Are internationally co-authored journal articles better quality?

Thelwall, Mike; Kousha, Kayvan; Abdoli, Mahshid; Stuart, Emma; Makita, Meiko; Wilson, Paul

DOI:
10.48550/arXiv.2212.05417 Focus to learn more

Publication date:
2022

Are internationally co-authored journal articles better quality? The UK case 2014-2020

Mike Thelwall
Statistical Cybermetrics and Research Evaluation Group, University of Wolverhampton, UK. https://orcid.org/0000-0001-6065-205X m.thelwall@wlv.ac.uk

Kayvan Kousha
Statistical Cybermetrics and Research Evaluation Group, University of Wolverhampton, UK. https://orcid.org/0000-0003-4827-971X k.kousha@wlv.ac.uk

Meiko Makita
Statistical Cybermetrics and Research Evaluation Group, University of Wolverhampton, UK. https://orcid.org/0000-0002-2284-0161 meikomakita@wlv.ac.uk

Mahshid Abdoli
Statistical Cybermetrics and Research Evaluation Group, University of Wolverhampton, UK. https://orcid.org/0000-0001-9251-5391 m.abdoli@wlv.ac.uk

Emma Stuart
Statistical Cybermetrics and Research Evaluation Group, University of Wolverhampton, UK. https://orcid.org/0000-0003-4807-7659 emma.stuart@wlv.ac.uk

Paul Wilson
Statistical Cybermetrics and Research Evaluation Group, University of Wolverhampton, UK. https://orcid.org/0000-0002-1265-543X pauljwilson@wlv.ac.uk

Jonathan Levitt
Statistical Cybermetrics and Research Evaluation Group, University of Wolverhampton, UK. https://orcid.org/0000-0002-4386-3813 j.m.levitt@wlv.ac.uk

International collaboration is sometimes encouraged in the belief that it generates higher quality research or is more capable of addressing societal problems. In support of this, there is evidence that the journal articles of international teams tend to be more cited than average. Reasons other than the benefits of international collaboration could explain this, however, such as increased national audiences from researcher networks. This article investigates research quality using 148,977 UK-based journal articles with post publication peer review scores from the 2021 Research Excellence Framework (REF). Based on an ordinal regression model controlling for collaboration, international partners increased the odds of higher quality scores in 27 out of 34 Units of Assessment (UoAs) and all four Main Panels. At the country level, the results suggests that UK collaboration with other advanced economies generates higher quality research, even if the countries produce lower citation impact journal articles than the UK. Conversely, collaborations with weaker economies tend to produce lower quality research, as judged by REF assessors. Overall, the results give the first large scale evidence of when international co-authorship for journal articles is beneficial, at least from a UK perspective, and support the continuation of research policies that promote it.

Keywords: Research collaboration; co-authorship; scientometrics; research policy; REF2021; Research assessment; research quality

Introduction

There are many types of research collaboration, not all of which lead to co-authored journal articles (Katz & Martin, 1997), but this study focuses articles having authors affiliated with at least two countries. This phenomenon has grown consistently over the last century (Larivière et al., 2015; Hsiehchen et al., 2018) and is widely encouraged by research funders in the belief that it is beneficial for scientific progress (Olechnicka et al., 2019). This idea is supported by evidence of the greater scientific impact of internationally co-authored research (Zhou et al., 2020). Nevertheless, international collaborations do not produce more novel research and it has been hypothesised that its citation advantage is an audience effect rather than a quality effect (Wagner et al., 2019). The audience effect hypothesis has not been fully tested yet, however, and international research may generate more citations because it is more rigorous or significant, which are two other important components of quality (Langfeldt et al., 2020). It may also be more successful at deepening knowledge about known problems (a less risky research strategy: Foster et al., 2015) or by pooling data (Thelwall et al., 2020). Moreover, a relatively unoriginal medical article might generate substantial impact by combining patient data from multiple countries to identify statistically significant conclusions that were not possible with previous single-country studies (e.g., Bolton et al., 2021; Tritschler et al., 2020), or may generate larger sample sizes for the same cost by recruiting in lower income countries (Thiers et al., 2008), for example.

It is not yet known whether internationally collaborative research tends to be better quality. Most studies of all types of research collaboration have used citations or journal impact factors as quality proxies or have directly examined citation rates without hypothesising about a relationship with research quality. One minor exception showed that collaborative research was more likely to be accepted for publication in a social psychology journal (Presser, 1980), although it did not investigate international collaboration. A larger scale study of 16,554 biomedical articles published before 2013 with at least two F1000Prime post-publication peer review quality scores found a weak Spearman correlation (0.07) between the number of countries and the average quality score of articles, but this may have been partly a collaboration effect. A regression analysis with the same data confirmed the positive but very weak effect of the number of authors (Bornmann, 2017). A stepwise logistic regression analysis of REF2014 journal articles to identify factors associating with the highest quality scores found internationality to be a significant predictor in only four out of 36 fields examined (Clinical Medicine; Earth Systems and Environmental Sciences; Computer Science and Informatics), but the regression included many independent variables that could have weakened the role of international collaboration. It also did not take a fractional approach (HEFCE, 2015). From a much wider perspective, an analysis of published research suggested that disruption of international research sharing (in the form of citing foreign publications), weakened the production of novel research during World War 1 (Iaria et al., 2018), underlining the importance of a different type of international cooperation.

In response to the lack of evidence of a relationship between research quality and international collaboration, with the partial exceptions of the weak effects in the two contexts noted above, this article investigates the issue for all fields of science and with the largest and most recent dataset yet. It focuses on the UK for the pragmatic reason of data availability. Whilst the quality of collaboration seems almost certain to vary between countries, this is

also addressed because it has not been investigated before empirically on a large scale. The following research questions drive the study.

1. In which fields are UK-authored journal articles higher quality when they are internationally co-authored?
2. Does the quality of an internationally collaborative article depend on which countries collaborate?
3. Are high impact international partners for the UK necessary for international research to be higher quality?

Background

International collaboration has been growing for at least a century although with substantial differences between countries (Hoekman et al., 2010; Larivière et al., 2015; Leydesdorff & Wagner, 2008; Ribeiro et al., 2018). The drivers and enablers of international collaboration seem to be complex, varying between fields (Wagner, 2005), but probably include funding, faster communication, cheaper travel, and access to expertise and resources (Adams et al., 2005; Freeman et al., 2014). International mobility and migration also helps (Kato & Ando, 2017).

An analysis of Web of Science journal articles and conference papers 1973-2009 found a tendency for countries to collaborate with other countries having dissimilar levels of citations per paper (Hsiehchen et al., 2018). For researchers in the Global South, engaging with academics and priorities from the Global North is often important to get international recognition for their work (Martinez & Sá, 2020). This might provide an incentive to collaborate with Global North researchers, partly explaining the dissimilar country collaboration tendency.

There are fundamental obstacles to international collaboration in some fields. In professional contexts, including some aspects of healthcare, nationally organised services are obstacles to international collaboration, because working practices are different (e.g., Gladstone et al., 2019). Law is perhaps an extreme example, with each country having different legal frameworks. This may be partly why international research collaboration is greater in the natural sciences than the social sciences, and rare in the humanities (Lariviere et al., 2006; Gazni et al., 2012).

National and field differences in the citation advantages of international teams

There is substantial evidence that international collaboration associates with higher citation rates (Hall et al., 2018; Larivière et al., 2015; Zhou et al., 2020) and almost all countries tend to attract more citations per paper for their international collaborations (Hsiehchen et al., 2018). It does not seem to increase productivity, however, at least based on fractional counting of Italian journal articles (Abramo et al., 2017). Research collaboration of any type usually associates with more citations (Thelwall & Maflahi, 2020), so it is important to distinguish between the general benefits of collaboration and those from international partners.

The citation advantage of international collaboration is unequal between countries. The country with the lowest citation impact in a collaboration tends to gain the most from it (Lancho-Barrantes et al., 2013; Guerrero Bote et al., 2013), except for collaboration with the USA (which has a unique bibliometric advantage in citation databases: Gingras & Khelfaoui, 2018). Collaboration with Iran may also not be beneficial overall, although it is not clear why (Guerrero Bote et al., 2013). A regression-based analysis of the number of citations or

Mendeley readers for biochemistry research articles found that having a US, British, German, French or Canadian author was an advantage but an Indian and Italian (readers only) author was a disadvantage (Sud & Thelwall, 2016). Differences in the citation rates of international collaborations may also be affected by the willingness of national funding bodies to include overseas partners in projects primarily targeting local researchers. In this context, the USA and China seem particularly willing to fund overseas research collaborators (Huang & Huang, 2018).

There are field differences in the citation advantage of international collaboration, as suggested by descriptive studies analysing various citation impact and collaboration indicators. One early study found that international research tended to be more cited in all six fields examined: Clinical medicine, Biomedical research, Biology, Chemistry, Physics, Mathematics, Engineering and Earth and space sciences. Nevertheless, in some country/field combinations, international collaborative research is less cited than domestic research (Glänzel, 2001), so the pattern is not universal. A later detailed analysis of international differences in the citation advantage of international collaboration in 2004 for 60 countries found the greatest advantages for social science and engineering journal articles. The international collaboration advantage for physics and medicine was lower but still substantial (Lancho-Barrantes et al., 2013). It is not clear why international collaboration should be most advantageous in social sciences and engineering, however. Whitley's (2000) theory of organisational dimensions of disciplinary differences suggests that impact in fields with little agreement on what is worth researching (i.e., high technical task uncertainty) might be harder to gain and so international networks might help with wider audiences. This seems likely to include most social sciences.

Reasons for the citation advantage of international co-authorship

International research is more cited than domestic research (Adams, 2013), perhaps because international collaboration is intrinsically better for research in some countries/fields or as a side effect of other factors. International collaboration should be widely encouraged if the former is true, so it is important to be clear about the difference (Kato & Ando, 2013).

The reasons why international co-authorship might be intrinsically better are related to input diversity. Diversity in teams has been shown to be advantageous in some contexts (Curşeu & Pluut, 2013; Joshi & Roh, 2009), so it is plausible that an international team of researchers might benefit from greater variety in ideas, skills, expertise, resources, scientific cultures, and background knowledge (Wagner et al., 2019). For example, two biologists from different countries might have learned different aspects of the discipline, know about different applications, be connected to different types of industry, and draw upon their wider educations. Nevertheless, there is limited evidence to support this in academia (Hall, et al., 2018), and diversity could also lead to misunderstandings. For social science research, international collaboration can give access to different social and professional contexts for comparative or more diverse studies, as well as the expertise of the researchers involved (Teune, 1966). Some aspects of reference novelty associate with higher citation impact according to one study (Uzzi et al., 2013), indirectly agreeing with the diversity hypothesis. Nevertheless, international research seems to be more conventional (Wagner et al., 2019), which directly undermines the diversity claim. In fields where there is a high degree of agreement on tasks and goals (i.e., technical task certainty: Whitley, 2000), such as medicine, novelty may be less important than other aspects of quality, and obtaining adequate funding may be of primary importance (Gottlieb et al., 2019).

In theory, international research might be more cited not because of any international diversity bonus, but because it is better funded, is primarily possible for more cited researchers, or attracts more attention. In support of the first hypothesis, some funding streams require international collaboration (e.g., most EU programmes), and some large-scale projects need to be international to attract sufficient funding for expensive equipment (Thelwall, 2020). In medicine, international studies can increase the size of a dataset either by combining resources for labour intensive, expensive data collection or for access to a greater number of subjects with a particular condition (Cooley et al., 2003). This can produce more definitive, and presumably more cited, results, but would need more funding than more limited national research. Indirectly supporting the funding hypothesis, better funded US researchers have been shown to have wider collaboration networks (Bozeman & Corley, 2004).

In support of the second hypothesis, highly cited researchers might be more successful at attracting large international grants (e.g., most EU funding) or international collaborators. For example, higher cited researchers might be better positioned to reach out to overseas researchers to gain access to expensive equipment (e.g., Beaver, 2001), complementary skills or high-level expertise. National funding programs may also aim to connect local excellent researchers with high quality experts from other countries (Bloch et al., 2015). There is empirical evidence from two contexts (Italy and chemistry) to suggest that the overall citation advantage of international research is at least partly from more highly cited researchers engaging in it (Abramo et al., 2011; Kato & Ando, 2013). For chemistry research, the citation level of the participating researchers is not enough to explain the citation rates of their international collaborations, however (Kato & Ando, 2013).

The citation advantage of international research has also been hypothesised to be its greater audience due to interest from the national networks of the collaborators (Kato & Ando, 2013; Wagner et al., 2019). This fits with evidence that authors disproportionately cite researchers that are geographically close (the USA: Börner et al., 2006; 20 large countries in 2004: Lancho Barrantes et al., 2012).

Alternatively, international research may tend to be more cited than similarly collaborative national research for other indirect reasons, including those related to the different types of collaboration (Katz & Martin, 1997). For example, it seems likely that PhD supervisor/student collaborations would be more prevalent for national than international research. If true, and if PhD student work tended to be less (more) cited than other collaborative work, then this would reduce (increase) the citation impact of national collaborations overall compared to international collaborations. It is also possible that more casual collaborations based on random conversations are less likely for geographically distant researchers: these tend to collaborate less (Hoekman et al., 2010).

Other reasons for promoting international collaboration

International collaboration may be promoted for reasons other than its perceived benefits. It may be regarded as developmentally positive for the researchers involved, helping them to build networks and experiences that will benefit their later careers (Corley et al., 2019; see also Dusdal & Powell, 2021). Overlapping with citation impact considerations, it may be necessary for some types of research that is too expensive for individual countries or when there is too little national data to obtain strong result about a medical or other issue. For particularly important problems, governments may also wish to pool and direct the expertise of the best available researchers, as in targeted funding calls (e.g., Edler, 2012). International

bodies, such as the European Union, may also want to promote international collaboration as part of a wider policy of integration, such as the European Research Area (Defazio, et al., 2009; Georghiou, 2001; Kwiek, 2021).

Methods

A regression approach was used to model the effects of national contributions to the quality of journal articles, using a large multidisciplinary set of quality scores for individual UK journal articles.

Data

The raw data consisted of 148,977 journal articles submitted to REF2021 by UK universities for assessment and first published between 2014 and 2020. Each UK academic with a research component of their job was allowed to submit up to 5 outputs, with each institution submitting an average of 2.5 outputs per full-time equivalent research active member of staff. Members of staff and their associated outputs were submitted to one of 34 Units of Assessment (UoAs), each of which encompasses a broad area of research (see the first graph below for a list of names). Whilst outputs in low numbered UoAs tended to be journal articles, in higher numbered UoAs they were more likely to be books, book chapters, performances, or artworks.

The REF articles had been assessed by at least two peer reviewers (field experts and usually senior UK academics) with the most relevant expertise from within the UoA subpanel team. They assigned a quality score: 1* nationally recognised, 2* internationally recognised, 3* internationally excellent or 4* world leading. Quality was judged for originality, significance, and rigour, although each sub-panel could interpret these differently (REF, 2019). The 318 articles with a score of 0 were discarded for being potentially ineligible for reasons unrelated to quality (e.g., the submitting author was judged not to have made a significant contribution). This study only includes the journal articles and their provisional REF2021 scores from March 2021, as supplied confidentially by the REF team. The raw data was deleted, as required, in May 2021.

A journal article was sometimes submitted by different co-authors to the same or a different UoA. Such duplicates were removed and, in cases where an article had received different scores, the median score was used, or a random selection when there were two medians. The UoAs are grouped thematically into the four REF Main Panels, A: UoAs 1-6; B: UoAs 7-12; C: UoAs 13-24; D: UoAs 25-34 for some analyses to give additional statistical power to analyse country-level differences. When grouped into main panels, duplicate articles were again eliminated using the same procedure as before for differing quality scores.

RQ1, RQ2: Regression Analysis

Apparent author contributions to REF scores were modelled using ordinal logistic regression (Harrell, 2015) applied to fractional author contributions by country. The natural log of the number of authors was added to control for general collaboration advantages. The details are as follows.

The underlying model is that the quality of an article is contributed to equally by all authors, with this contribution potentially varying by affiliation country. For example, an article with seven UK authors and three Kenyan authors would have 70% UK and 30% Kenyan contributions. According to this model, the probability that this article has any given REF star

rating depends on these two inputs, as well as the log of the total number of authors: ln(10) in this case. The country component of this model is a simplification because the first author probably had the most responsibility for the article, or the corresponding author in some fields (e.g., as believed in medicine: Bhandari et al., 2014), and the last author may have had more responsibility than others (Duffy, 2017). Nevertheless, there is no agreed weighting scheme for author contributions and non-standard author ordering, such as alphabetical and partial alphabetical, are also common (Mongeon, et al., 2017). There did not seem to be a strong alternative to the equal contribution hypothesis because of these factors.

The team size logarithm term helps to account for larger teams tending to generate work with higher citation impact (Larivière et al., 2015) and possibly also a higher research quality. Thus, the model factors out the assumed collaboration quality bonus from any national bonus. Moreover, although the fractional author counting by country takes into account country contributions, if any country's UK co-authorships were primarily small roles in massively multi-authored papers, this could translate into an apparent country advantage (e.g., Guerrero Bote et al., 2013). The teams size factor included reduces the chance that this happens. A logarithm term is a reasonable approximation to the pattern that the citation increment of additional authors decreases as the team size increases (e.g., Thelwall & Maflahi, 2020).

Ordinal logistic regression was used because the scores 1*, 2*, 3*, and 4* are ordered but not necessarily on the numerical scale 1-2-3-4, so applying linear regression would require an additional assumption about the distance apart of the four scores. Financially, 1* and 2* are equivalent and 4* has four times the value of 3*, so a linear assumption would not fit with how the data is used for research funding. Ordinal logistic regression treats the four scores as ordered but not necessarily equidistant. The main assumption of the ordinal logistic regression approach used is that factors affecting the odds of an article attaining a higher score rather than a lower score do not depend on where the cut-off is set. Whilst it is not known whether this is true, it is a necessary assumption to run this kind of regression and it seems to be a reasonable hypothesis, in the absence of evidence to the contrary.

Ordinal logistic regression is similar to logistic regression except that the dependant variable is an ordered set. Like logistic regression, the regression coefficients can be transformed into odds ratios. An exponentiated regression coefficient (as shown in the graphs) of 1 indicates that the specified country has no effect on the probability of an article attaining any particular REF score. Values above 1 increase the odds that an article gets a higher score and values below 1 decrease these odds.

RQ3: Correlation analysis

The citation impact advantage ordinal regression coefficients of countries were correlated against three indicators to seek evidence of a pattern in the types of countries that had larger or more positive international research quality benefits in their collaborations with the UK. The obvious indicator to start with is one for the average citation impact of the collaborating country, as used in previous investigations into the citation advantage of international collaboration (Lancho-Barrantes et al., 2013; Guerrero Bote et al., 2013).

Average country citation impact was calculated using a copy of Scopus downloaded in January-February 2022. The Mean Normalised Log-transformed Citation Score (MNLCS) was calculated for each country (Thelwall, 2017). This uses field and year normalisation for comparability between fields, so that a value of 1 always indicates world average citation rate, irrespective of field and year. All citation counts were log-transformed with ln(1+x) in the

calculation to reduce the influence of individual highly cited articles. Departing from standard MLNCS practice but in keeping with the regression approach used (see below), country MNLCS were calculated on a fractional counting basis, so an article with 70% Kenyan authors would count as 0.7 of an article from Kenya. For each country the formula is:

$$MNLCS = \left(\sum_{i=1}^{n} \text{fraction}_\text{i} \times \ln(1 + citations_i) \,/field_i\right) / \left(\sum_{i=1}^{n} \text{fraction}_\text{i}\right)$$

Where $fraction_i$ is the fractional contribution of the country to article *i*, and $field_i$ is the average of the log normalised citations of all articles published in the same field and year as article *i*, irrespective of the authors' countries.

Since other factors may affect the strength of an international collaboration, three additional indicators were obtained to analyse in parallel with country MNLCS values. Per capita Gross Domestic Product (GDP) was obtained from the International Monetary Fund website (www.imf.org/en/Publications/WEO/weo-database/2022/April/) as an indicator of national wealth since this may influence funding for international projects. The United Nations Development Programme Human Development Index (HDI) and its education component were obtained from the HDR website (hdr.undp.org/en/data). These reflect overall human development and its educational component, both of which are relevant to the capability of a nation's researchers. The educational component mainly reflects the proportion of the population engaging in various levels of schooling and the overall HDI incorporates a wide range of factors related to levels of health, education, and standard of living throughout the population.

Results

RQ1 International collaboration overall

International collaboration, after controlling for collaboration in general, associates statistically significantly with an increased probability of higher REF quality scores in all main panels (Figure 1). There is an international collaboration advantage in most (27 out of 34) UoAs, and it is statistically significant in just under half: 15 out of 34. International collaboration seems to have the greatest advantage in business and economics, and it is statistically significantly apparently a disadvantage only in the hybrid UoA 34: Communication, Cultural and Media Studies, Library and Information Management. The control variable for collaboration increases the odds of a higher quality score in 27 out of 34 UoAs and all four panels (not shown). The exceptions are UoAs 16, 17, 22, 28, 30, 31, and 32.

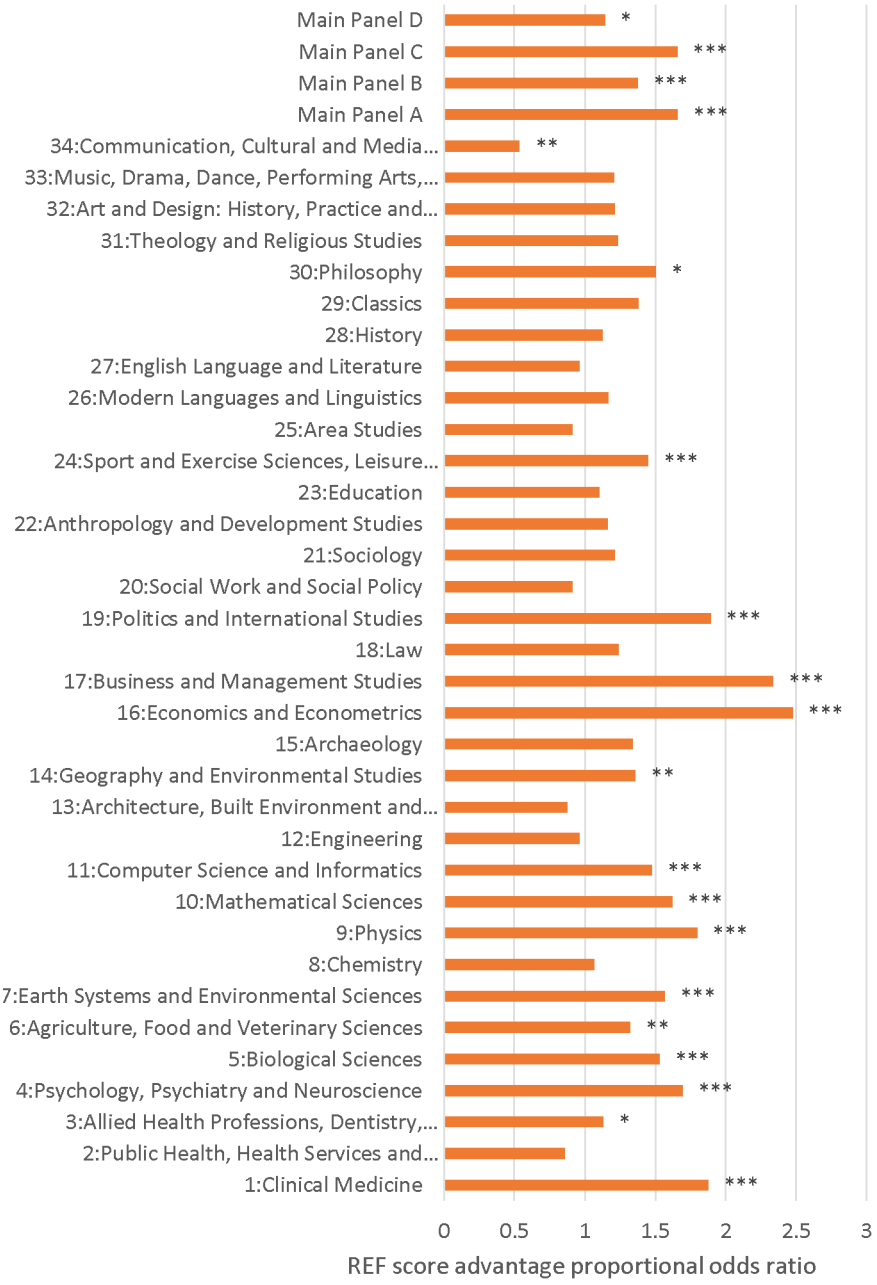

Figure 1. The average score advantage for articles with non-UK authors, as calculated through ordinal logistic regression with log(authors) as a control variable. The reference score is the UK average. A value of 1 indicates no advantage compared to UK-only authors. Stars indicate statistical significance: *** p<0.001; ** p<0.01; * p<0.05.

RQ2 International collaboration by country

From a descriptive statistics perspective without controlling for collaboration or using regression, average quality scores by country (fractional counting) gives an overview of international quality differences in UK co-authored outputs. For Main Panel A (UoAs 1-6), which mainly covers the health and life sciences, co-authorship with many countries associates with higher average research quality scores than domestic UK authorship (Figure 2). For example, articles with an author from Iceland, on average, scored 0.46 higher than articles from only UK authors and articles from Malaysia, on average scored 0.42 lower, using

fractional counting. Most countries have an apparent score advantage compared to the UK, partly because of the collaboration advantage of international authorship, which is not controlled for in Figure 2.

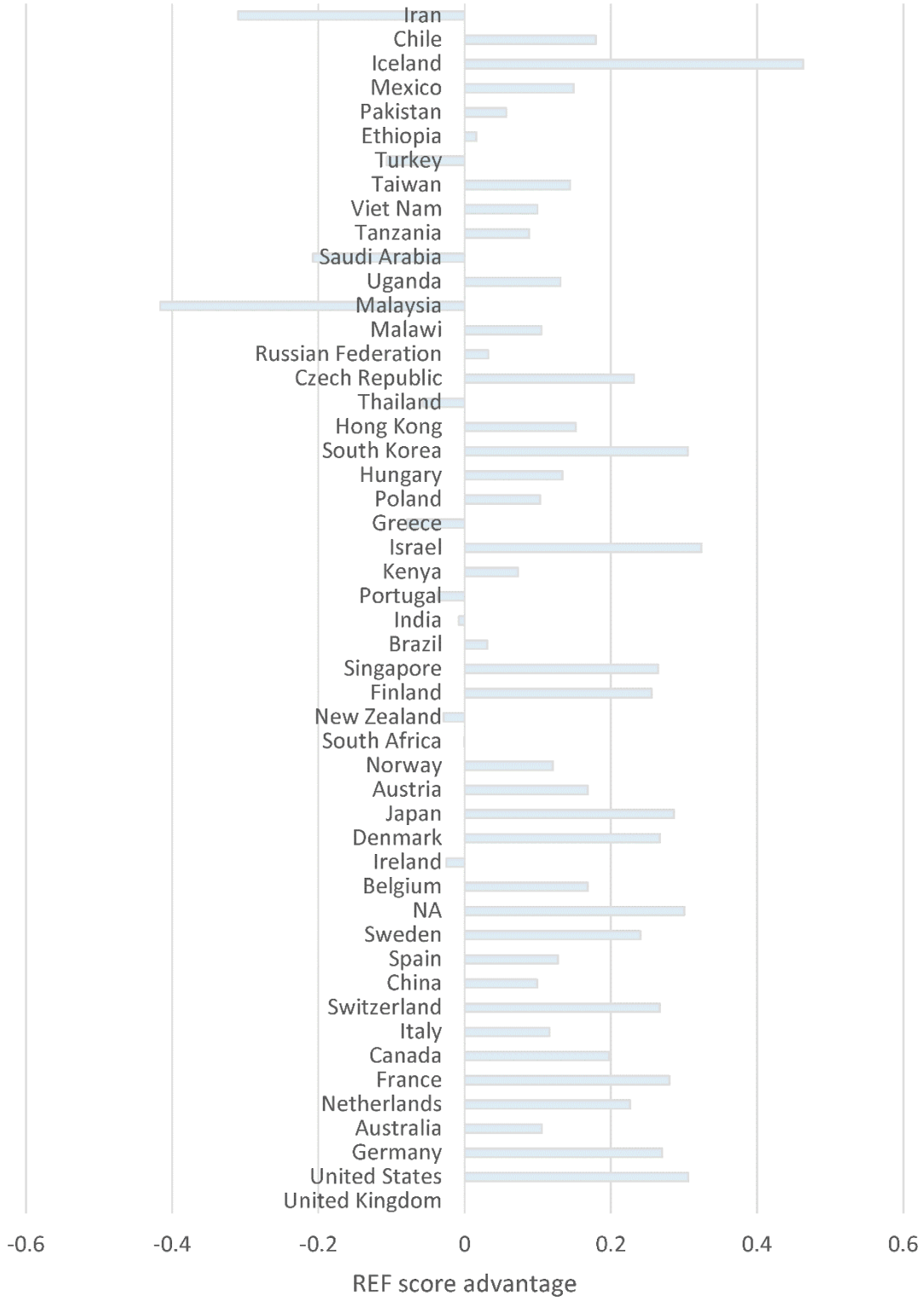

Figure 2. The weighted mean average score advantage for articles with an author from the specified country for articles submitted to Main Panel A (mainly health and life sciences). The

reference score is the UK average. Qualification: the 50 countries co-authoring the most Main Panel A journal articles, using fractional author counting. NA: no country assigned in Scopus.

The ordinal regression model gives superior information to the above descriptive statistics about the average REF score advantage because it takes into account the relative contributions of all countries to an article, as well as controlling for collaboration in general. The regression confirms that author contributions from some countries increase the odds of a higher REF score in Main Panel A but contributions from other countries decrease those odds (Figure 3). Note that whilst articles with an author from China tend to score above UK-only articles (Figure 2), there is little or no differences in the logistic regression (Figure 3), showing the importance of the regression approach. This could occur, for example, if collaborations with China tended also to include the USA.

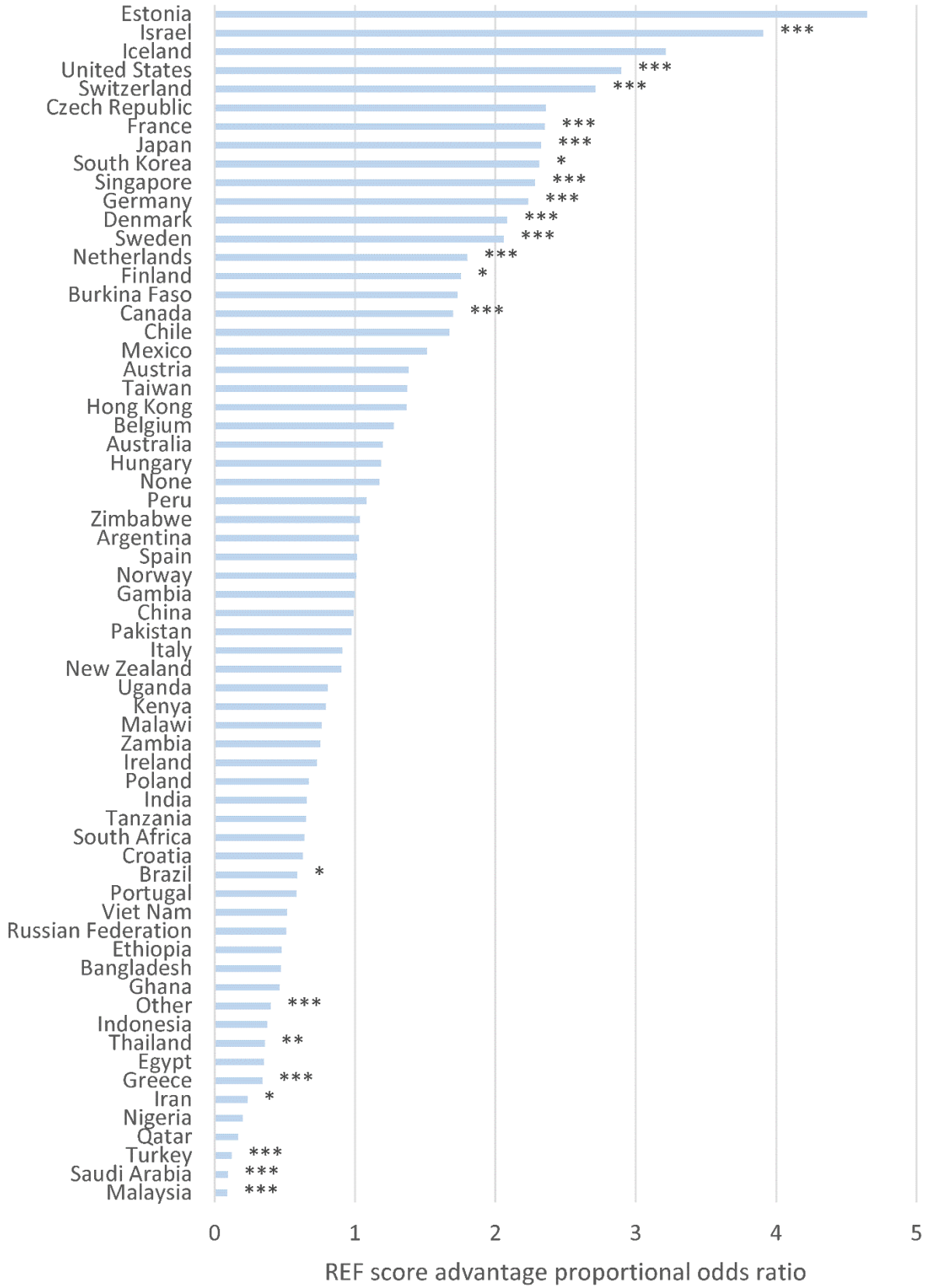

Figure 3. Proportional odds ratios from an ordinal logistic regression for REF score for journal articles from Main Panel A, with log(authors) as a control variable. Qualification: At least 10 journal articles using fractional author counting. Stars indicate statistical significance: *** p<0.001; ** p<0.01; * p<0.05. The logged authors proportional odds ratio is 2.11.

For the mainly physical science and engineering articles of Main Panel B, the contributions of countries outside the UK are less likely to associate with a higher REF score overall than for Main Panel A. this is clearest from Figure A1 against Figure 2, but is also evident by comparing Figure 4 to Figure 3. There are considerable overlaps in terms of countries with advantages in both Figures 3 and 4, or with disadvantages in both.

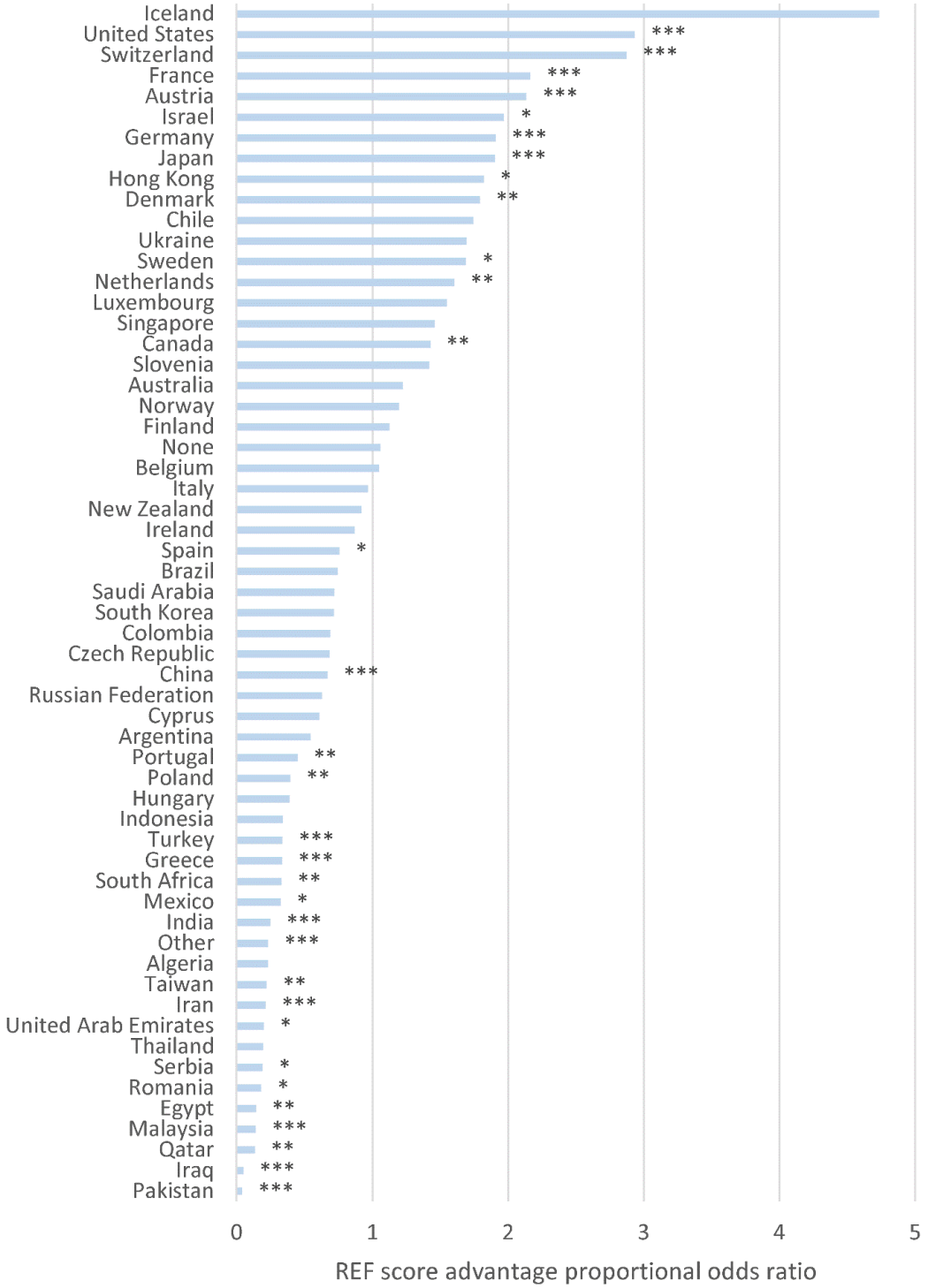

Figure 4. As Figure 3 for Main Panel B. The logged authors proportional odds ratio is 1.31.

The predominantly social science research of Main Panel C shows advantages for many countries (Figures A2, 5). There are again considerable overlaps with Main Panel A and Main Panel B for countries with advantages in two of the three, or disadvantages in two of the three.

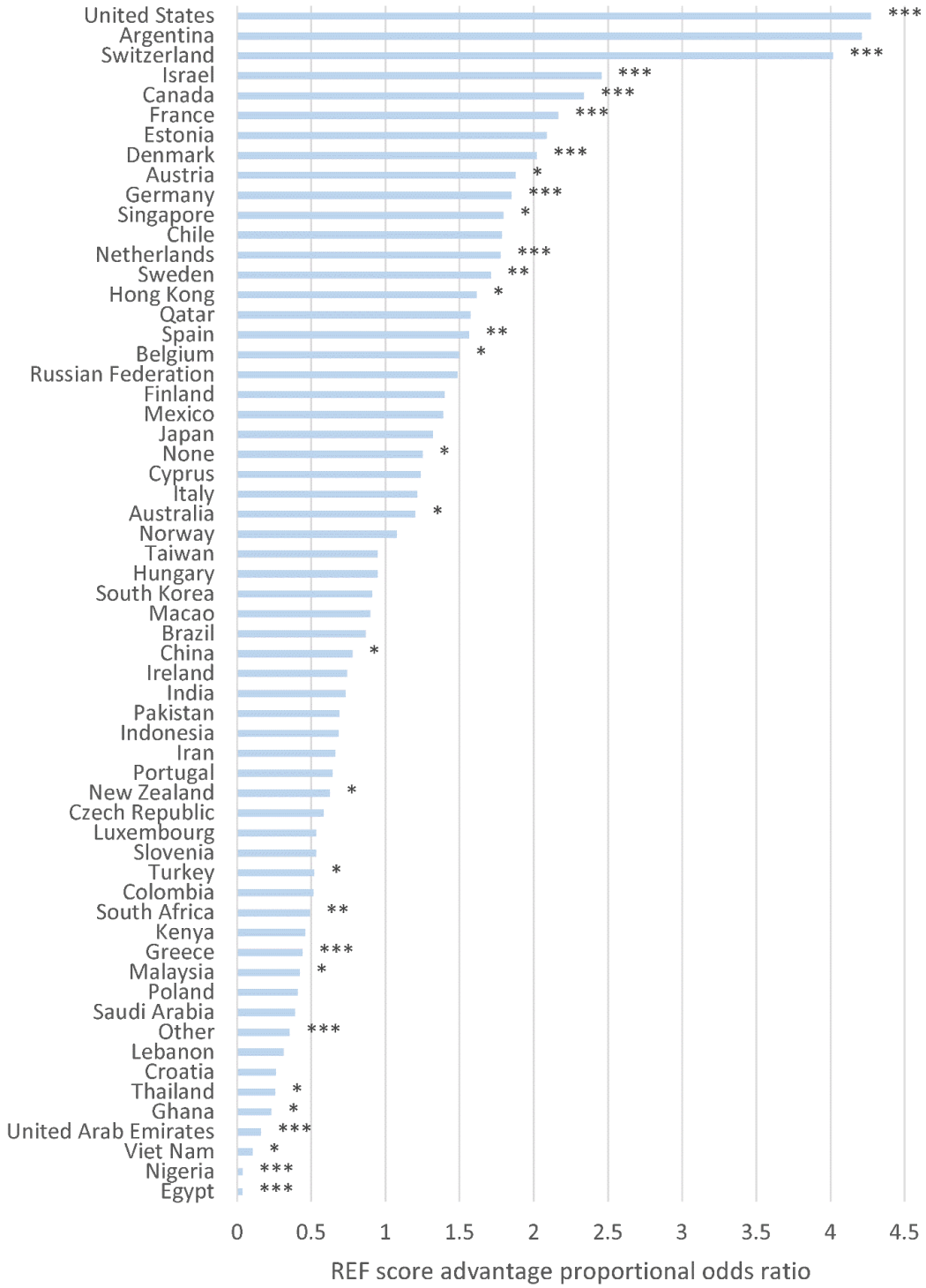

Figure 5. As Figure 3 for Main Panel C. The logged authors proportional odds ratio is 1.24.

Main Panel D includes mainly arts and humanities research, with some social sciences fields too. It has relatively few journal articles because the UoAs tend to be small and researchers

tend to submit books, book chapters, performances or artworks instead. There is a weak tendency for non-UK contributions to benefit journal article quality (Figures A3, 6), but with the clearest evidence from a different set of countries to the other Main Panels.

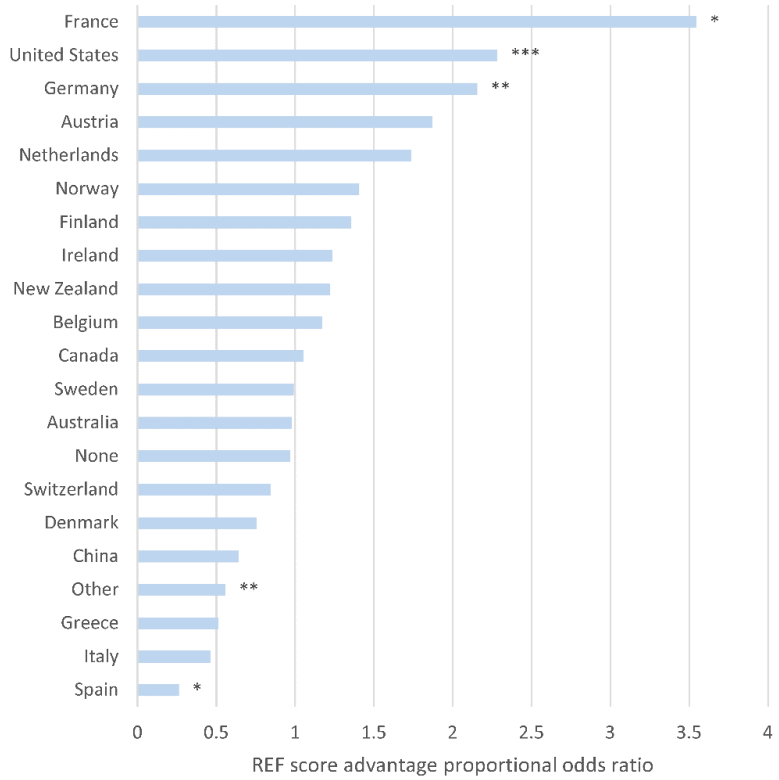

Figure 6. As Figure 3 for Main Panel D. The logged authors proportional odds ratio is 1.14.

RQ3 Reasons for differing country collaboration advantages

There is a moderate positive correlation between the average citation impact of a country (MNLCS) and the ordinal regression coefficients for that country in Main Panels A, B, and C. The correlation is strong if only statistically significant coefficients are included (Table 1). This confirms that collaboration with countries independently producing higher impact research tends to give a greater likelihood of higher quality collaborative research.

Nevertheless, correlations with regression coefficients tend to be higher for per capita GDP than for MNLCS, suggesting that national wealth may be more important for collaboration partners than highly cited research (Table 1). The correlations are even higher for HDI and its education component, especially for Main Panels A and B, suggesting that population education, health and standard of living are even more important than national wealth. Although citation rates can sometimes be misleading for small countries due to the dominance of important international collaborations or foreign-backed research institutes (Confraria et al., 2017), most of the countries analysed publish enough for this not to be a likely alternative explanation.

Table 1. Pearson correlations between regression coefficients and country MNLCS with fractional counting, per capita GDP, HDI, and Education for significant regression coefficients (all regression coefficients). The highest correlation in each row is bold.

Panel	MNLCS	Per capita GDP	HDI	Education	Countries
A	0.557 (0.334)	0.592 (0.397)	**0.780** (0.444)	0.727 (0.463)	20 (62)

B	0.626 (0.459)	0.562 (0.441)	**0.674** (0.581)	0.649 (0.564)	31 (55)
C	0.617 (0.337)	0.648 (0.326)	0.655 (0.471)	**0.664** (0.478)	27 (56)
D	-0.093 (0.004)	**0.442** (0.220)	0.051 (0.201)	0.021 (0.178)	4 (19)

At the individual country level, in Main Panels A to C, the international collaboration advantage applies to all countries that exceed the UK's MNLCS (1.17) and many that don't (Figures 10-12). The same is true for Main Panel D (only four countries and not shown). The closest to an exception is Qatar in Main Panel B, which has an MNLCS close to the UKs but its collaborations with the UK tend to reduce the odds of a higher quality article, from the UK perspective. In broad terms, the results suggest that UK collaborations with other advanced countries tend to be beneficial for the quality of research, irrespective of the citation impact of the collaborating country. There are exceptions to this rule, however, such as Spain for Main Panel B (a disadvantage) although Spanish collaboration is an advantage for Main Panel C.

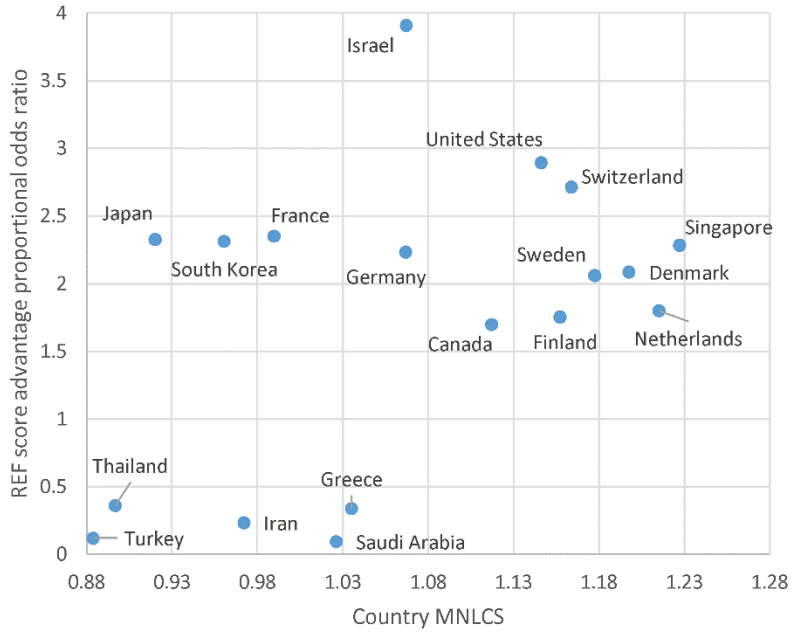

Figure 7. REF scorer advantage proportional odds ratio against country MNLCS (fractional counting) for Main Panel A UoAs combined, with only statistically significant country ratios shown. The UK MNLCS is 1.17

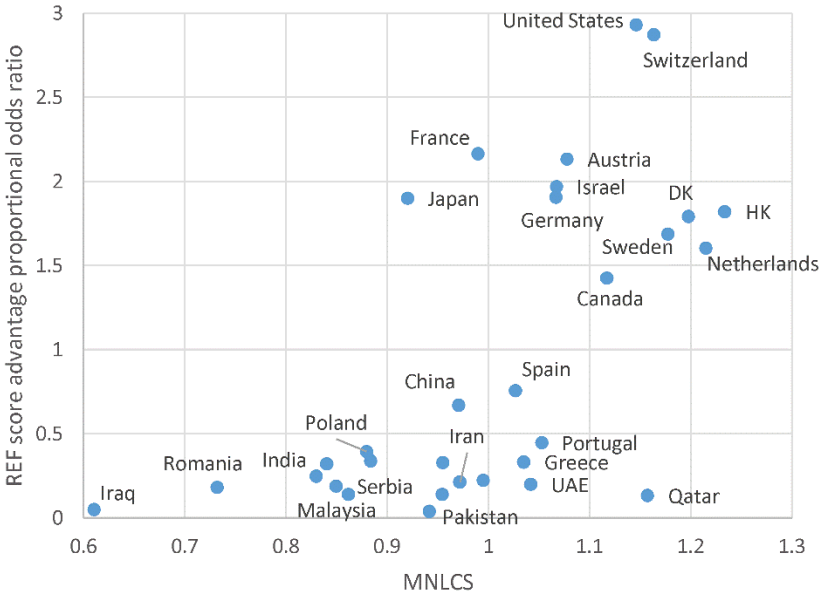

Figure 8. As figure 7 for Main Panel B. DK=Denmark; HK= Hong Kong (China).

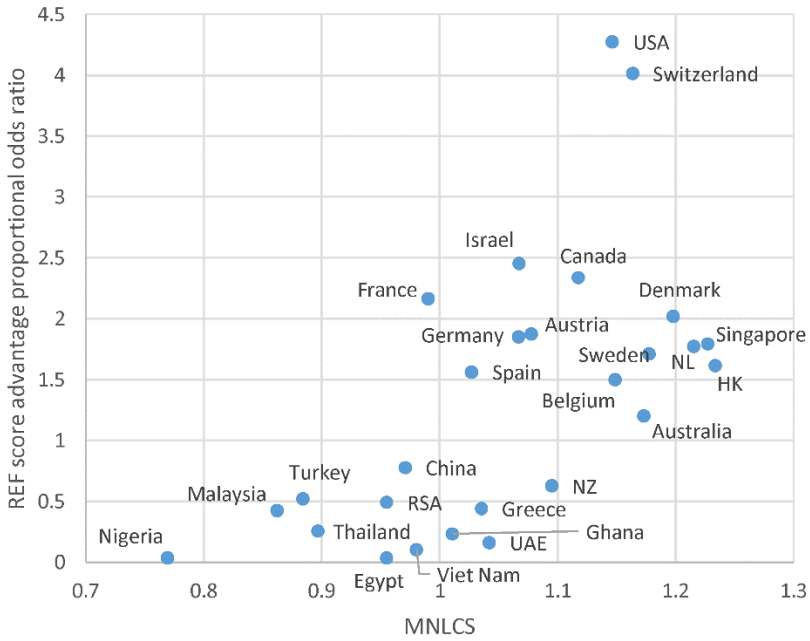

Figure 9. As Figure 10 for Main Panel C.

Discussion

The results are limited to the UK although they may apply to broadly similar countries in terms of the national average citation impact, per capita GDP or HDI. The results are also limited to the self-selected UK REF submissions, which represent articles considered to be the authors' best and therefore underrepresent weaker articles. In UoA 8 Chemistry and UoA 9 Physics, researchers may have often written far more articles than they submitted. If they had co-authored many low-quality international articles, then these would be hidden from the REF due the selection process.

The results do not prove a cause-and-effect relationship between international collaboration and research quality, for the reasons discussed above. For example, a researcher may use their best idea to attract funding for an international project rather than

the international collaboration generating the good idea or otherwise stronger research. Similarly, better regarded UK researchers might find it easier to attract international partners for collaborative research. From a different perspective, since there are many funding programmes exclusively for international research, such as most from the EU framework programmes, international research may be better funded than national research, or may attract better participants because they were successful in funding competitions. In these cases, the international dimension would not improve the result directly but would be an effect of the underlying cause.

The strong positive relationship between international research collaboration and research quality contrasts with the very weak relationship found in the most relevant prior study, although it did not use fractional counting (Bornmann, 2017). The F1000Prime articles would probably fit mainly in UoA 1, which had a strong positive association with international collaboration. The difference may be due to the UK focus here, changes over time, the different model for accounting for the effect of international collaboration, or the different purposes of the quality scores used. The F1000Prime system used three ratings, 1: good; 2: very good; 3: exceptional and these were awarded by at least two peer reviewers, so this is a close parallel with the REF (four ratings, but 1* is rare, and usually two reviewers). Presumably the F1000Prime reviewers would be less uniformly senior than REF assessors and their judgments may have been less seriously considered and less norm referenced. F1000Prime articles are selected by the reviewers so may be more uniformly high quality than REF articles. This would reduce the statistical power of any tests. Thus, the main differences are probably the scope (author-selected UK articles vs. reviewer-selected international articles), the fractional counting model, and the quality of the reviewing processes. The stronger results compared to a previous REF study (HEFCE, 2015) are probably due to the fractional counting model and parsimonious independent variables.

The finding that the outputs of collaborations with countries that are less wealthy, and with lower human development, tend to be of lower quality than similarly collaborative national UK research is worrying. There are multiple possible explanations. Global South research is less cited (Confraria et al., 2017) and may be lower quality than Global North research due to lower English proficiency, less funding (Manet al., 2004), weaker research training, and less research support. On this basis, a Global South co-author would tend to contribute less expertise and resources than a Global North contributor. Collaborations with UK researchers might then have a partly translatory goal of helping Global South researchers to access a Global North audience (Martinez & Sá, 2020) and might receive less funding on this basis. The collaborative articles may thus be steps towards independence for the Global South researcher in a way that the UK REF evaluation system does not recognise (although it might in the research environment component of the REF). Development funding agencies are influenced by political considerations and often emphasise ambitious non-academic goals, which may undermine the academic scoring of such projects (Currie-Alder, 2015). Alternatively, the UK REF might not recognise the priorities of Global South researchers and unfairly award lower quality scores to research that they led or participated in. A similar issue has been raised for Maori research in New Zealand (Roa et al., 2009). To give an extreme example, Dengue science might be considered much less important than cancer research in the UK, even though this mosquito-borne viral infection is an increasing threat in tropical and subtropical countries (Wellekens et al., 2022). Global South research might also sometimes have perspectives that some Global North evaluators might find unpalatable (Giwa, 2015;

Openjuru et al., 2015) or inappropriate, such as criticism of colonial legacies or ongoing imperialism, or a greater focus on ethics and societal value.

Conclusion

Although not definitive, the results give the first large scale direct evidence that journal articles written by international teams tend to be higher quality than national research, even after accounting for team size. At the country level, this seems to apply to nearly all advanced economies, even those with low citation impact research. These findings apply primarily to the UK but seem likely to also be true for other countries that tend to author high impact research and may apply to most or all advanced economies. The overall international collaboration advantage for the UK is due to collaboration with advanced economies being more frequent than collaboration with other economies, since these tend to produce lower quality research, at least as judged by REF panel members.

From a pure research quality perspective, as judged by Global North standards, collaborations with less advanced countries are not optimal for research quality. Nevertheless, these collaborations seem likely to satisfy wider goals, such as creating international connections that may be useful for other purposes (Wagner, 2008) or trying to overcome the legacies of colonialism (especially for the UK) (Allpress et al., 2010). Even from a UK REF perspective, researchers may be rewarded for achieving these wider goals through the environment and impact case study REF components.

In terms of research policy implications, the results suggest that funders and research managers should continue to encourage international collaborative research because it not only tends to be more cited but also tends to be higher quality, at least in terms of journal articles. Policy makers should also be aware that international collaboration may be disadvantageous in some fields, however (Figure 1) and have little effect in others. In this context, it seems sensible to be judicious in selecting fields to target for international research funding and to allow national research to blossom in fields where it is more likely to be excellent.

Acknowledgement

This study was funded by Research England, Scottish Funding Council, Higher Education Funding Council for Wales, and Department for the Economy, Northern Ireland as part of the Future Research Assessment Programme (https://www.jisc.ac.uk/future-research-assessment-programme). The funders had no role in the design or execution of this study. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funders.

References

Abramo, G., D'Angelo, A. C., & Murgia, G. (2017). The relationship among research productivity, research collaboration, and their determinants. Journal of Informetrics, 11(4), 1016-1030.

Abramo, G., D'Angelo, C. A., & Solazzi, M. (2011). The relationship between scientists' research performance and the degree of internationalization of their research. Scientometrics, 86(3), 629-643.

Adams, J. (2013). The fourth age of research. Nature, 497(7451), 557-560.

Adams, J. D., Black, G. C., Clemmons, J. R., & Stephan, P. E. (2005). Scientific teams and institutional collaborations: Evidence from US universities, 1981–1999. Research policy, 34(3), 259-285.

Allpress, J. A., Barlow, F. K., Brown, R., & Louis, W. R. (2010). Atoning for colonial injustices: Group-based shame and guilt motivate support for reparation. International Journal of Conflict and Violence, 4(1), 75-88.

Beaver, D. (2001). Reflections on scientific collaboration (and its study): past, present, and future. Scientometrics, 52(3), 365-377.

Bhandari, M., Guyatt, G. H., Kulkarni, A. V., Devereaux, P. J., Leece, P., Bajammal, S., & Busse, J. W. (2014). Perceptions of authors' contributions are influenced by both byline order and designation of corresponding author. Journal of clinical epidemiology, 67(9), 1049-1054.

Bloch, C., & Sørensen, M. P. (2015). The size of research funding: Trends and implications. *Science and Public Policy*, 42(1), 30-43.

Bolton, W. S., Chapman, S. J., Corrigan, N., Croft, J., Collinson, F., Brown, J. M., & Jayne, D. G. (2021). The incidence of low anterior resection syndrome as assessed in an international randomized controlled trial (MRC/NIHR ROLARR). Annals of Surgery, 274(6), e1223-e1229.

Börner, K., Penumarthy, S., Meiss, M., & Ke, W. (2006). Mapping the diffusion of scholarly knowledge among major US research institutions. Scientometrics, 68(3), 415-426.

Bornmann, L. (2017). Is collaboration among scientists related to the citation impact of papers because their quality increases with collaboration? An analysis based on data from F1000Prime and normalized citation scores. Journal of the Association for Information Science and Technology, 68(4), 1036-1047.

Bozeman, B., & Corley, E. (2004). Scientists' collaboration strategies: implications for scientific and technical human capital. *Research Policy*, 33(4), 599-616.

Confraria, H., Godinho, M. M., & Wang, L. (2017). Determinants of citation impact: A comparative analysis of the Global South versus the Global North. Research Policy, 46(1), 265-279.

Cooley, M. E., Sarna, L., Brown, J. K., Williams, R. D., Chernecky, C., Padilla, G., & Danao, L. L. (2003). Challenges of recruitment and retention in multisite clinical research. Cancer nursing, 26(5), 376-386.

Corley, E. A., Bozeman, B., Zhang, X., & Tsai, C. C. (2019). The expanded scientific and technical human capital model: the addition of a cultural dimension. The Journal of Technology Transfer, 44(3), 681-699.

Currie-Alder, B. (2015). Research for the developing world: Public funding from Australia, Canada, and the UK. Oxford, UK: Oxford University Press.

Curşeu, P. L., & Pluut, H. (2013). Student groups as learning entities: The effect of group diversity and teamwork quality on groups' cognitive complexity. Studies in Higher Education, 38(1), 87-103.

Defazio, D., Lockett, A., & Wright, M. (2009). Funding incentives, collaborative dynamics and scientific productivity: Evidence from the EU framework program. *Research Policy*, 38(2), 293-305.

Duffy, M. A. (2017). Last and corresponding authorship practices in ecology. Ecology and Evolution, 7(21), 8876-8887.

Dusdal, J., & Powell, J. (2021). Benefits, motivations, and challenges of international collaborative research: A sociology of science case study. *Science and Public Policy*, 48(2), 235-245.

Edler, J. (2012). Toward variable funding for international science. Science, 338(6105), 331-332.

Foster, J. G., Rzhetsky, A., & Evans, J. A. (2015). Tradition and innovation in scientists' research strategies. American Sociological Review, 80(5), 875-908.

Freeman, R. B., Ganguli, I., & Murciano-Goroff, R. (2014). Why and wherefore of increased scientific collaboration. In A. Jaffe & B. Jones (Eds.), The changing frontier: Rethinking science and innovation policy (pp. 17–48). Chicago, IL: University of Chicago Press.

Gazni, A., Sugimoto, C. R., & Didegah, F. (2012). Mapping world scientific collaboration: Authors, institutions, and countries. Journal of the American Society for Information Science and Technology, 63(2), 323-335.

Georghiou, L. (2001). Evolving frameworks for European collaboration in research and technology. Research Policy, 30(6), 891-903.

Gingras, Y., & Khelfaoui, M. (2018). Assessing the effect of the United States'"citation advantage" on other countries' scientific impact as measured in the Web of Science (WoS) database. Scientometrics, 114(2), 517-532.

Giwa, A. (2015). Insider/outsider issues for development researchers from the Global South. Geography Compass, 9(6), 316-326.

Glänzel, W. (2001). National characteristics in international scientific co-authorship relations. Scientometrics, 51(1), 69-115.

Gladstone, A., Landsbury, R., Stieber, J., Treu, T., & Weiss, M. (Eds.). (2019). Current issues in labour relations: An international perspective. Berlin, Germany: Walter de Gruyter GmbH & Co KG.

Gottlieb, M., Lee, S., Burkhardt, J., Carlson, J. N., King, A. M., Wong, A. H., & Santen, S. A. (2019). Show me the money: successfully obtaining grant funding in medical education. Western Journal of Emergency Medicine, 20(1), 71-77.

Guerrero Bote, V. P., Olmeda-Gómez, C., & de Moya-Anegón, F. (2013). Quantifying the benefits of international scientific collaboration. Journal of the American Society for Information Science and Technology, 64(2), 392-404.

Hall, K. L., Vogel, A. L., Huang, G. C., Serrano, K. J., Rice, E. L., Tsakraklides, S. P., & Fiore, S. M. (2018). The science of team science: A review of the empirical evidence and research gaps on collaboration in science. American psychologist, 73(4), 532.

Harrell, F. E. (2015). Ordinal logistic regression. In: Harrell, F. E. (ed.) Regression modeling strategies (2ed.) (pp. 311-325). Berlin, Germany: Springer.

HEFCE (2015). Correlation analysis of REF2014 scores and metrics (Supplementary Report II to the independent Review of the Role of Metrics in Research Assessment and Management). London, UK: Higher Education Funding Council for England.

Hoekman, J., Frenken, K., & Tijssen, R. J. (2010). Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe. Research Policy, 39(5), 662-673.

Hsiehchen, D., Espinoza, M., & Hsieh, A. (2018). Evolution of collaboration and optimization of impact: self-organization in multinational research. Scientometrics, 117(1), 391-407.

Huang, M. H., & Huang, M. J. (2018). An analysis of global research funding from subject field and funding agencies perspectives in the G9 countries. Scientometrics, 115(2), 833-847.

Iaria, A., Schwarz, C., & Waldinger, F. (2018). Frontier knowledge and scientific production: evidence from the collapse of international science. The Quarterly Journal of Economics, 133(2), 927-991.

Joshi, A., & Roh, H. (2009). The role of context in work team diversity research: A meta-analytic review. Academy of management journal, 52(3), 599-627.

Kato, M., & Ando, A. (2013). The relationship between research performance and international collaboration in chemistry. Scientometrics, 97(3), 535-553.

Kato, M., & Ando, A. (2017). National ties of international scientific collaboration and researcher mobility found in Nature and Science. Scientometrics, 110(2), 673-694.

Katz, J. S., & Martin, B. R. (1997). What is research collaboration? Research Policy, 26(1), 1-18.

Kwiek, M. (2021). What large-scale publication and citation data tell us about international research collaboration in Europe: Changing national patterns in global contexts. Studies in Higher Education, 46(12), 2629-2649.

Lancho-Barrantes, B. S., Guerrero-Bote, V. P., & de Moya-Anegón, F. (2013). Citation increments between collaborating countries. Scientometrics, 94(3), 817-831.

Lancho Barrantes, B. S., Guerrero Bote, V. P., Rodríguez, Z. C., & de Moya Anegón, F. (2012). Citation flows in the zones of influence of scientific collaborations. Journal of the American Society for Information Science and Technology, 63(3), 481-489.

Langfeldt, L., Nedeva, M., Sörlin, S., & Thomas, D. A. (2020). Co-existing notions of research quality: A framework to study context-specific understandings of good research. *Minerva*, 58(1), 115-137.

Lariviere, V., Gingras, Y., & Archambault, É. (2006). Canadian collaboration networks: A comparative analysis of the natural sciences, social sciences and the humanities. Scientometrics, 68(3), 519-533.

Larivière, V., Gingras, Y., Sugimoto, C. R., & Tsou, A. (2015). Team size matters: Collaboration and scientific impact since 1900. Journal of the Association for Information Science and Technology, 66(7), 1323-1332.

Leydesdorff, L., & Wagner, C. S. (2008). International collaboration in science and the formation of a core group. Journal of informetrics, 2(4), 317-325.

Man, J. P., Weinkauf, J. G., Tsang, M., & Sin, J. H. D. D. (2004). Why do some countries publish more than others? An international comparison of research funding, English proficiency and publication output in highly ranked general medical journals. European Journal of Epidemiology, 19(8), 811-817.

Martinez, M., & Sá, C. (2020). Highly cited in the south: International collaboration and research recognition among Brazil's highly cited researchers. Journal of Studies in International Education, 24(1), 39-58.

Mongeon, P., Smith, E., Joyal, B., & Larivière, V. (2017). The rise of the middle author: Investigating collaboration and division of labor in biomedical research using partial alphabetical authorship. PloS One, 12(9), e0184601.

Olechnicka, A., Ploszaj, A., & Celińska-Janowicz, D. (2019). The geography of scientific collaboration. Oxford, UK: Taylor & Francis.

Openjuru, G. L., Jaitli, N., Tandon, R., & Hall, B. (2015). Despite knowledge democracy and community-based participatory action research: Voices from the global south and excluded north still missing. Action Research, 13(3), 219-229.

Presser, S. (1980). Collaboration and the quality of research. *Social Studies of Science*, 10(1), 95-101.

REF (2019). Panel criteria and working methods. https://www.ref.ac.uk/publications-and-reports/panel-criteria-and-working-methods-201902/

Ribeiro, L. C., Rapini, M. S., Silva, L. A., & Albuquerque, E. M. (2018). Growth patterns of the network of international collaboration in science. Scientometrics, 114(1), 159-179.

Roa, T., Beggs, J. R., Williams, J., & Moller, H. (2009). New Zealand's performance based research funding (PBRF) model undermines Maori research. Journal of the Royal Society of New Zealand, 39(4), 233-238. DOI: 10.1080/03014220909510587

Sud, P., & Thelwall, M. (2016). Not all international collaboration is beneficial: The Mendeley readership and citation impact of biochemical research collaboration. *Journal of the Association for Information Science and Technology*, 67(8), 1849-1857.

Teune, H. (1966). Advantages and problems of international collaboration in social science research. Background, 10(2), 177-192.

Thelwall, M., & Maflahi, N. (2020). Academic collaboration rates and citation associations vary substantially between countries and fields. Journal of the Association for Information Science and Technology, 71(8), 968-978.

Thelwall, M. Munafò, M., Mas-Bleda, A., Stuart, E., Makita, M., Weigert, V., Khan, N., Drax, K. & Kousha, K. (2020). Is useful research data usually shared? An investigation of genome-wide association study summary statistics. Plos One, 15(2): e0229578. https://doi.org/10.1371/journal.pone.0229578

Thelwall, M. (2017). Three practical field normalised alternative indicator formulae for research evaluation. Journal of Informetrics, 11(1), 128–151. 10.1016/j.joi.2016.12.002

Thelwall, M. (2020). Large publishing consortia produce higher citation impact research but coauthor contributions are hard to evaluate. Quantitative Science Studies, 1(1), 290-302.

Thiers, F. A., Sinskey, A. J., & Berndt, E. R. (2008). Trends in the globalization of clinical trials. Nature Reviews Drug Discovery, 7(1), 13-14.

Tritschler, T., Mathieu, M. E., Skeith, L., Rodger, M., Middeldorp, S., Brighton, T., & International Network of VENous Thromboembolism Clinical Research Networks INVENT-VTE. (2020). Anticoagulant interventions in hospitalized patients with COVID-19: A scoping review of randomized controlled trials and call for international collaboration. Journal of Thrombosis and Haemostasis, 18(11), 2958-2967.

Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013). Atypical combinations and scientific impact. Science, 342(6157), 468-472.

Wagner, C. S. (2005). Six case studies of international collaboration in science. Scientometrics, 62(1), 3-26.

Wagner, C. S. (2008). The new invisible college: Science for development. Washington DC: Brookings Institution Press.

Wagner, C. S., Whetsell, T. A., & Mukherjee, S. (2019). International research collaboration: Novelty, conventionality, and atypicality in knowledge recombination. *Research Policy*, 48(5), 1260-1270.

Wellekens, K., Betrains, A., De Munter, P., & Peetermans, W. (2022). Dengue: current state one year before WHO 2010–2020 goals. Acta Clinica Belgica, 77(2), 436-444.

Whitley, R. (2000). The intellectual and social organization of the sciences (2 ed). Oxford, UK: Oxford University Press.

Zhou, P., Cai, X., & Lyu, X. (2020). An in-depth analysis of government funding and international collaboration in scientific research. Scientometrics, 125(2), 1331-1347.

Appendix 1: Country

Figure A1. As Figure 2 for Main Panel B (mainly physical science and engineering).

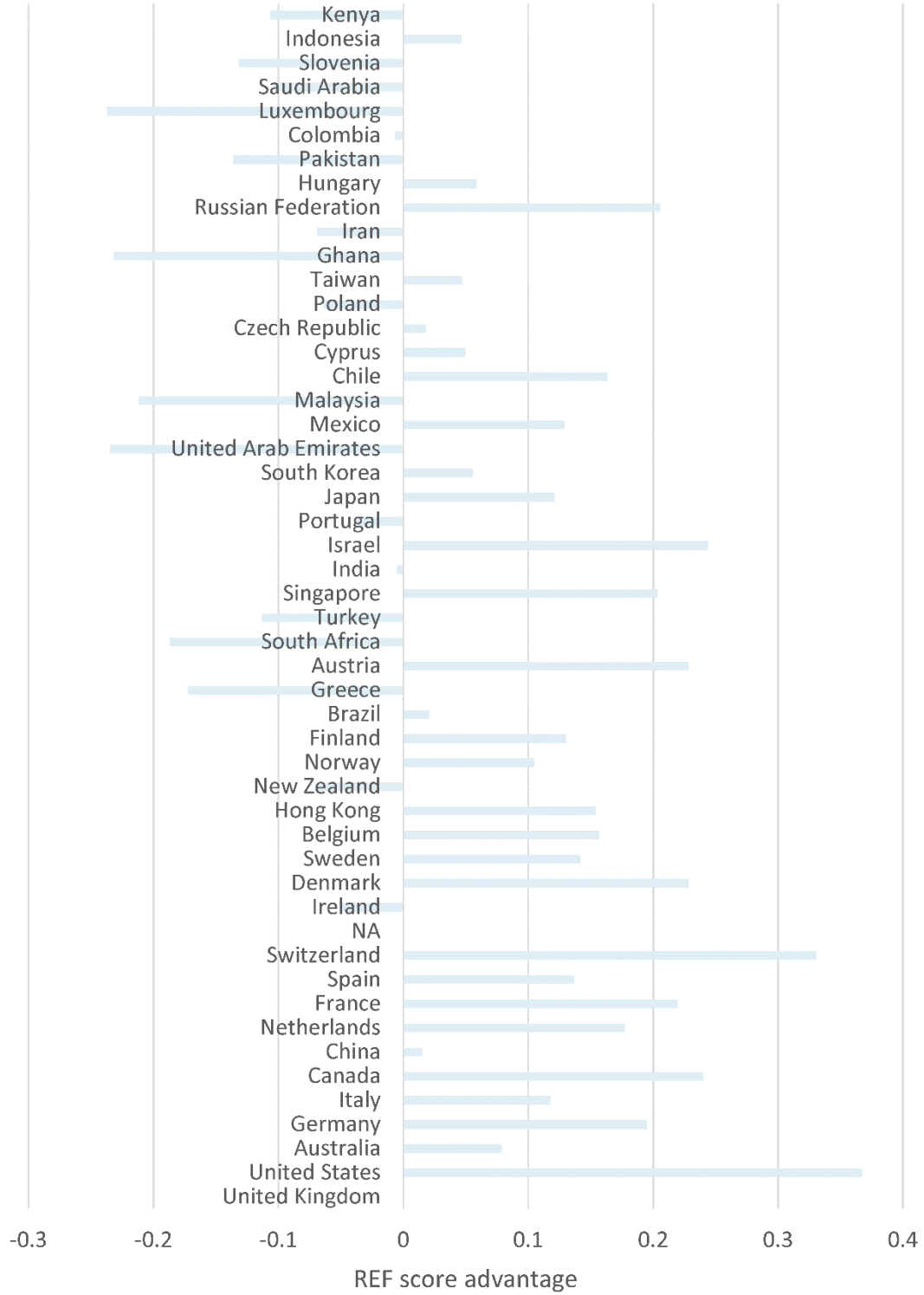

Figure A2. As Figure 2 for Main Panel C (mainly social sciences).

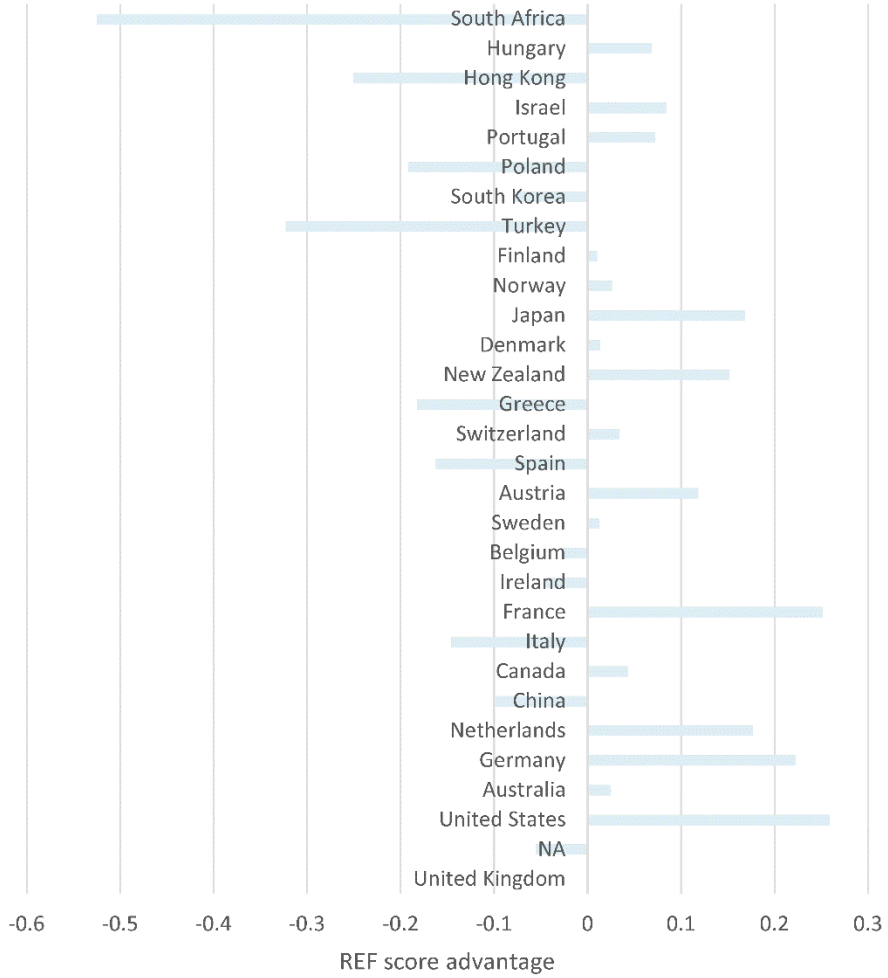

Figure A3. As Figure 2 for Main Panel D (mainly arts and humanities), but for the top 30 only.