**Response to Reviewers’ Comments**

**Date**: February 14, 2023

**Manuscript Number**: SCIM-D-22-00945R2

**Title of Article**: Temporal trends in academic performance and career duration of principal investigators in ecology and evolutionary biology in Taiwan

**Corresponding Author**: Syuan-Jyun Sun

**Email**: sjs243@ntu.edu.tw

---------------------------------------------------------------------------------------------------------------------

Dear Dr. Lin Zhang,

Thank you for inviting us to submit a revised version of the manuscript. We greatly appreciate the valuable comments and feedback from the reviewers. We have incorporated many of the suggestions and the revision has substantially improved the manuscript. In particular, we have made the following major changes:

* Revised Table 1 by providing more details on the models results and fit measures (model specifications, sample sizes, regression coefficients, standard errors, 95% confidence limits).

Please see the following section for our detailed point-by-point responses. All line numbers pertaining to the changes refer to the revised manuscript.

Sincerely,

Syuan-Jyun Sun (corresponding author)

On behalf of Gen-Chang Hsu and Wei-Jiun Lin

---------------------------------------------------------------------------------------------------------------------

**Reviewer 1's Comments to the Author(s):**

**Comment 1** > The authors have responded all the suggestions and added some paragraphs which make this article more complete and readable.

**Response** > Thanks for the positive comments.

**Reviewer 2's Comments to the Author(s):**

The authors have made some changes and improvements to the manuscript. However, there are still many unresolved issues and the added information have also revealed new questions and issues. Below is my questions and comments. Hopefully they will help the authors to improve the manuscript.

**Comment 1** > I recommended that the authors include a more extensive, accurate and nuanced summary of the state the research the authors are aiming to make a contribution to. This is not fully answered in the revised manuscript. The authors have added some of the references I suggested in a sentence but that is not enough to cover the field.

**Response** > Maybe add a paragraph summarizing the current state of the field?

“Previous studies have focused on how various bibliometric indicators predict researchers’ future academic excellence and scientific contributions. The number of publications, top journal publications, publication rates, the number of distinct journals, and the number of citations are all important determinants of academic performance (Acuna et al. 2012; Danell 2011; Lindahl 2018). Academic performance is critical for researchers’ future success as publication requirements for recruitment as a new faculty member and promotion to full professor have surged in recent years, yet empirical quantification of how performance affects the duration before recruitment and promotion over time remains unexplored.”

**Comment 2** > The description of the data collection with the publish or perish software is better in the revised manuscript. However, some questions remain: (1) At row 155, the authors write: "We also cross-referenced the results item with the updated curriculum vitae online to ensure the accuracy of search results". How was the accuracy of the search results? And did the authors correct the bibliometric data when mis-matches etcetera was identified? (2) The authors definition of a PI the field of ecology and evolutionary biology is not clearly stated. The selection seems to be based on affiliation. However, the field the authors are investigating is not clearly defined. The readers of Scientometrics may not have deep knowledge of this field. Therefore, I recommend the authors to provide a concise section where the field of ecology and evolutionary biology is defined and described in terms of relevant characteristics, e.g., formal scientific communication, publication praxis, collaboration, interdisciplinarity, etcetera. As of now several questions arise when I read the data collection section. Are the different departments the authors refer to sub-fields in the field of inquiry? Is publication praxis and citation rates similar in these sub-fields? Did all PIs publish in the field of ecology and evolutionary biology during their whole career or did some publish in other fields earlier in their careers and then transitioned to this field? Did all PIs do their PhDs in this field?

**Response** >

1. Accuracy of the search results
2. Definition of EEB
3. Department and sub-fields
4. PI publication patterns in EEB
5. PhD of these PIs

**Comment 3** > The issue of not using a normalized bibliometric indicator is unfortunately only partly resolved since the authors choose to continue with the h-index. First, the authors claim that normalization by field is not necessary since all PIs was faculty at one of 8 HEIs at time of data collection, but they have not provided any proper arguments or evidence in their manuscript to back this claim. Normalization might be relevant also at the sub-field level (Van Eck et al., 2013). Second, the authors write "By including both journal articles and book (chapters) in the calculation of h-index rather than calculating h-indexes separately for each of them, the potential variation in h-index among the two publication types were reduced". I do not understand this reasoning because it is not in line with the logic behind normalizing for document type in the scientometric literature. Are the publication rates similar for authors that publish more chapters and authors that publish more articles? Are citation rates similar for chapters and articles in the field they are investigating and in the sample? I recommend that the authors clarify their reasoning here and provide good references and arguments for how this works.

**Response** > Thanks for pointing out the issue regarding the variations in h-index at the sub-field level. Yes, as Van Eck et al. (2013) found, there were substantial differences in citation practices among areas of basic and diagnostic research and clinical intervention research in medical sciences. On the other hand, the majority of publications in the field of ecology and evolutionary biology are basic research. Moreover, Van Eck et al. (2013) compared the citation patterns among three WoS subjects *Cardiac & cardiovascular systems*, *Clinical neurology*, and *Surgery* within medical field (in other words, these three subjects were the sub-fields in their study), whereas our study focuses mainly on a single WoS subject *Ecology* (which is equivalent to a sub-field in Van Eck et al. [2013]). Therefore, the within-field differences should be relatively lower compared to those revealed by Van Eck et al. (2013). Nonetheless, we do acknowledge that within-field variations may still exist in our study and we have added this potential caveat to the discussion section (Line XXX):

“*Finally, our analyses were based on PIs in ecology and evolutionary biology, within which variations in publication performance and citation patterns may exist.*”

Regarding the issue with citation differences among publication types, we examined the citation rates of articles and book chapters for the PIs in our study and found that the differences were minor. Moreover, books generally constituted only a minor part of the PIs’ total publications. Therefore, including both articles and books in our calculation of h-index should not largely bias our results.

\* Revise the methods to show that the citation rates are similar and show this in the supplementary methods.

**Comment 4** > The authors have included a section about google scholar but they have not provided convincing arguments. The argument of free access is irrelevant in this context. The argument that the lack of good metadata in google scholar does not matter for the authors study since they are not using it, is not really true. If the authors had used a citation database with good metadata etcetera, e.g., Scopus, it would have been easier to use a normalized bibliometric indicator as the dependent variable which would have strengthened the study.

**Response** >

The critical bibliometric indexes for our analysis are the number of publications, annual citations, and year of publication. These are readily available on Google Scholar Profile. On the other hand, the metadata for researchers’ publications concern mainly DOI, affiliation data, funding information, etc., which are not directly relevant to our study regardless of which search engine is used. In fact, as mentioned in our previous response, studies have shown that Google Scholar has a wider coverage compared to WoS or Scopus, and if metadata are not of substantial importance, Google Scholar might be preferable.

Martín-Martín et al. (2021): *“… the final decision about which source to use may depend on properties of the sources other than coverage, such as metadata quality and bulk access options. If these factors are not of overriding importance, however, then Google Scholar is the best choice in almost all subject areas for those needing the most comprehensive citation counts but not needing complete lists of citing sources.”*

**References:**

Martín-Martín, A., Thelwall, M., Orduna-Malea, E. & Delgado López-Cózar, E. (2021). Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations’ COCI: a multidisciplinary comparison of coverage via citations. *Scientometrics*, 126, 871-906.

**Comment 5** > The authors have now included information about missingness in the manuscript. At row 486, Tables and figures, the size of the sample used for each model is presented. In the model with the lowest number of missing values the missingness is 7.6 % (model 1). In model 2, 4, and 6 is very high. In model 2 the sample size is 58 out of 145, i.e, 60% missing values, in model 4 there are 62% missing values, and in model 6 there are 68% missing values. Listwise deletion might be a reasonable method to handle missingness below 5% (Graham, 2009) depending on the mechanism for the missingness. When the missingness in the range presented by the authors the missingness needs to be addressed with some more advanced method than list wise deletion to handle the missingness, e.g., multiple imputation with chained equations. However, given the small sample in combination with a missingness of this magnitude, I doubt that the authors can provide a convincing solution other than collecting more data. How can we be sure that the authors results are not biased by the large amount of missingness?

**Response** > Thanks for the suggestion and the reference for handling missing data. In fact, the missing entries in our data set were not true missingness due to incomplete records. Rather, these missing entries simply did not exist because not all PIs in our data set were full professors. For assistant or associate professors, there would be no data during the promotion phase. This is why Model 2, 4, and 6 had much lower sample sizes compared to Model 1, 3, and 5.

Because of the nature of the missing entries, it may not be suitable to conduct multiple imputation in our case: the promotion data for PIs who are not full professors cannot be simply modeled/predicted by the full professors. Moreover, these missing entries should not introduce bias into our analysis because these entries were not truly missing but non-existent. That said, we agree that the sample sizes for Model 2, 4, and 6 might not be satisfactory. Nonetheless, given that ecological and evolutionary biology (EEB) in Taiwan is relatively small, and the selected PIs in our study actually constituted XX% of the PIs in EEB in Taiwan, we feel that our analysis could still provide useful information for this field.

**Comment 6** > I do not think the authors conclusions are supported by the data. Given the small sample size and the descriptive statistics I believe the authors need to show that their sample is representative by (1) defining a population; (2) Compare the variables in the sample with the equivalent variables in the population (e.g., how the proportion of men and women in the sample compare with the proportion of men and women in the population; How does the percentages of full, associate, and assistant professors compare with the population, etcetera). The authors should also show that the 8 Higher Education Institutions (HEI) are representative and how the PIs are distributed over the HEIs and compare the share of PIs in the HEIs in the sample with the share in the population. How many HEIs are there in the sampling frame? Why have the authors only sampled from top ranked HEIs? How does that comply with the representativity of the sample?

**Response** >

1. Summary of PIs in EEB in all universities in Taiwan.
2. Compare the PIs in our analysis to those in the defined population.
3. Provide this info in the supplementary methods.

The table below shows the number of PIs in each of the eight university/institute in our analysis (Note that there were only a few PIs from National Tsing Hua University and National Yang Ming Chiao Tung University because ecology and evolutionary biology is not a main research focus at these two universities).

|  |  |
| --- | --- |
| University/institute | Number of PIs |
| National Taiwan University | 47 |
| National Chung Hsing University | 28 |
| Academia Sinica | 26 |
| National Sun Yat-sen University | 16 |
| National Taiwan Normal University | 13 |
| National Cheng Kung University | 9 |
| National Tsing Hua University | 5 |
| National Yang Ming Chiao Tung University | 1 |

We only included PIs from higher-ranked universities/institutes for the following reasons:

1. The PIs in EEB from the eight top-ranked universities/institutes are fairly representative of those in Taiwan.
2. The research environment and funding resources differ between higher-ranked and lower-ranked universities/institutes (generally having fewer funding opportunities), and such inherent difference among universities/institutes could affect the research performance of PIs. As a result, including PIs from lower-ranked universities/institutes might introduce bias into our results.
3. Most of the applicants will set higher-ranked universities/institutes as their top priorities. Thus, from the applicants’ perspective, it would be more relevant to focus on the patterns in these universities/institutes.

**Comment 7** > At row 208 I cannot see any regression diagnostics for the poisson regressions. Why did the authors choose poisson over negative binomial regression? Have the authors tested the poisson models for over dispersion?

**Response** > We did check the assumptions of equal variance and normality for all six models in our analysis in our previous revision. However, there seemed to be a mistake in the online submission system and the figures for model diagnostics in our response document were lost. We provided them again below. Overall, there is no severe violation of the assumptions. The residuals in Model 5 seem to deviate from the line at the top-right corner. Nonetheless, this should not be a major issue as studies have shown that regression models are fairly robust to moderate degree of non-normality (Knief and Forstmeier 2021; Schielzeth et al. 2020). (The diagnostic plots below were generated using the R package “performance”.)

|  |
| --- |
| Model 1  Performance_recruitment |
| Model 2  Performance_promotion |
| Model 3  Duration_recruitment |
| Model 4  Duration_promotion |
| Model 5  Diff_recruitment |
| Model 6  Diff_promotion |

We also checked for overdispersion in our Poisson models (Model 1. 2, 3, and 4) using the function “check\_overdispersion()” in the R package “performance”. The test results show that the Model 1 is overdispersed, and therefore we refit the model using a negative binomial distribution and a log link function. The relevant contents in the manuscript was updated accordingly:

1. Methods:

*“Model 1 was fitted with a negative binomial error distribution and a log link function as the response is non-negative integers with significant overdispersion (χ2 = 199.59, df = 130, P < 0.001); Model 2 was fitted with a Poisson error distribution and a log link function as no significant overdispersion was detected (χ2 = 64.01, df = 51, P = 0.10).”* (Line XXX)

*“Both Model 3 and 4 were fitted with a Poisson error distribution and a log link function as the response is non-negative integers without significant overdispersion (Model 3: χ2 = 149.58, df = 125, P = 0.07; Model 4: χ2 = 32.98, df = 47, P = 0.94).”* (Line XXX)

1. Figure 3a.
2. The statistical results for Model 1 in Table 1.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Dispersion ratio | *χ2* | *df* | *P* |
| Model 1 | 1.57 | 199.59 | 130 | < 0.001 |
| Model 2 | 1.26 | 64.01 | 51 | 0.10 |
| Model 3 | 1.20 | 149.58 | 125 | 0.07 |
| Model 4 | 0.70 | 32.98 | 47 | 0.94 |

**References**

Knief, U. & Forstmeier, W. (2021). Violating the normality assumption may be the lesser of two evils. *Behavior Research Methods*, 53, 2576-2590.

Schielzeth, H., Dingemanse, N.J., Nakagawa, S., Westneat, D.F., Allegue, H., Teplitsky, C., Réale, D., Dochtermann, N.A., Garamszegi, L.Z. & Araya‐Ajoy, Y.G. (2020). Robustness of linear mixed‐effects models to violations of distributional assumptions. *Methods in ecology and evolution*, 11, 1141-1152.

**References suggested by Reviewer 2**

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. Annual review of psychology, 60, 549-576.

Van Eck, N. J., Waltman, L., van Raan, A. F., Klautz, R. J., & Peul, W. C. (2013). Citation analysis may severely underestimate the impact of clinical research as compared to basic research. PloS one, 8(4), e62395.