**Response to Reviewers’ Comments**

**Date**: December 30, 2022

**Manuscript Number**: SCIM-D-22-00945R1

**Title of Article**: Increased academic performance and prolonged career duration among 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 most of the suggestions and the revision has substantially improved the manuscript. In particular, we have made the following major changes:

* Corrected the citation format issue and added several recent articles to the manuscript to better reflect the current status of IGP research.

Please also 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

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

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

This article collects various types of open data and investigates the academic job market in Taiwan, which is an important topic deserving further future studies. However, there are some questions or suggestions as follows:

*(1) About data collection*

**Comment 1** > In line 123 to 126, this research includes 145 PIs who had an updated CV. In other words, researchers who hadn't updated CV were excluded and those PIs' academic careers whether represent specific patterns or not. Hence, what's the representativeness of these 145 PIs in this study?

**Response** >

* Seven major research universities and institutes, should be representative
* For those without an up-to-date CV, we will search online and find any information available to complete the necessary data
* How many faculty members are there in these seven institutes? The percentage of them that were excluded?
* Ask Wei-Jun for more details

**Comment 2** > Besides the institutional/departmental websites, and ORCID, other open data or database could be further consideration, such as Web of Science, Scopus, or Academic Research Service Portal Researcher Query of National Science and Technology Council.

**Response** >

* Yes, these are good sources to obtain PIs’ academic publication records. However, their education backgrounds and academic track records are usually not available on these websites. Therefore, the main sources of our data are still institutional/departmental websites, and ORCID, other open data or database.
* Ask Wei-Jun for more details.

1. *About literatures*

**Comment 3** > This study includes those variables such as year of recruitment, gender, PhD university origin, PhD university ranking, year of promotion and so on (shown as Table 1). What's the theoretical basis of relationship between those variables? If this article supplied the section of literature review, readers would more understand the existing related researches of this topic even the theoretical basis.

**Response** >

* Thanks for bringing this up. Yes, it would be beneficial to provide some background information for the readers so that they can better understand the context and objectives of our study. We have now added this in our revised manuscript.
* Year of recruitment and promotion
* Gender
* PhD university origin
* PhD university ranking

**Comment 4** > According to the results and discussions, what's the concrete suggestions to higher education policy, recruitment of university's teaching and research staffs, or PhD students who aim to academic careers?

**Response** >

* Thanks for this critical comment. One of the main goals of our study is to provide references for the academic communities. We have now included this in our conclusion:
* For students pursuing a PhD degree:
* For PhD students/postdocs/early-stage researchers:
* For faculty members:
* For employment departments and institutes:

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

**Comment 1** > The authors examine "how academic performance as well as duration before recruitment as a new principal inverstigator (PI) and promotion to full professor changed over time, and how PhD university origin, PhD university ranking, and gender affected the career success". The manuscript has potential to make a contribution to the literature. However, the manuscript has some problems which makes me recommend major revision. Hopefully my questions and comments can help the authors to improve the manuscript.

**Response** >

* Thanks for the positive attitude towards our study. We have revised our manuscript based on the following comments.

*Title*

**Comment 2** > The title does not really convey what the study is about.

**Response** >

* We have now changes the title to.

*Abstract*

**Comment 3** > The first sentence in the abstract "Academic job markets have become increasingly challenging worldwide, yet it remains poorly characterized how competitively-successful candidates should be and what the underlying determinants of their success are" seem unsubstantiated. There are numerous studies that have examined determinants of academic success (see. e.g., Hirsch, 2007; Danell, 2011; Acuna et al., 2012; Havemann and Larsen, 2015; Bornmann and Williams, 2017a; Lindahl, 2018). My recommendation is that the authors include a more extensive literature review on previous research in the field and provide a more accurate and nuanced summary of the state of this research.

**Response** >

* Read the provided references and add relevant information to the introduction.
* Previous studies have looked at how academic track records (publication rate, top journal publications, and top 10% publications) may predict individual research performance and excellence.
* But so far there is little research on how education background, gender affect researchers’ performance.

*Materials and Methods*

Measurement of academic performance

**Comment 4** > The data collection with the Publish or Perish software for the h-index need to be described and presented much more and in greater detail. There is not enough detail to be able to review the data collection or the data for calculating the h-index. As a reader I'm not sure how the authors collected the publications for the authors. Did they conduct searches through Publish or Perish at the publication level or the author level? How was the search queries formulated, i.e., did the authors conduct searches on the basis of publication titles, persistent identifiers, etc? My recommendation is (1) that the authors provide the search queries in the manuscript or as an appendix and (2) that they provide much more detail about the data collection procedure and what they have done including how they handle the CV data etcetera.

**Response** >

* Ask Wei-Jun about the data collection and searching procedure

**Comment 5** > I cannot see how many documents that are included in the final dataset? This should be included in the manuscript.

**Response** >

* Ask Wei-Jun about the data collection and searching procedure.
* Manually extract the relevant information and combine them with the data downloaded from the Publish and Perish database as a single datasheet.

**Comment 6** > What do the authors mean with "regardless of authorship for" in the sentence on page 7 row 1-3?

**Response** >

* Publications of any authorship position, not limited to first or corresponsing author.

**Comment 7** > The authors use the h-index to measure research performance. The h-index is not a normalized indicator of research performance, i.e., it do not adjust for, e.g., research area, publication year, and publication type, and do not live up to best practice in scientometric research. See e.g., Waltman (2016) for a review of citation indicators. To use non-normalized bibliometric indicators as measures of research performance can lead to severe biases in the analyses. The h-index has been heavily criticized in the scientometric literature (Bornmann, & Daniel, 2007; Bornmann, & Daniel, 2009) and it is not recommended to use to measure research performance at the individual level (Waltman, & Van Eck,2012). My recommendation is that the authors change their dependent variable to a normalized bibliometric indicator that is in accordance with best practice in scientometric research or provide good arguments for why the use of h-index should be used in this case. Another potential solution is to use a variation of the h-index that adjust for the problems with the h-index and fit the context of the authors study (see e.g., Alonso et al., 2009, for a review of h-index and its variant).

**Response** >

* Justifications of h-index: within the same field and so relatively fine, normalization is not a main concern in our study, also relatively easy and straightforward, and readily available.
* See **h-Index: A review focused in its variants, computation and standardization for different scientific fields** for a comprehensive review of the advantanges and disadvantages of h-index.

**What do we know about the h index?**

“The h index is seen to have the advantage that it gives a robust estimate of the broad impact of a scientist's cumulative research contributions (Hirsch, [2005](https://onlinelibrary.wiley.com/doi/10.1002/asi.20609#bib26)a). This means that the h index is insensitive to a set of lowly cited (noncited) papers or to one or several highly cited papers: A scientist with very few highly cited papers (a “one-hit wonder”) or, alternatively, many lowly cited papers will have a weak h index (Cronin & Meho, [2006](https://onlinelibrary.wiley.com/doi/10.1002/asi.20609#bib15); Egghe, [2006](https://onlinelibrary.wiley.com/doi/10.1002/asi.20609#bib17)b, [2006](https://onlinelibrary.wiley.com/doi/10.1002/asi.20609#bib18)c). As a rule, the index favors enduring performers that publish a continuous stream of papers with lasting and above-average impact (Anonymous, [2005](https://onlinelibrary.wiley.com/doi/10.1002/asi.20609#bib1)).”

“A further advantage seen for the h index is that the necessary data for calculation is easy to access”

“Since h values (that is, published papers and the citations papers receive) increase over time (Egghe, [2006](https://onlinelibrary.wiley.com/doi/10.1002/asi.20609#bib16)a; Hirsch, [2005](https://onlinelibrary.wiley.com/doi/10.1002/asi.20609#bib26)a), it is apparent that a scientist's h index depends on the person's scientific age (that is, years publishing, Glänzel, [2006](https://onlinelibrary.wiley.com/doi/10.1002/asi.20609#bib24)b; Roediger, [2006](https://onlinelibrary.wiley.com/doi/10.1002/asi.20609#bib34)). Therefore, in ranking scientists, the h index always puts newcomers at a disadvantage and older, well-established scientists at a advantage (Cronin & Meho, [2006](https://onlinelibrary.wiley.com/doi/10.1002/asi.20609#bib15); Glänzel, [2006](https://onlinelibrary.wiley.com/doi/10.1002/asi.20609#bib24)b). It should also be considered that when using the h index for comparison purposes, there are discipline-dependent citation patterns in science”

**The state of h index research - Is the h index the ideal way to measure research performance?**

“If the h index is used for the evaluation of research performance, it should always be taken into account that, similar to other bibliometric measures, it is dependent on the length of an academic career and the field of study in which the papers are published and cited. For this reason, the index should only be used to compare researchers of a similar age and within the same field of study.”

**A review of the literature on citation impact indicators**

“For practical purposes, there often is a need to make comparisons between publications that are from different fields or different years or that have different document types. Normalized citation impact indicators have been developed to make such comparisons.”

“This is because there are large differences among fields in citation density, that is, in the average number of citations per publication. In addition to comparisons between publications from different fields, one should also be careful with comparisons between publications from different years. Even within the same field, a publication from 2005 with 25 citations cannot necessarily be considered to have a higher citation impact than a publication from 2010 with ten citations.”

**Comment 8** > The authors need to discuss the pros and cons of using google scholar. Why use Google Scholar instead of a citation database, e.g., Scopus or the citation indices accessible through Web or Science? Harzing state that Web of Science and Scopus have higher accuracy so why not use them (see <https://harzing.com/resources/publish-or-perish/manual/using/query-results/accuracy)?> How might the use of Google Scholar affect the results? The authors should provide a discussion in the manuscript where the pros and cons of using Google Scholar become transparent for the reader.

**Response** >

* Discuss the pros and cons.
* Freely available compared to WoS (paid services).
* See “Comparing Google Scholar with Web of Science and Scopus” in Waltman 2016.

**Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations’ COCI: a multidisciplinary comparison of coverage via citations**

“Unlike WoS and Scopus, Google Scholar follows an inclusive and automated approach, indexing any seemingly academic document that its crawlers can find and access on the web, including those behind paywalls through agreements with their publishers (Van Noorden [2014](https://link.springer.com/article/10.1007/s11192-020-03690-4#ref-CR60)). Additionally, Google Scholar is free to access, allowing users to access a comprehensive and multidisciplinary citation index without charge.”

“Additionally, there is a significant amount of extra coverage in Google Scholar that is not found in any of the other data sources (26% of all citations across all data sources). Google Scholar could therefore make an important contribution to the scientific community by opening its bibliographic and citation data, which would also facilitate the identification of errors such as coverage fluctuations.”

“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.”  
  
**Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories**

“This study provides evidence that GS finds significantly more citations than the WoS Core Collection and Scopus across all subject areas. ”

“In conclusion, this study gives the first systematic evidence to confirm prior speculation ([Harzing, 2013](https://www.sciencedirect.com/science/article/pii/S1751157718303249" \l "bib0085); [Martín-Martín et al., 2018](https://www.sciencedirect.com/science/article/pii/S1751157718303249" \l "bib0145); [Mingers & Lipitakis, 2010](https://www.sciencedirect.com/science/article/pii/S1751157718303249" \l "bib0160); [Prins et al., 2016](https://www.sciencedirect.com/science/article/pii/S1751157718303249" \l "bib0200)) that citation data in GS has reached a high level of comprehensiveness, because the gaps of coverage in GS found by the earliest studies that analysed GS data have now been filled. It surpasses WoS and Scopus numerically in all areas of research, and is greatly superior in the areas where WoS and Scopus have a poor coverage, including the Social Sciences and Humanities. However, at this point there is no reliable and [scalable method](https://www.sciencedirect.com/topics/computer-science/scalable-method) to extract data from GS, and the metadata offered by the platform is still very limited,”

**A review of the literature on citation impact indicators**

“Google Scholar was also launched in 2004. It indexes scholarly literature that is available online on the web. This includes not only publications in journals and conference proceedings, but also for instance books, theses, preprints, and technical reports. Google Scholar is made freely available by Google. It should be emphasized that Google Scholar is of a very different nature than WoS and Scopus. It is primarily a search engine for scholarly literature, and it provides only very limited bibliographic meta data on publications.”

“A general impression obtained from the literature is that Google Scholar suffers from a lack of quality control. Many inaccuracies in Google Scholar are reported in the literature. [Jacsó, 2005](https://www.sciencedirect.com/science/article/pii/S1751157715300900" \l "bib0770), [Jacsó, 2006](https://www.sciencedirect.com/science/article/pii/S1751157715300900" \l "bib0775), [Jacsó, 2010](https://www.sciencedirect.com/science/article/pii/S1751157715300900" \l "bib0780) for instance discusses problems related to content gaps, incorrect citation counts, and phantom data. The possibility of manipulating citation counts in Google Scholar is discussed by [Beel and Gipp (2010)](https://www.sciencedirect.com/science/article/pii/S1751157715300900" \l "bib0155), [Labbé (2010)](https://www.sciencedirect.com/science/article/pii/S1751157715300900" \l "bib0850), and [López-Cózar, Robinson-García, and Torres-Salinas (2014)](https://www.sciencedirect.com/science/article/pii/S1751157715300900" \l "bib0305). Google Scholar is also criticized for its lack of transparency (e.g., [Jacsó, 2005](https://www.sciencedirect.com/science/article/pii/S1751157715300900" \l "bib0770), [Wouters and Costas, 2012](https://www.sciencedirect.com/science/article/pii/S1751157715300900" \l "bib1690)). It is unclear what is covered by Google Scholar and what is not.”

Statistical analyses  
  
**Comment 9** > CV data usually comes with a lot of missing values. However, I cannot find anything about missing values in the text. Is there no missing values in the data? If there are missing values a wonder how have the authors handled the missingness.

**Response** >

* Yes there are missing values, and not analyzed for respective models
* If the relevant information in not presented, we will not included in our dataset (also see our response to Reviewer 1’s comment)

**Comment 10** > Regarding the LMMs the authors are referencing R-packages which is fine. But I recommend the authors to also provide relevant references for the actual methods they use.

**Response** > Cite LMMs.

**Comment 11** > What do the authors mean with the following sentence: "Non-significant interactions (p > 0.05) were dropped from our final model results. Did the authors first try all possible interactions for each model and then in the final models they only included the significant interactions. Or did the authors include all interactions in the final model but only show the significant interaction in the results (i.e., Table 1)?

**Response** >

* The first one. We have revised this part to clarify this.

**Comment 12** > The authors write that they log-transformed the dependent variables "to meet the assumption of normality". (page 9, row190-191). Did the authors test the assumption of normality on the transformed variables?

**Response** >

* Test the assumption using SW tests on residual and qqplots.

*Results*  
  
**Comment 13** > The authors should provide descriptive statistics for their data and variables. Either in the results section or in the Materials and Methods section. This is important so that the reader can get an overview of the data and its properties.

**Response** >

* Add data overview in the first part of the results.

**Comment 14** > The authors should be clearer about how they use p-values and how they interpret them, especially since their data is not a random sample. For example, the authors write that "PhD university origin, ranking, and gender had no effect on the duration either before recruitment or before promotion (page 10, row 49-54). In Table 1 I can observe that, e.g., the coefficient for the "PhD university origin" is 3.48 which indicates a positive effect and that the p-value is 0.06. I take it that the chosen significance level in the manuscript is 0.05. 0.06 is not that much higher than 0.05. From my perspective I would say that there is a positive effect but that the p-value indicates that there are some uncertainty and that interpretation should therefore be done with some caution. There are two issues here. First, how does the authors define and use p-values in the manuscript? This is not clear. Second, the sample is not a random sample so it seems a bit strange to be super strict about the p-values and e.g., conclude that there is an effect if the p-value is 0.04999 and conclude that there is no effect of the p-value is 0.05, regardless of the size of the coefficient, taking sample size into consideration, etc. The use of p-values does in either case not live up to the required assumptions for making real inference due to the non-random sample. Overall (i.e., this is a recommendation for all the results and not just for the example I provided regarding the "PhD university origin"). The sample size is small (N = 145) and a larger sample size would likely produce significant results.  My suggestion for the authors is to adopt a less dichotomous and more nuanced strategy for interpreting the results of their analyses, e.g., using confidence intervals for determining uncertainty together with the p-values, and not dogmatically approve or disprove of an effect just by looking at the p-values.

**Response** >

* This is a great suggestion.
* Add CI to the table
* Re-interpret the results

**Comment 15** > Can the authors complement the analyses with effect sizes so that it become easier for the reader to understand the size of the effects?

**Response** >

* Add effect sizes to the table

**Comment 16** > Can the authors transform back the coefficients so that it becomes easier to interpret the actual effects.

**Response** >

* Not all vars are log-transformed if the residual are normal.
* Show the original coefficients to assess the relative effects of each predictor in each models, not to make predictions about new data, so we think presenting the original coefficients is fine.

**Comment 17** > I believe that there is to little information in Table 1 regarding the models and outcomes. As a reader it is difficult to properly assess the results of the analysis. Standard errors and confidence intervals should be included. Some kind of model of fit measure should be included.

**Response** >

* Revise the table to include slope, SE, n, CI.

**Comment 18** > As I understand it Table 1 present 6 models. This should be more clearly presented in the table, i.e., that each dependent variable in the first column denotes a specific model.

**Response** >

* Revise the table.

**References**

Acuna, D. E., Allesina, S., & Kording, K. P. (2012). Predicting scientific success: Daniel E. Acuna, Stefano Allesina and Konrad P. Kording present a formula to estimate the future h-index of life scientists. Nature, 489(7415), 201-202. <https://doi.org/10.1038/489201a>

Alonso, S., Cabrerizo, F. J., Herrera-Viedma, E., & Herrera, F. (2009). h-Index: A review focused in its variants, computation and standardization for different scientific fields. Journal of informetrics, 3(4), 273-289.

Bornmann, L., & Daniel, H. D. (2007). What do we know about the h index?. Journal of the American Society for Information Science and technology, 58(9), 1381-1385.

Bornmann, L., & Daniel, H. D. (2009). The state of h index research: Is the h index the ideal way to measure research performance?. EMBO reports, 10(1), 2-6.

Bornmann, L., & Williams, R. (2017). Can the journal impact factor be used as a criterion for the selection of junior researchers? A large-scale empirical study based on ResearcherID data. Journal of Informetrics, 11(3), 788-799. <https://doi.org/10.1016/j.joi.2017.06.001>

Danell, R. (2011). Can the Quality of Scientific Work Be Predicted Using Information on the Author's Track Record? Journal of the American Society for Information Science and Technology, 62(1), 50-60. <https://doi.org/10.1002/asi.21454>

Havemann, F., & Larsen, B. (2015). Bibliometric indicators of young authors in astrophysics: Can later stars be predicted? Scientometrics, 102(2), 1413-1434. <https://doi.org/10.1007/s11192-014-1476-3>

Hirsch, J. E. (2007). Does the h index have predictive power? Proceedings of the National Academy of Sciences, 104(49), 19193-19198. <https://doi.org/10.1073/pnas.0707962104>

Lindahl, J. (2018). Predicting research excellence at the individual level: The importance of publication rate, top journal publications, and top 10% publications in the case of early career mathematicians. Journal of Informetrics, 12(2), 518-533.

Waltman, L., & Van Eck, N. J. (2012). The inconsistency of the h‐index. Journal of the American Society for Information Science and Technology, 63(2), 406-415.