follow Benford’s law. If at least one of these assumptions does not hold true, a rejection of the combined

***Corresponding author: Peter Winker**, Faculty of Economics and Business Studies, Justus-Liebig-University Giessen, Licher Str. 64, 35394 Giessen, Germany, E-mail: peter.winker@wirtschaft.uni-giessen.de

hypothesis does not allow for any conclusion regarding the initial hypothesis. Thus, I will argue in the following why the additional assumptions might not hold.

- a. For an existing city there is an obvious incentive to report high population numbers as these are positively linked to participation in tax income. This incentive does not change over time. Therefore, the argument that the same results are found for different years does not exclude such a bias in reporting, which obviously would also have to show up at the disaggregated level of statistical units. Looking at the argument from the alternative, i.e. for the case when the city does not exist, one would expect that numbers are invented in order to resemble as closely as possible those of existing numbers. Since population numbers for states, regions and cities are known to often follow Benford's law, one would expect that fabricated population numbers for statistical districts would also follow this distribution given that the Bielefeld conspiracy implies a lot of resources available for generating "good data".
- b. The second assumption concerns the distribution of first digits of population numbers. The author correctly refers to the literature stating that often one finds Benford's law to hold for population numbers of states, regions, and cities. However, he does not provide any reference to statistical districts within a city. As the "growing or shrinking of an area" (see third paragraph of Section 2) is not possible for these regions, the argument given in the literature for a Benford distribution does not hold true. Thus, we just do not know whether such data are likely to follow Benford's distribution. The fact, that this assumption cannot be rejected for other cities, does not imply that it is true. Finally, a simple phone call to the city of Bielefeld could provide some further insights into the data. In fact, the statistical districts of Bielefeld have been reorganized in 2017 following a requirement from the federal employment agency in order to allow for matching labor market data at the statistical district level. To this end, each statistical district has to contain at least 1000 inhabitants. Therefore, several smaller units had to be merged, while maintaining the special structure. Furthermore, statistical districts do rarely exceed 10,000 inhabitants. Consequently, the numbers cannot be assumed to be the result of an organic process of growing and shrinking, but is rather the result of a quite recent intervention. Nevertheless, even when using the data prior to the reorganization of statistical districts, the numbers still fail to follow closely the Benford distribution. However, the departures become much smaller.

To sum up: Proper empirical analysis requires a clear theoretical background. When testing joint hypotheses, a rejection does not provide clear insights onto which of the components of the joint hypothesis might be incorrect. Finally, when doing empirical research, it always pays off to look at the data, to look at the data, to look another time at the data, and then to ask the people producing the data about further details!