

How international is internationally collaborated research? Heritage composition of Russia's international collaboration network

Maria Karaulova¹ and Abdullah Gök²

¹*maria.karaulova@manchester.ac.uk*

Manchester Institute of Innovation Research (MIOIR), University of Manchester, Oxford Road, M13 9PL,
Manchester (UK)

²*abdullah.gok@strath.ac.uk*

Hunter Centre for Entrepreneurship, Strathclyde Business School, University of
Strathclyde, 130 Rottenrow, G4 0GE, Glasgow (UK)

Abstract

International research performance indicators attain increased attention in science policy. More studies point to ongoing global bias in production, composition and assessment of research performance metrics (Rafols et al., 2012; van Leeuwen et al., 2001). This research examines international collaboration. It is an indicator routinely operationalised as co-authorship of articles between organisations by taking organisational address as a proxy of the collaborating country (Katz and Martin, 1997). We use geographical approximation of author heritage rooted in the morphology of the surname and find that in a significant minority of internationally collaborated papers, co-authors are likely to have the same origin. In other words, we observe an overestimation in the international collaboration indicator. The findings indicate that if a significant share of international collaborations of a national research system occurs with researchers previously affiliated with this system, internationally co-authored publications are therefore only 'inter-national' on a formal inter-organisational level. This contributes to the evidence that stresses more complex nature of scientific collaboration (Bozeman and Corley, 2004) and may have fundamental implications on the use of international collaboration indicator.

Introduction

In this research, we highlight and analyse the heritage bias in indicators for internationally collaborated research. International research performance indicators are seen to reflect relative competitiveness of a country in producing leading research (in terms of cited papers) and its commercialisation (in terms of assigned patents). An increasing number of studies point to ongoing global bias in production, composition and assessment of research metrics (Rafols et al., 2012; van Leeuwen et al., 2001). When funding, assessment and human resource decisions are made on the basis of indicators, we need to understand what exactly they do and do not tell us.

In science and innovation policy, especially when the issues of development are concerned, countries that have extensive links with other nations are appraised positively by international bodies and in national science policy. Internationalisation is a policy goal in itself, it is seen as important in advancing scientific research (Science Europe, 2014). International collaboration can also be a means to other policy goals. For instance, international collaboration in the European Research Area has been regarded, alongside with mobility, as a sign of increasing integration, and is also designed to level out research capacities of the Western and the Eastern European countries (EC, 2012).

International collaboration as an indicator is not a neutral metric and contains multiple assumptions about what is assumed to be a 'scientific collaboration'. A landmark paper by Katz and Martin (1997) alongside with later research by Bozeman and Corley (2004) bring forth these assumptions. Most notably, these authors highlight how not every co-authored publication is a result of a collaboration and not every collaboration results in co-authored publications (see further Youtie and Bozeman, 2014). As these studies expose rich underlying

social processes that are not fully captured in instances of co-publication, co-authoring of research articles, books and conference papers has come to be regarded as the dominant way in which scientific collaboration is operationalised and measured bibliometrically.

International scientific collaborations identified this way have been growing at explosive rates and is widely reported (Glänzel, 2001). The assumptions behind what the growth of this metric means have been linked mainly to research internationalisation, knowledge exchange and other benefits for the collaborating authors, organisations and countries, because the diversity increases creativity and allows to find better solutions for complex problems (Melkers and Kiopa, 2010). International collaboration is recognised as a capacity-building factor of domestic research indicating the increase in research quality (Bornmann et al., 2015). The criticisms of the growth of internationally collaborated research focus mostly around its uneven distribution, uneven contributions by different authors, and uneven benefits that the authors in central and in the peripheral regions gain from it (Schubert and Sooryamoorthy, 2009). A further example is that patterns of international cooperation in nanotechnology are still centred on the developed countries, which are key nodes in international networks (Shapira and Wang, 2010).

In this paper, we further problematise the assumptions between the supposed gains from international collaboration and the way it is measured. An accepted view of a *bibliometric definition* of international collaboration is an instance when a research output is co-authored by two or more authors affiliated with organisations located in two or more countries. This definition of international collaboration is used in all publications we reviewed for this research. Single-authored outputs with multiple affiliations are not usually regarded as instances of international collaboration. The assumption behind this measurement is, as noted above is to say that when a scientist located in Spain co-authors an article with a scientist located in Germany, it is in effect a Spanish and a German scientists collaborating with each other. Such collaboration entails positive effects related to creativity and diversity (Gkypali et al., 2017), and improves the cohesion between the German and the Spanish research systems. The bias behind these assumptions is the objective of our research.

Research Objectives

We explore the extent and the potential implications of same-heritage international scientific collaboration, taking Russian international collaboration network as a case study. Within the broader goal to unpack the types of interactions and exchanges that occur in international collaborations, we investigate the extent and the role of the bias behind the bibliometric definition's assumption that when authors located in two countries collaborate, these authors are nationals of those respective countries (see Figure 1). This issue becomes increasingly important in the case of when the developing countries aim to connect to the global scientific knowledge flows.

First, we expect to uncover the heritage bias in international collaboration, and we expect it to be substantial. That is to say, that a significant share of internationally collaborated publications between Russia and the rest of the world means co-publication between Russian scientists in Russia and Russian scientists who reside in other countries.

Second, we will investigate the structure of this bias. As shared heritage points to cognitive proximity between two scientists and makes collaboration easier to accomplish, we expect that the extent of same-heritage collaboration will vary depending on the collaborating country, discipline, it will change with time, it will vary depending on which domestic organisation is collaborating, both in terms of the type of the organisation and its geographical location, and

depending on the funding received. We also expect to see different same-heritage collaboration dynamics in the centres of excellence and among the ‘star’ scientists.

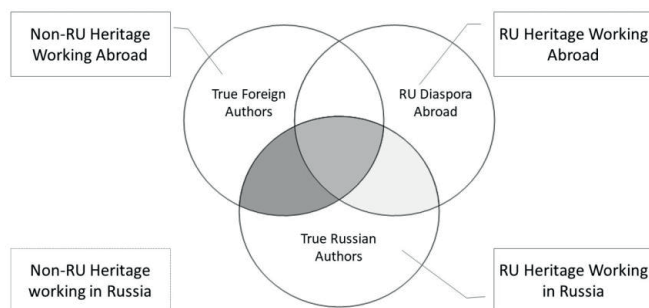


Figure 1 International Collaboration Bias Structure

By unpacking the inherent international collaboration bias, we, therefore, question the assumed relationship between co-authors in established international collaboration metrics. Ultimately, international collaboration indicators may point to reproduction and reinforcement of relationships between global centres and peripheries, and to knowledge channelling, rather to knowledge exchange.

Background

Heritage links memory, language and places with the construction of identity, values and communities (Smith, 2006). Heritage can be associated with the concept of human capital and links knowledge, skills accumulated in the duration of a person’s life with their behaviour. Studying heritage refers more to accounting for intrinsic and specific features of cultural capital within social groups, which make these groups distinctive in how they might approach problems and look for solutions.

In science studies, heritage has a twofold operationalisation. First, a researcher’s heritage is located in the organisations where she was first socialised into scientific profession. Through the process of socialisation, doctoral researchers learn about the rules, norms and expectations associated with doing scientific research (Knorr-Cetina, 1999). These norms are localised in organisations and countries. Second, heritage refers to the broad shared culture of social groups, which, most importantly, includes language. Second generation migrant researchers and entrepreneurs maintain their parents’ country of origin heritage and are able to act as mediators in the global innovation system (Khadria, 1999). Such broad understanding of heritage may also have an effect on the dynamics of international collaboration.

In the case of Russia, the country has developed a peculiar research and innovation system during the Soviet period, and even over 20 years after the USSR broke down, the research system is heavily path-dependent (Karaulova et al., 2016). At the same time, throughout the 1990s, Russia experienced high rates of ‘brain drain’, especially of the best recognised scientists in physics and mathematics (Graham and Dezhina, 2008). As of the mid-2000s onward, the Russian state set out the course towards internationalisation of domestic research, with the purpose to link the country with the global scientific community and benefit from knowledge exchange.

The set of related studies that analyse similar dynamics have so far focused mostly on the role of Chinese overseas diaspora in gatekeeping and/or mediating international collaboration (Freeman and Huang, 2015; Jin et al., 2007). This research has suggested some avenues of analysis, but has not discussed the implications of what the structural differences in international collaboration dynamics mean for how we should interpret science indicators.

Methodology

In our previously published research, we developed a method to identify the national heritage of authors based on the morphology of their surnames (Karaulova et al., 2019). By employing this approach, we infer Russian heritage from author surnames in internationally co-authored publications. Surname data has been used in bibliometric analyses to determine contribution of recognisable ethnic groups to the development of particular discipline (Kissin, 2011), to determine effects of inter-ethnic collaboration on quality of publications (Freeman and Huang, 2014), or to highlight the contribution of ethnic and gender minorities (Lewison, 2001). Taking Russia as a case study of this research has another benefit: in a country that was internationally isolated for the large part of the 20th century, geographical approximation of ‘Russian’ surnames is consistent with the actual population, i.e. most Russians still live in Russia (Revazov et al., 1986).

We take advantage of the fact that Russian surnames have persistent morphological regularities, and the majority of them can be identified from a small set of specific suffixes. Combined with first name data, this procedure has very high rates of recall and precision. By using the author name data, we can distinguish the heritage of a researcher from their work address and therefore analyse the extent to which Russian scientists in Russia collaborate with Russian scientists abroad, and what such international collaborations might mean.

We analyse publications indexed in the Web of Science that have at least one author with an address in Russia. The dataset includes 709,360 publications, the date range is 1995-2015. All disciplines, languages and document types are included in the data. Within the dataset 82.2% of publications have a co-author and 34.1% of those collaborated papers have a co-author with an address outside of Russia.

Initial Findings

Using the surname-based lexicological method, each author in the dataset is marked either as Russian heritage or non-Russian heritage. After applying the two-step Russian heritage identification procedure, we classified 95,7% of the records with the address in Russia as “Russian heritage authors”, which broadly corresponds with our previous estimates that the Russian science system does not employ many non-Russian researchers (Karaulova et al., 2019).

The findings from a small pilot study of a random sample of records published in 2015 showcase that the method outlined above can be used to address our research objectives. In the random sample, we found a significant bias in internationally collaborating research: only 18% of the internationally co-authored publications were co-authored by authors in Russia and by non-Russian authors abroad, whilst the vast majority of these papers had at least one Russian heritage author based in Russia, one Russian heritage author based abroad and one non-Russian heritage author based abroad, which possibly suggests a mediation function.

Countries that have extensive international collaboration networks and are the ‘core’ of science globalisation, such as the USA (Wagner and Leydesdorff, 2005), have lower rates of overseas diaspora involvement in the structure of collaboration networks with Russia. While

the share of research papers collaborated with the participation of Russian heritage authors reached 40% for major international partners of Russia, the results are more telling for minor partners. Countries that have relatively strong science base, but do not have traditionally close links with Russia, such as Portugal, Belgium or Australia, demonstrate very high level of overseas diaspora involvement in the share of publications co-authored with Russian scientists.

Discussion and Conclusions

If a significant share of international collaborations of a national research system occurs with researchers previously affiliated with this system, internationally co-authored publications are therefore only 'inter-national' on a formal inter-organisational level, but in fact occur between co-authors that share academic upbringing and culture. This finding contributes to the evidence that stresses more complex nature of scientific collaboration (Bozeman and Corley, 2004) and may have fundamental implications on the use of international collaboration indicator and on science policy decisions.

We found that in a significant minority of internationally collaborated papers, co-authors are likely to have the same origin. In other words, we observe an overestimation in the international collaboration indicator. This inherent bias in the established international collaboration indicator may overestimate the impact of international collaboration on periphery countries in comparison with its impact on advanced core countries. This paper makes a call for revision and further detalisation of the indicator that is sensitive to unequal science development dynamics.

When bibliometric tools are used to measure international collaboration and cooperation, invariably, assumptions are made about the social reality of these tools. Globally, the findings of this study are valid for national science policy of countries that rely on international collaboration networks to foster the development of domestic science and technology through knowledge transfer and spillovers.

References

- Bornmann, L., Wagner, C., Leydesdorff, L., 2015. BRICS countries and scientific excellence: A bibliometric analysis of most frequently cited papers. *Journal of the Association for Information Science and Technology* 66, 1507–1513.
- Bozeman, B., Corley, E., 2004. Scientists' collaboration strategies: Implications for scientific and technical human capital. *Res Policy* 33, 599–616. <https://doi.org/10.1016/j.respol.2004.01.008>
- EC, 2012. Enhancing and focusing EU international cooperation in research and innovation: A strategic approach. Brussels.
- Freeman, R.B., Huang, W., 2014. Collaborating With People Like Me: Ethnic co-authorship within the US (Working Paper No. 19905). National Bureau of Economic Research.
- Freeman, R.B., Huang, W., 2015. China's "Great Leap Forward" in Science and Engineering, in: Geuna, A. (Ed.), *Global Mobility of Research Scientists: The Economics of Who Goes Where and Why*. Academic Press, pp. 155–177.
- Gkypali, A., Filiou, D., Tsekouras, K., 2017. R&D collaborations: Is diversity enhancing innovation performance? *Technological Forecasting and Social Change* 118, 143–152. <https://doi.org/10.1016/j.techfore.2017.02.015>
- Glänzel, W., 2001. National characteristics in international scientific co-authorship relations. *Scientometrics* 51, 69–115. <https://doi.org/10.1023/A:1010512628145>
- Graham, L.R., Dezhina, I., 2008. *Science in the New Russia: Crisis, Aid, Reform*. Indiana University Press.

- Jin, B., Rousseau, R., Suttmeier, R.P., Cao, C., 2007. The role of ethnic ties in international collaboration: The overseas chinese phenomenon. *Int Soc Scientometrics & Informetrics-Issi*, Leuven.
- Karaulova, M., Gök, A., Shackleton, O., Shapira, P., 2016. Science system path-dependencies and their influences: nanotechnology research in Russia. *Scientometrics* 107, 645–670. <https://doi.org/10.1007/s11192-016-1916-3>
- Karaulova, M., Gök, A., Shapira, P., 2019. Identifying author heritage using surname data: An application for russian surnames. *Journal of the Association for Information Science and Technology* 0. <https://doi.org/10.1002/asi.24104>
- Katz, J.S., Martin, B.R., 1997. What is research collaboration? *Res. Policy* 26, 1–18. [https://doi.org/10.1016/S0048-7333\(96\)00917-1](https://doi.org/10.1016/S0048-7333(96)00917-1)
- Khadria, B., 1999. *The migration of knowledge workers: second-generation effects of India's brain drain*. Sage Publications Pvt. Ltd.
- Kissin, I., 2011. A surname-based bibliometric indicator: publications in biomedical journal. *Scientometrics* 89, 273–280. <https://doi.org/10.1007/s11192-011-0437-3>
- Knorr-Cetina, K.D., 1999. *Epistemic Cultures: How the Sciences Make Knowledge*. Harvard University Press, Boston, MA.
- Lewisson, G., 2001. The quantity and quality of female researchers: A bibliometric study of Iceland. *Scientometrics* 52, 29–43. <https://doi.org/10.1023/A:1012794810883>
- Melkers, J., Kiopa, A., 2010. The Social Capital of Global Ties in Science: The Added Value of International Collaboration. *Review of Policy Research* 27, 389–414. <https://doi.org/10.1111/j.1541-1338.2010.00448.x>
- Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P., Stirling, A., 2012. How journal rankings can suppress interdisciplinary research: A comparison between *Innovation Studies* and *Business & Management*. *Res. Policy* 41, 1262–1282. <https://doi.org/10.1016/j.respol.2012.03.015>
- Revazov, A., Paradeeva, G., Rusakova, G., 1986. Suitability of Russian surnames as a “quasi-genetic” marker. *Genetika* 22, 699–704.
- Schubert, T., Sooryamoorthy, R., 2009. Can the centre–periphery model explain patterns of international scientific collaboration among threshold and industrialised countries? The case of South Africa and Germany. *Scientometrics* 83, 181–203. <https://doi.org/10.1007/s11192-009-0074-2>
- Science Europe, 2014. *The Importance of International Collaboration for Fostering Frontier Research*.
- Shapira, P., Wang, J., 2010. Follow the money. *Nature* 468, 627–628. <https://doi.org/10.1038/468627a>
- Smith, L., 2006. *Uses of Heritage*. Routledge.
- van Leeuwen, T.N., Moed, H.F., Tijssen, R.J.W., Visser, M.S., van Raan, A.F.J., 2001. Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance. *Scientometrics* 51, 335–346. <https://doi.org/10.1023/A:1010549719484>
- Wagner, C.S., Leydesdorff, L., 2005. Network structure, self-organization, and the growth of international collaboration in science. *Res. Policy* 34, 1608–1618. <https://doi.org/10.1016/j.respol.2005.08.002>
- Youtie, J., Bozeman, B., 2014. Social dynamics of research collaboration: norms, practices, and ethical issues in determining co-authorship rights. *Scientometrics* 101, 953–962. <https://doi.org/10.1007/s11192-014-1391-7>