
**STI 2018 Conference Proceedings**

*Proceedings of the 23rd International Conference on Science and Technology Indicators*

All papers published in this conference proceedings have been peer reviewed through a peer review process administered by the proceedings Editors. Reviews were conducted by expert referees to the professional and scientific standards expected of a conference proceedings.

**Chair of the Conference**

Paul Wouters

**Scientific Editors**

Rodrigo Costas
Thomas Franssen
Alfredo Yegros-Yegros

**Layout**

Andrea Reyes Elizondo
Suze van der Luijt-Jansen

The articles of this collection can be accessed at https://hdl.handle.net/1887/64521

ISBN: 978-90-9031204-0

© of the text: the authors
© 2018 Centre for Science and Technology Studies (CWTS), Leiden University, The Netherlands


This ARTICLE is licensed under a Creative Commons Atribution-NonCommercial-NonDetivates 4.0 International Licensed

23rd International Conference on Science and Technology Indicators (STI 2018)
"Science, Technology and Innovation indicators in transition"
12 - 14 September 2018 | Leiden, The Netherlands
#STI18LDN

# A statistical approach to infer causal effects of the Chinese publication growth on bibliometric impact measures[1]

Stephan Stahlschmidt[*] and Sybille Hinze[**]

[*] *stahlschmidt@dzhw.eu*
Department 2 – Research System and Science Dynamics, German Centre for Higher Education Research and Science Studies (DZHW), Schützenstr. 6a, Berlin, 10117 (Germany)

[**] *hinze@dzhw.eu*
Department 2 – Research System and Science Dynamics, German Centre for Higher Education Research and Science Studies (DZHW), Schützenstr. 6a, Berlin, 10117 (Germany)

## Introduction

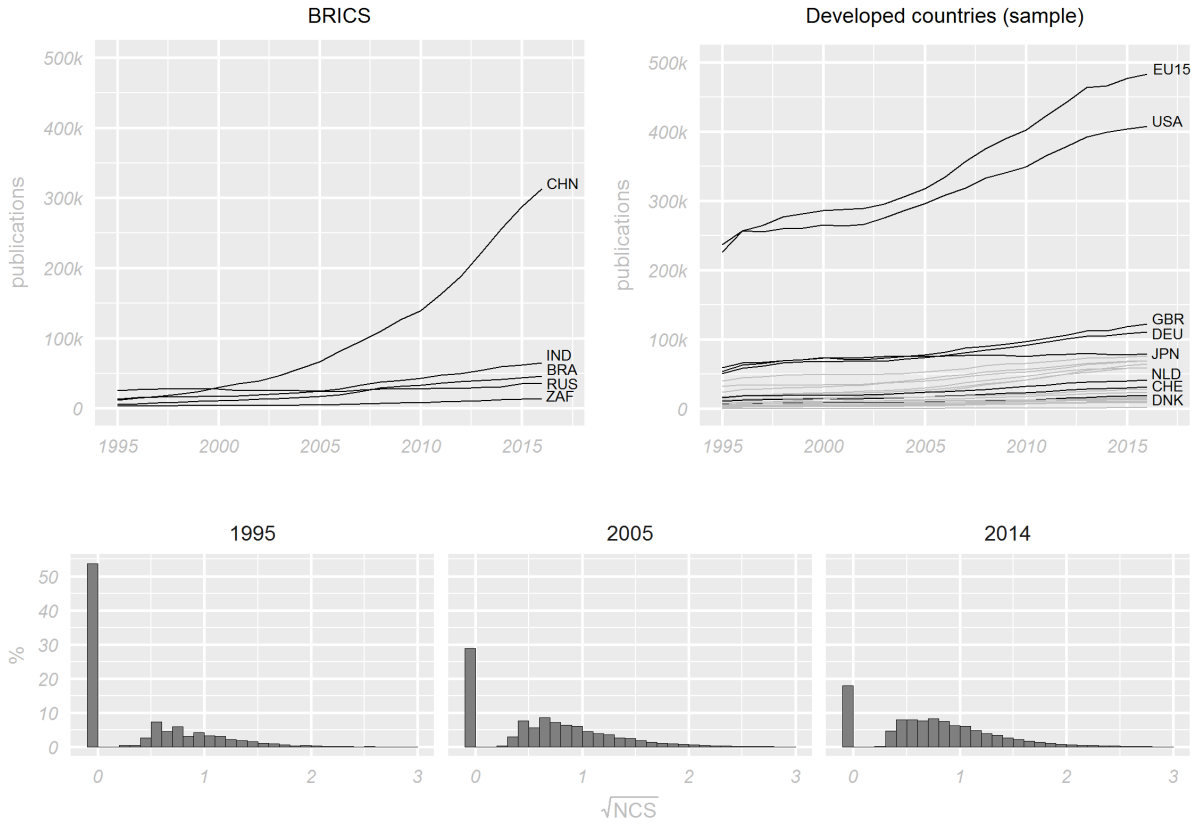
Since 2006 China counts as the second biggest single "producer" of publications trailing only the United States and consequently increasingly competes for space in *Web of Science* indexed journals[2]. The increase of Chinese publications is extensive and thus beyond what could be considered marginal. It rather affects the whole database. Furthermore it is stimulated by economic growth and political factors (Zhou and Leydesdorff, 2016) rendering it an external macro-level intervention on the science system. Consequently, the question arises, how this unprecedented growth of contributions from a single country with its specific bibliometric characteristics shown in Figure 1 affects bibliometric impact measures, which relate an entity's publications to the general publication universe. I.e. we analyse macro-level effects resulting from the Chinese publication growth attempting to explain how these effects stem from changes on the micro level of publications.

We approach these questions of causal inference from a statistical perspective and do not address wider implications like the true or ultimate cause. Instead we limit our scope to a statistically feasible task as we strive to measure an effect of a putative cause proposed by bibliometric measurement theory and express the findings via probability statements.

The pronounced investment of the Chinese state in the national science system is understood as a macro level intervention, which alters the evolution of the bibliometric universe. Resulting causal effects may be inferred by contrasting the actual and the counterfactual, i.e. a fictive setting without the aforementioned intervention, world. However, such a comparison is confronted with interfering factors, which arise either from the data collection process or from the subject matter of bibliometrics. Especially confounding mechanisms pose the main challenges to a statistical analysis of causation. They were at first mastered comprehensively by Roland Fisher's treatment of experiments (Fisher, 1935). Such randomized experiments

[1] This work was supported by German Federal Ministry of Education and Research (Grant 01PQ13001 and 01PQ17001)
[2] This paper is based on a more elaborated analysis detailing how the Chinese publication growth affects impact indicators of a large set of OECD countries (Stahlschmidt and Hinze, 2018) and former work of the first author (Stahlschmidt, 2015)

**Figure 1: Growth of national publications among BRICS countries (upper left panel), national publication growth of a sample of OECD countries (upper right panel) and yearly histograms (lower panel) of the normalised citation score for Chinese publications (Data is truncated in the upper tail)**




still denote the gold standard in the statistical analysis of causation, but may not be applied to the ubiquitous observational data in bibliometrics.

On such data the causal analysis relies crucially on the corresponding subject matter theory to model and consequently control the bias arising from confounding mechanisms. Given such subject matter theory and an observed correlation between two variables, e.g. a policy intervention and a bibliometric impact indicator, the statistical approach to causation excludes interfering factors and declares any remaining residual association as being causal. Consequently causation arises as an interpretation in statistics, because "causality is not proven so much as other possibilities are rejected" (Schield (1995), p. 3).

This design approach to causation can also be identified in randomized experiments, where randomization can be understood as a clever device to account at once for confounding mechanisms and limited domain knowledge. On observational data, many techniques on causal inference are based on a detailed model of the causal structure in the particular subject matter domain and are consequently termed structural models (e.g. Wright (1921), Haavelmo (1943), Sims (1980)).

Treatment effect models with their utilization of untreated units of the population (Rubin1974, Imbens2009) do not pose such strong requirements on available subject matter theory, as they require less information on the mechanism transferring the policy intervention into macro-level effects. However, given this smaller information set they might not be as precise as well-defined structural models.

Given these options we present a treatment effect approach, as the exact mechanism on how the Chinese publication increase affects bibliometric impact measures has not been established and consequently the necessary theory for a structural model has not been available.

**Constructing a counterfactual bibliometric world**

Our analysis is based on the *Web of Science* raw data[3] and in order to cover the apparent start of the rise of Chinese publications in the late 1990s we commence our analysis in 1995 and extend it to 2016 while applying a three-year citation window.

Author affiliation data was used to identify Chinese publications. We build our analysis upon the classical assumption, that any listed author has contributed an essential part to a publication and consequently assign any publication to the set of Chinese publications with at least one listed author affiliated to an institution residing in China.

Based upon this definition we quantify the effect of the Chinese publication increase on bibliometric impact measures by inferring what would have happened without them. Consequently we constructed a counterfactual bibliometric world without Chinese publications and contrasted this with the actual bibliometric universe allowing for an assessment of the effect Chinese publications exhibit on the bibliometric universe.

As stated before, this approach borrows from the treatment effect literature in Economics (Rubin (1974), Imbens and Wooldridge (2009)). Based upon observational data, treatment effects models infer if and how a treatment causally affects a target audience. Ideally these models compare the same observational units with and without the treatment on some outcome variable and declare any difference to denote a causal effect of the treatment. Obviously any unit can either be exposed or not be exposed to the treatment and a direct comparison on the same unit is infeasible (Holland, 1986). Consequently treatment effects models apply carefully constructed substitute comparisons exploiting the untreated units of the population. However, as the Chinese publications affect the whole *Web of Science* publications universe, no unaffected units are available, but have to be constructed artificially.

Hence, we recount citation links after excluding the aforementioned set of Chinese publications. Next these counts are applied to re-compute the *Web of Science Subject Categories* based expected citation counts and we subsequently compare each non-Chinese publication to these counterfactual statistics to obtain national Mean Normalized Citation Scores ($MNCS_{counterfactual}$).

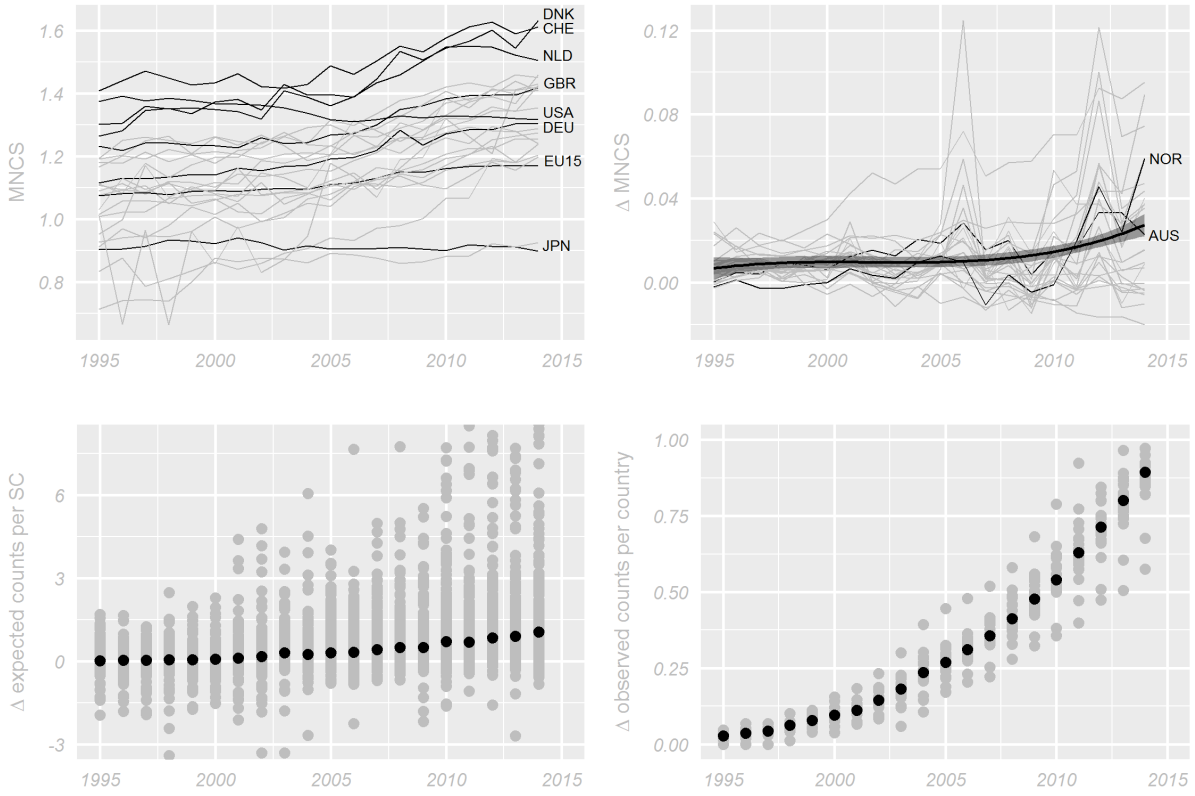
Finally these values are contrasted with the actual national MNCS for country $i$ via

$$\Delta\, MNCS^{(i)} = MNCS^{(i)}_{actual} - MNCS^{(i)}_{counterfactual}$$

where $\Delta MNCS^{(i)}$ quantifies how national bibliometric impact indicators are affected by Chinese publications.

[3] Data is provided by the German Competence Centre for Bibiliometrics (www.bibliometrie.info).

**Figure 2: Actual national MNCS values (upper left panel), changes in MNCS (upper right panel) and differences in in expected counts (lower left panel), respectively observed citations (lower right panel) due to the increase in Chinese publications**




This comparison between the actual and counterfactual impact indicators is facilitated by the particular nature of the MNCS, as it evaluates publications by comparing the respective citation counts with citations in a pre-determined environment. Consequently the resulting absolute decline in publications and citations imposed by the exclusion of Chinese publications might potentially affect both the obtained and the expected citations. However, these level shifts do not inhibit the aforementioned comparison if we assume that the actual and counterfactual worlds truthfully describe the complete national publication output in the respective settings. Consequently any difference between the actual and counterfactual national impact indicators results solely from the change in the environment driven by Chinese publications.

## Contrasting the factual with the counterfactual world

While the evolution of national MNCS for a set of OECD countries[4] is depicted in the upper left panel of Figure 2, the top right panel shows the effect China exhibits as it illustrates the resulting differences $\Delta MNCS^{(i)}$. Every grey line represents a single country $i$, while the black line denotes the average of these national values accompanied by a confidence interval. As most values and especially the average values are positive we conclude that the MNCS of the listed countries benefit from the Chinese publications. This effect, albeit rather small in the beginning, increases over time and affects the single countries to varying degree. As the MNCS relates obtained to expected citations, we depict in the lower left (right) panel of

[4] Countries shown in graph: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Great Britain, Greece, Ireland, Israel, Italy, Japan, Luxemburg, Netherlands, Norway, New Zealand, Portugal, South Korea, Spain, Sweden, Switzerland, United States of America

**Figure 3: Scatterplot of Subject Categories concerning their average reference list length and expected citation counts in 2014 (upper left panel), normalized length of Chinese reference lists (solid line) and normalized length of Chinese reference lists (dashed line) utilized in citation window (lower left panel), distribution of cited article over time from citing articles from 2014 (upper right panel) and stacked yearly count of utilized citations from non-standard document types (lower right panel). Besides upper left panel Chinese values are depicted in black, while grey values denote the sample-based, non-Chinese equivalents.**

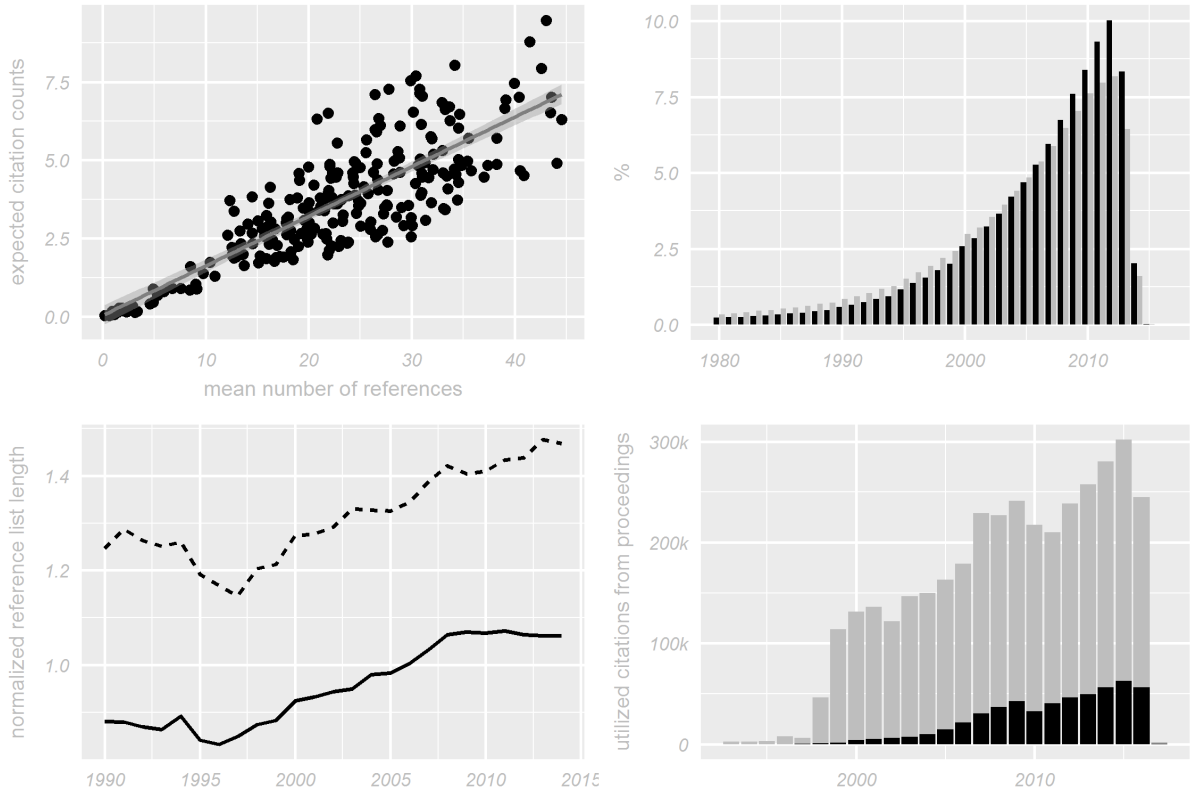


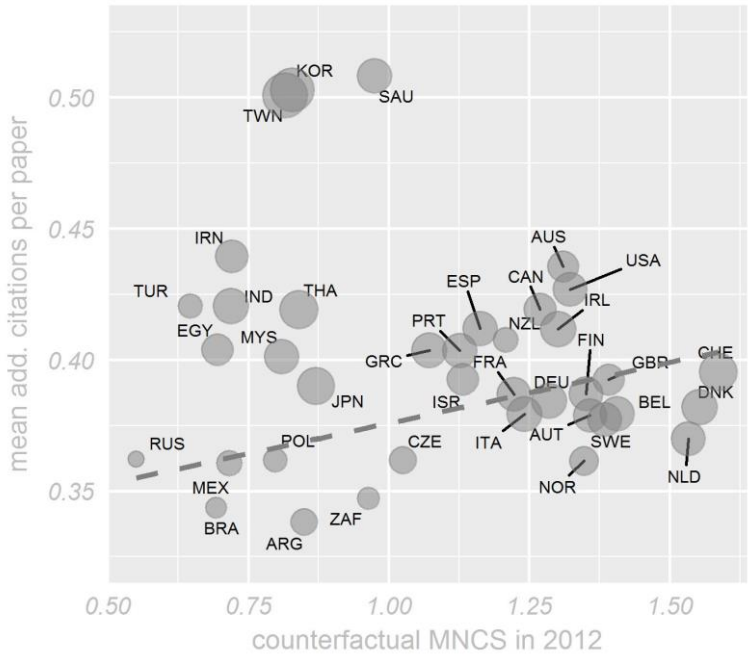

Figure 2 the increase in expected (obtained) citations counts resulting from the inclusion of Chinese publications (Stahlschmidt and Hinze, 2016) and detailing the computational mechanism.

However, these changes in the obtained, respectively expected citations represent empirical symptoms of the underlying mechanisms, which can be exemplified via a highly stylized bibliometric toy model. We omit for a moment any citations across time or disciplines and any influence non-source items might have. In such a perfectly encapsulated setting citations are distributed as a zero-sum game from reference lists and the expected citation count equals the average reference list length. Consequently any expansion in terms of additional publications might only increase (decrease) the expected citation counts if these additional publications include more (less) than usual references.

We depict the counterfactual reference list lengths and counterfactual expected citation counts for all *Web of Science Subject Categories* in the year 2014 in the upper left panel of Figure 3. A profound positive relation might be identified, which is, however, obscured by the frequent citation links across time and the less frequent citation links across disciplines. Furthermore an enlarged base of citing publications including also proceedings interferes as well.

The solid line in the lower left panel of Figure 3 depicts the average of normalized reference list lengths of Chinese publications. These lists have been shorter than the global average in the beginning of the observation period, but due to substantial growth around the millennium

**Figure 4: National difference between counterfactual national impact and average, normalized, additional citations from China in 2012 with correlation line for non-Asian countries. Countries receiving more than 1500 citations are abbreviated by three-letter codes. Singapore is omitted.**




change the average normalized Chinese reference list has exceeded the world average since 2006. However, the observed increase in expected citation counts begins much earlier. As the enlarging effect on expected citation counts can be observed uniformly across all *Subject Categories*, Chinese publication seem unlikely to differ strongly in their use of citations across disciplines. Contrary citations crossing the yearly time periods are common and the varying time focus of Chinese and non-Chinese publications can exemplarily be observed for citing publications from 2014 in the top right panel of Figure 3. Chinese publications focus much stronger on more recent publications, as their share of references to publications from 2007 to 2014 surpasses the non-Chinese shares and trails them for all preceding years.

But the applied three-year citation window curtails the count of relevant citations to cited publications not older than two years and consequently favours the Chinese focus on more recent literature. Accordingly the dashed line in the lower left panel in Figure 3 restricts the count to references within the three-year citation window. Thus Chinese reference lists exceed the non-Chinese ones in terms of citations utilized in the whole observation period. A further explanation of the increase in expected citation counts stems from our extended definition of the citing side, which is shown in the lower right panel of Figure 3. While comparably small in size a substantial growth of these citations can be identified.

In general any country will benefit from Chinese publications if the additional citations received will outweigh the rise in expected counts. Two factors might influence to what extent other countries will receive additional citations from Chinese publications. First an outward looking China, which cites foreign papers relatively more often than national publications. Second a non-uniform spread of these citations among countries.

As the share of Chinese national self-citations stays constant, we analyse how outward citations from Chinese publications are distributed among countries. Figure 4 graphically relates the additional Chinese citations received by other countries in 2012 to their general scientific impact expressed via the counterfactual MNCS and their publication output. In

detail the y-axis describes how many additional normalized citations a paper receives from Chinese publications on average once it is cited by a Chinese publication and the point size denotes the share of national publications being cited by Chinese publications. According to this graph Chinese publication cite especially the literature of scientifically leading countries as these obtain higher additional citations as well as a larger share of their publications being cited. However this relation is obscured by a strong regional focus in which many Asian countries obtain more citation than stipulated by this reasoning.

**Implications and limitations**

The resulting aggregate effects on impact indicators originate from changes on the micro scale of publications. Our definition of the counterfactual state resizes the set of publications and consequently any statistics on a macro level computed on this modified micro base of publications. However, the resulting effects might only be meaningful interpreted at the macro level, as on the micro level of publications we miss cited publications excluded in our counterfactual state. These missing knowledge claims render the identification of counterfactual actions by individual scientists infeasible and consequently the counterfactual state of single publications cannot be deduced. We also allow for an imaginative overcapacity in manpower, as non-Chinese co-authors of a joint Chinese paper are separated from their contributions.

However, as parallel discoveries testify scientific disciplines move forward as a whole, so it might be assumed, that the residual manpower in the counterfactual world would partially cover the missing cited publications. Furthermore national contributions to the knowledge set seem to be underpinned by robust steady trends. Hence, we assume that national impact indicators in the counterfactual world do depict a realistic picture on a macro scale and consequently validate the observed differences.

Still the constructed counterfactual bibliometric world does not constitute a perfectly known alternative setting but rather expresses a *Gedankenexperiement* – a thought experiment – which helps to understand the mechanism driving the evolution of the database and the resulting consequences for national impact indicators. For example in the unlikely case of Chinese publications displacing completely likewise publications, we would report purely artificial effects. Naturally our knowledge on the counterfactual state is limited, although it defines the subsequent effect size due to the applied contrast between both settings.

We also restrict our scope of analysis, as it focuses exclusively on publication metadata relevant to impact indicators. This rather technical approach explores a specific cause-effect relation proposed by bibliometric (measurement) theory, namely the relative nature of currently applied impact indicators. Consequently it forgoes a comprehensive analysis of further effects caused by the Chinese intervention. For example the applied definition of Chinese publications misses out on Chinese scientists working in non-Chinese institutions and does not explain the variation in the observed effect.

Crucially we also stipulate that no confounding takes places and the Chinese policy intervention constitutes an isolated, exogenous effect without a sizeable relation to other concurrent effects. As the Chinese state is predominately funding scientific institutions in China we assume no sizeable backdoor for a confounding mechanism. Only international coordinated research programmes might interfere, as the resulting joint author publications are assigned exclusively to the set of Chinese publications, although the cooperating non-

Chinese national science system also reallocates necessary funds. Compared to the overall spending we assume this eventual effect to be of minor importance.

However, causal inference arises as an interpretation in statistics and crucially depends on the particular definition of the counterfactual state. This observation not only holds for our analysis, but the obtained results actually impose restriction on other attempts to single out causal effects from policy interventions in bibliometric terms. As we have seen, the Chinese publication growth is large enough to structurally affect the whole bibliometric universe and every included country. Therefore it constitutes a confounding mechanism for further evaluation exercises, as its effect on bibliometric entities varies and must be controlled for in subsequent counterfactual comparisons. In general the publication universe is concurrently affected by several structural effects and this unsteady base complicates on a micro and macro level the analysis of any particular cause-effect relation of interest.

**References**

Fisher, R.A. (1935). The Design of Experiments Edinburgh: Oliver and Boyd.

Haavelmo, T. (1943). The statistical implications of a system of simultaneous equations. *Econometrica*, 11, 1–12.

Holland, P.W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81, 946–960.

Imbens, G.W. and Wooldridge, J.M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47, 5–86.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66, 688–701.

Schield, M. (1995). Correlation, determination and causality in introductory statistics. *Proceedings of the Joint Statistical Meeting 1995*, Section on statistical education.

Sims, C. (1980). Macroeconomics and reality. *Econometrica*, 48, 1–48.

Stahlschmidt, S. (2015). Towards Causal Reasoning: Notes on Dimension Reduction, Graphical Models and Treatment Effects. Dissertation. Berlin: Humboldt-Universität zu Berlin

Stahlschmidt, S., and Hinze, S. (2016). How does the scientific progress in developing countries affect bibliometric impact measures of developed countries? A counterfactual case study on China. In Ràfols, I., Molas-Gallart, J., Castro-Martínez, E., & Woolley, R. (Eds.), *Peripheries, frontiers and beyond. Proceedings of the 21st International Conference on Science and Technology Indicators*. València: Editorial Universitat Politècnica de València.

Stahlschmidt, S. and Hinze, S. (2018). The dynamically changing publication universe as a reference point in national impact evaluation: A counterfactual case study on the Chinese publication growth. *arXiv*, 1801.05208 (submitted)

Wright, S (1921). Correlation and Causation. *Journal of Agricultural Research*, 20, 558-585.

Zhou, P. and Leydesdorff, L. (2016). A Comparative Study of the Citation Impact of Chinese Journals with Government Priority Support. *Frontiers in Research Metrics and Analytics*, 1, 3.