Science and Public Policy Advance Access published March 17, 2015

Science and Public Policy (2015) pp. 1–22 doi:10.1093/scipol/scv007

On the relationship between gender disparities in scholarly communication and country-level development indicators

Cassidy R. Sugimoto[1,*], Chaoqun Ni[2] and Vincent Larivière[3]

[1]*School of Informatics and Computing, Indiana University Bloomington, 1320 E. 10th St, Bloomington, IN 47405, USA.*
[2]*Graduate School of Library and Information Science, Simmons College, 300 The Fenway, Boston, MA 02115, USA.*
[3]*École de bibliothéconomie et des sciences de l'information, Université de Montréal, C.P. 6128, Succ. Centre-ville, Montréal, QC, H3C 3J7, Canada and also at Observatoire des Sciences et des Technologies, Centre Interuniversitaire de Recherche sur la Science et la Technologie, Université du Québec à Montréal, CP 8888, Succ. Centre-Ville, Montréal, QC, H3C 3P8, Canada.*
**Corresponding author. Email: sugimoto@indiana.edu.*

Gender disparities in science remain, despite decades of policies aimed at achieving gender parity. Yet, little is known about the macro-level factors affecting such disparities. This paper examines the degree to which country-level human development indicators (HDI) and gender inequality indicators (GII) gathered by the United Nations Development Report can reveal systemic gender inequalities in scholarship. Countries 'low' in HDI and GII had the lowest contribution of female participation in science and highest degree of international collaboration. Research from highly developed countries was more cited, although gender disparities remained. For HDI, gross national income was a strong predictor of scientific output and impact (and, to a lesser degree, collaboration). The rate of women in the labor force was the strongest predictive variable in GII, explaining differences in output, collaboration, and impact. However, predictive variables differed by HDI/GII quartile, suggesting that monolithic policies may not be appropriate for addressing gender disparities in science.

Keywords: gender disparity; science policy; scientometrics; human development; scholarly communication; citation analysis.

1 Introduction

For a country to be scientifically competitive, it needs to maximize its human intellectual capital–base. However, gender disparities in science remain, despite more than a decade of policies aimed at achieving gender parity (Hagmann 1999) and promises of a 'coming gender revolution in science' (Etzkowitz et al. 2008). It is clear from previous studies (Larivière et al. 2013) that more needs to be done in terms of capital formation. The matriculation of students in tertiary education is not the problem: the OECD (2012) reports more female entrants than men in all but four countries (Germany, Turkey, Korea, and Japan). Rather, disparities are seen further downstream: women drop out of academe at a higher rate than men, creating more opportunities for men to progress through the ranks. As an editorial in *Nature* noted, in the USA and Europe, 'barely one-fifth of full professors are women' and progress on gender equality in science 'seems to have stalled' (Anonymous 2013).

Some might suggest that it is not enough to focus policies on tertiary education, as the problem may well be grounded in cultural values embedded in primary and secondary education. However, girls score significantly better than boys in primary and secondary education in countries such as Oman, Qatar, UAE, Georgia, and Saudi Arabia (Provasnik et al. 2012)—countries where men dominate scientific production (Larivière et al. 2013). Disciplinary differences in gender composition may also be dictated at the country level: Iran has introduced a stipulation excluding women from studying in the areas of nuclear physics, electrical engineering, and mining engineering (International Labour Organization 2013), thus perpetuating gender differences in these countries and disciplines (Larivière et al. 2013). This is exacerbated by differences in gendered propensities to engage in international collaboration: for example, less than 15% of Iranian women's papers are the result of international collaboration; while, in contrast, 44% of female-authored papers in France are the result of international collaboration. These stark differences across the globe suggest that that country effects moderate gendered participation in science.

In a comprehensive and contemporary study of gender disparities in science, Larivière et al. (2013) found that women were underrepresented in scholarly output, had more domestic collaboration portfolios than men, and received fewer citations than male colleagues. They concluded by noting the complexity of inferring policy actions from macro-level scientometric data in isolation. They state:

> . . . if there were a simple solution or programme that could improve matters, this issue would already be solved. Unfortunately, behind this global imbalance lie local and historical forces that subtly contribute to the systemic inequalities that hinder women's access to and progress in science. Any realistic policy to enhance women's participation in the scientific workforce must take into account the variety of social, cultural, economic and political contexts in which students learn science and scientific work is performed. Each country should carefully identify the micro-mechanisms that contribute to reproducing the past order. No country can afford to neglect the intellectual contributions of half its population. (Larivière et al. 2013)

The objective of this paper is to examine the degree to which country-level human development (HDI) and gender inequality indicators (GII) can reveal potential 'systemic inequalities' related to the contribution of women to scholarly communication. This work will examine relationships between country-level indicators gathered by the United Nations (UN) Development Report (Human Development Report 2013a; 2013b) and scientometric measures of output, collaboration, and impact. Specifically, the results will use the variables and ranking of the HDI and GII.

The results of this study will inform policy development aimed at reducing gender disparities in science. In particular, this study will investigate whether there are systematic factors relating to disparities across countries or whether particularistic factors at the country-level dominate, which factors are most explanatory, and how the relative development of the country factors into this analysis. As Long and Fox (1995) noted in their review of gender disparities in science:

> It is not enough to ask whether particularism or universalism operates in science. . .Rather, given the patterns observed, it is important to understand more fully the processes leading to the lower participation and performance of women and minorities in science.

It is the goal of this paper to lend some evidence to this call.

2. Literature review and research questions

The production of journal articles is but one place where gender disparities in science can be seen: male students are given preference for positions as lab managers (Moss-Racusin et al. 2012); female academics earn less than their male counterparts (European Commission 2007; Shen 2013) and are not as satisfied (Holden 2001; Seifert and Umbach 2008); women patent and engage less in other entrepreneurial activities (Abreu and Grinevich 2013; Ding 2006; Ding and Choi 2011; Hauessler and Colyvas 2011; Hunt et al. 2013; Ostergaard et al. 2011) and are not proportionately represented on scientific advisory boards (McCook 2013); and women fare poorer in terms of funding allocations (Larivière et al. 2011; Ley and Hamilton 2008). These items are fairly interrelated: high degrees of collaboration can increase productivity, productivity is likely to garner additional citations, and collaborative work is more cited than sole-authored work (Larivière et al. 2014). Each of these contributes to the social and scientific 'capital' (Bozeman et al. 2001; Pezzoni et al. 2012) of the scholar, which in turn generates additional rewards such as: funding, opportunities for technology transfer, and selection for prestigious positions. Therefore, this work will focus on these foundational indicators of scientific 'capital' (i.e., productivity, collaboration, and impact) and the relationship between these items, gender, and country-level development indicators. Each research question and associated hypothesis will be introduced in the context of the motivating literature. However, a full table of research questions, hypotheses, and the degree to which they were supported by the results can be found in the Appendix (Appendix 5).

2.1 Research output and productivity

The so-called 'productivity puzzle' has been the focus of much research on gender and science, demonstrating that men, on average, produce more than women in studies that have, respectively, spanned time periods, countries, and fields (Council of Canadian Academies 2012; Fox 2005; Larivière et al. 2011; Prpic 2002; West et al. 2013; Xie and Shauman 1998, 2003; Zuckerman 1991; Paul-Hus et al. in press), although variations by fields/subfields can be found (for a review, see van Arensbergen et al. 2012). Studies have also looked at the degree to which men and women contribute overall—that is, calculating the ratio of female and male output, rather than the individual productivity. Larivière et al. (2013) is the largest of such studies to-date and demonstrates that men have a higher research output than women across most countries and disciplines. However, no studies have examined the relationship between output and levels of development and equity. Therefore, using HDI as an indicator of levels of development and GII as an indicator of gender equity in a country, we will assess the following questions:

RQ1a. What is the relationship between gender parity in scientific output and levels of human development?
H1a. We would expect that levels of gender parity in scientific output and levels of human development would be related; that is, that gender parity would increase in countries with higher levels of human development.

RQ1b. What is the relationship between gender parity in scientific output and levels of gender equity?
H1b. We would expect that levels of gender parity in scientific output and levels of gender equity would be related; that is, that gender parity in scientific output would increase in countries with a higher level of gender equity.

RQ2a. What is the relationship between a country's capacity to produce (i.e., population) and overall production, gender parity in scientific output, and levels of human development?
H2a. We would expect that the most productive countries (normalized by population size) would tend to be those in highly developed countries and that these countries would have higher levels of gender parity in scientific output.

RQ2b. What is the relationship between a country's capacity to produce (i.e., population) and overall production, gender parity in scientific production, and levels of gender equity?
H2b. We would expect that the most productive countries (normalized by population size) would tend to be those in countries with high degrees of gender equity and that these countries would also be those trending towards gender parity in scientific output.

2.2 Collaboration

Collaboration rates continue to grow across disciplines and countries (Ganzi et al. 2012). However, women's participation in such activities is not equal to that of men. Men predominate in the prestigious first and last author positions and women are significantly underrepresented as single authors (West et al. 2013). Women are less likely to be involved in collaborative research projects and less likely to participate in collaborations that lead to publication (Long 1992; Mauléon and Bordons 2006; Sonnert and Holton 1995; Xie and Shauman 1998 citing Cole and Zuckerman 1984; Zuckerman 1991). Female researchers with children were often less likely to collaborate (Long 1992) with mentors expressing less likelihood of collaborating with junior scholars with children (Ezkowitz et al. 2000). In a study of Quebec scholars, Larivière et al. (2011) demonstrated that women were more likely to collaborate nationally than internationally, resulting in a more restricted collaboration network. This result was replicated in a large-scale global study (Larivière et al. 2013), finding that women in all the most productive countries had more domestic portfolios than their male compatriots. Given this, we will investigate how the relative difference between men and women in international and domestic collaboration varies by levels of development and gender equity in a country.

RQ3a. What is the relationship between gendered participation in collaboration and levels of human development?
H3a. We expect that this difference will be smaller in highly developed countries and more pronounced in lesser developed countries. Furthermore, we expect higher levels of international collaboration in highly developed countries.

RQ3b. What is the relationship between gendered participation in collaboration and levels of gender equity?
H3a. We expect that women's portfolios will be more domestic than their male compatriots (as shown in Larivière et al. 2013); however, we expect that this difference will be more pronounced in countries with low gender equity (i.e., low GII rankings).

2.3 Scientific impact

There has been considerable disagreement regarding the relative impact (as demonstrated through citations) of women's work compared to men's (van Arensbergen et al. 2012). Studies have shown similar impact between men and women (Bordons et al. 2003; Gonzalez-Brambila and Veloso 2007; Long and Fox 1995; Mauleón and Bordons 2006; Zuckerman 1991 citing Cole and Zuckerman 1984), women with higher impact (Long 1992) and men with higher impact (Bordons et al. 2003; Larivière et al. 2011; Mauleón and Bordons 2006; Peñas and Willett 2006). However, these studies have been highly localized and those investigating more than one discipline have shown wide disciplinary variation. However, in an analysis of 5.4 million research articles, Larivière et al. (2013) found that articles sole-authored by women had a lower impact than men's, as did articles with women in dominant author positions (i.e., first and last). Based upon this work, we will examine the relationship between the gender gap in impact and levels of development and gender equity.

RQ4a. What is the relationship between gender, collaboration, and impact, by levels of human development?
H4a. Previous research (Larivière et al. 2013) has shown that articles with women in dominant author positions are less cited than those with men in these positions. We expect that this difference will be smaller in more developed countries. We also

expect to see higher levels of overall impact for women coming from more developed countries.

RQ4b. What is the relationship between gender, collaboration, and impact, by levels of gender equity in a country?
H4b. We expect to see higher impact for articles with women in dominant author positions in countries with high levels of gender equity and smaller differences between dominant male/female authors in countries with higher gender equity.

2.4 Explanations and solutions

Decades of research have been devoted to finding explanations of gender inequality in science (Cole and Zuckerman 1984). Bias is often suggested as an explanation, but others (Ceci and Williams 2011) prefer to focus on the 'differing biological realities'—that is, the fact that women bear children. A validation of this explanation was found by a study noting a connection between childcare systems in the country and the share of women's scientific output (Frietsch et al. 2009). Another explanation for gender differences is research specialization. Leahey (2006), for instance, found that women were working on a more diverse set of topics than men, and that this lower specialization has a negative effect on their visibility and, in turn, on their academic capital.

Some have suggested that quotas are the answer to pervasive gender inequalities. To this end, the European Commission has committed to upholding a quota to ensure that at least 40% of the members of its funding advisory board for 2014–20 are women (Muhlenbruch and Jochimsen 2013). However, it is feared that this may place additional burdens on already overworked women in science and may fail to produce demonstrable outcomes (Mühlenbruch and Jochimsen 2013) as both men and women are biased in their evaluation of females applying for positions in science (Moss-Racusin et al. 2012). It has also been suggested that the lack of female role models and mentors serve as detriments to female progression and success in science (Wellcome Trust 2013).

Despite the long history of research in this area, there have been relatively few large-scale global studies of gender and scientific productivity, collaboration, and impact. West et al. (2013) and Larivière et al. (2013) are two exceptions. However, each of these only provided macro-level scientometric data and did not examine potential explanations for these findings or the relationship between these findings and additional information about the discipline and countries studied. Therefore, the goal of this paper is to examine the relationship between these macro-level scientometric measures and country-level human development and GII, in order to identify potential systemic issues contributing to disparities in scholarly communication. Specifically, we hope to address the following research questions and hypotheses:

RQ5a. What is the relationship between gender parity in scientific output, collaboration, and impact and each HDI indicator, by levels of development?

H5a. We would anticipate that those indicators most related to education would be the most related to gender parity in science. We would expect that these issues would be constant across levels of development.

RQ5b. What is the relationship between gender parity in scientific output, collaboration, and impact and each GII indicator, by indicator rank?

H5b. We would anticipate that those indicators most related to female participation in the labor market would be most related to the gender parity in scientific output. We would expect that these issues would be constant across levels of gender equity.

3. Material and methods

3.1 Bibliometric data

Data for this project are drawn from Thomson Reuters' Web of Science (WoS) database, covering the Science Citation Index Expanded, the Social Sciences Citation Index and the Arts and Humanities Citation Index. All articles (research and review articles) for the period 2008–12 were included in the analysis. Other types of documents (e.g., editorials, letters to the editor, and book reviews) are excluded from the analysis because they are generally not peer-reviewed, nor considered to be original contributions to scholarly knowledge (Moed 1996). In total, 5,423,287 articles (research and review articles) with 27,037,168 authorships (i.e., counting each instance of an author on a byline as a single authorship) were analyzed. The raw data were transformed into a relational database on an SQL server in order to perform the various analyses.

3.2 Gender and development indices

Two types of indices, both constructed by the UN Development Programme (UNDP), were used in this analysis. The Human Development Index:

> . . . is a summary composite index that measures a country's average achievements in three basic aspects of human development: health, knowledge, and income. (UN Development Programme 2013)

It was introduced into the Human Development Report in 1990 (UN Development Programme 2013). As shown in Fig. 1, the HDI has three dimensions—health, education, and living standards—comprised of four indicators: life expectancy at birth, mean school years, expected schooling years, and gross national income (GNI) per capita.

The GII:

> . . . reflects women's disadvantage in three dimensions—reproductive health, empowerment and the labour market. (Human Development Report 2013)

These dimensions include seven indicators: female/male labor force participation, female/male educational attainment (secondary level and above), parliamentary representation (of women), adolescent fertility, and maternal mortality (see Fig. 1).

Data were downloaded from the Human Development Report (2013) (that is, data representing the state of the country in 2012). These data are freely available at <http://hdr.undp.org/opendata> and contain detailed explanations of the calculation of each indicator. The HDI is explicitly grouped into quartiles within the report and country association with one of the quartiles (very high, high, medium, and low) was used for the analysis. GII values were not identified in quartiles. Therefore, the values were ranked and grouped into quartiles manually to match the HDI analysis. Although our article data is from the period 2008–12, we chose the 2012 HDI data as the most recent and most comprehensive dataset: 18 countries were added to the HDI in the period 2011–2. In addition, the index tends to be fairly stable. For example, within the most productive (i.e., those which produced the highest number of papers in our dataset) 50 countries, no country changed quartiles within the HDI distribution in the period 2011–2. In total, 187 countries were included in the 2012 HDI.

The GII was first used in 2010 to replace the gender-related development index and the gender empowerment measure. Fewer countries produce reliable statistics on these variables than the HDI: 148 countries are included in the GII. The GII is similar to the HDI in the stability of quartiles: only two of the 50 most productive countries (accounting for 97.2% of all papers) in the GII changed quartile in the period 2011–2 (Turkey and Mexico; from medium to high). Therefore, 2012 was also used for the GII.

Of the 148 countries with both HDI and GII data available in 2012, 42 countries fall into different quartiles between the indices. For example, China is ranked medium in HDI, but very high in GII; the USA is considered very high according to the HDI, but high by GII. The Spearman correlation between HDI and GII is 0.876 (significant at the 0.01 level). However, despite this high correlation, results from both will be reported, given the slight difference in classification of high producing countries such as China and the USA. Furthermore, the indicators are composed of different variables which will be used in the regression analyses.

3.3 Matching countries with WoS

Thomson Reuters indexes the institutional address (institution, country, city, etc.) of each author, which allows for precise geographical assignation of articles by gender. A list of 203 countries/territories was extracted from WoS based on the author address information provided by each publication. The name of countries (in English) provided by the WoS database was used. Names as provided by International Organization for Standardization (ISO) Standard 3166 were used instead of WoS names. For instance, the Democratic Republic of the Congo is Zaire in the WoS database, which officially refers to the state that existed in the period 1971–97. South Korea is the name from WoS, while it is the Republic of Korea according to the ISO Standard 3122.

Of the 187 countries reported in the HDI, 183 had publications in WoS during the period 2008–12. Four entries appeared in HDI, but were not in WoS: Hong Kong, Occupied Palestinian Territory, South Sudan, and Timor-Leste. Hong Kong was reported separately from China in the UNDP as a Special Administrative Region. However, only one paper was associated with Hong Kong (as a separate country) in this time period in WoS; therefore, this paper was merged with China. Taiwan was not reported by the UNDP, but was listed as an individual country in WoS. Therefore, it was not included in the analysis for either GII or HDI.

Additionally, we excluded countries which had less than 20 papers and fewer than 50% of those 20 articles in the country had at least one gender assigned. These included: St. Vincent and the Grenadines, Samoa, and Andorra. After these exclusions, 180 countries were included in the HDI analysis, representing 98.37% of the total papers in the original dataset. Countries were not recategorized into quartiles, but kept in the original quartiles. Given this, there were 45 countries in very high, 46 in high, 43 in medium, and 46 in low.

As noted earlier, 148 countries were included in the original GII report, all of which have corresponding HDI values. All these 148 countries were included in the analysis regarding GII, accounting for 97.8% of the papers in the original dataset. There were 37 countries in each quartile.

3.4 Name–gender assignment

Since 2008, WoS has included the full first name of authors (when appearing on the byline of scientific papers), which allows the gender classification of authors. An author–name list was prepared that contains the given names of authors indexed in WoS. The given name was provided in a separate field, but not in a unified form. Some given names are initials instead of complete names, or contain special characters like '()', '-', '.' or a space. In order to match with the source lists introduced above, the author-name set was preprocessed as follows:

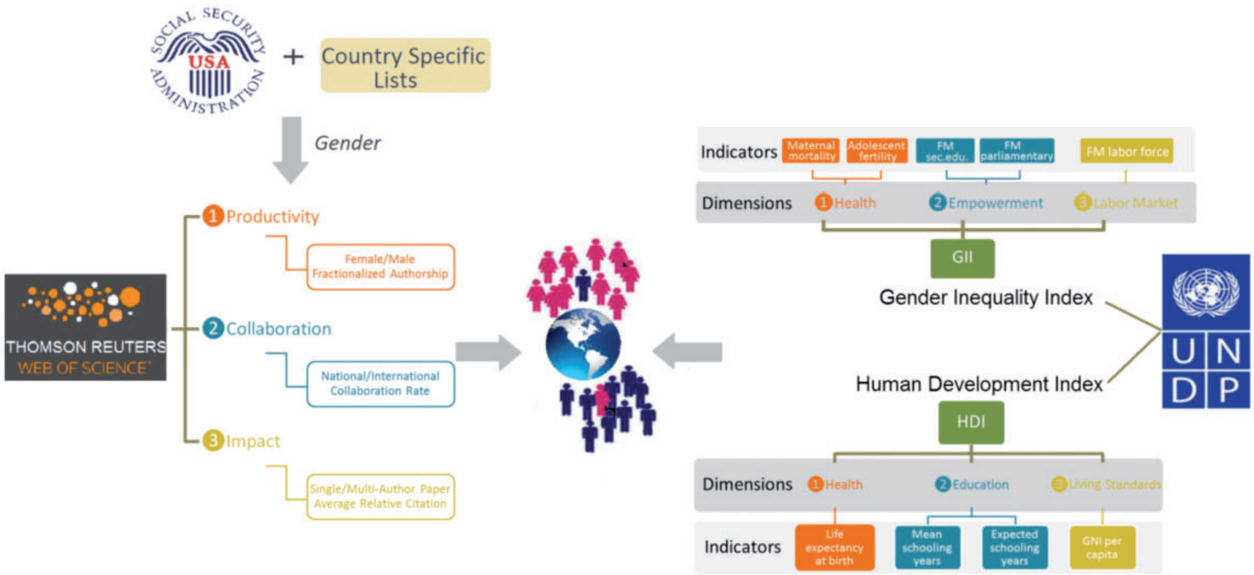


Figure 1. Schematic of data integration.

- All characters in '()' of a given name were extracted and treated as nicknames;
  - Identify initials:
    - Calculate the '.' in the given name:
      - If no '.', calculate the space
      - If there is '.', calculate the length of whole string
        - If the length of a string is smaller than 3 times the number of '.', then they are treated as initials.
        - If not: leave for next step
- For names that are not initials, split given name to several parts by space;
- Replace all hyphens in each part into a space: for instance, 'Jean-Pierre' will be converted to 'Jean Pierre'.

It should be noted that we identified authorships, not individuals—we were interested in identifying the gender of each author, but were not concerned with matching authors across papers. That is, we were interested in the gender of each author on each given paper, but not on how many papers were authored by that individual author. Our analysis is on the aggregate level: how many papers had a female or male author, not only how many papers were authored by each individual female or male author.

Perceived gender was identified by matching the first names of authors with universal and country-specific name lists. Universal lists were applied to the entire set of WoS authors, and country-specific lists were applied to subsets of WoS authors associated with particular countries. The dominant list was based upon US Census data. The US Census provides lists of given names and the percentage of the population with a specific given name and associated gender. Therefore, with the given names of authors obtained from WoS data, each author was coded for possible gender using these lists. In cases where a name was used for both genders, it was only attributed to a specific gender when it was used at least ten times more frequently for one gender than the other. Otherwise it was categorized as a 'unisex' name. Only names classed as male or female were used in the final analysis.

Other universal lists (e.g., WikiName, WikiPedia) were only used for names that could not be categorized using the US Census list (as male, female, or unisex). We incorporated 22 other lists to provide country-specific name information (see Lariviere et al. (2013) for links to all lists). For many countries, rules were instituted in addition to relying on name–gender lists. For example, in Korea, names ending with -jae are typically male names, while names with -mi are typically associated with female names. First names were the primary unit of analysis. However, in some cases surnames were used to inform assignment where surnames carry gender information (e.g., Russia and former Soviet countries[1]). Lastly, manual analysis was applied to countries where translation caused difficulties in name assignment. Specifically, a subset of frequently occurring Chinese and Taiwanese names was manually disambiguated by native speakers. Details regarding these lists, rules, and processes can be found in the supplemental material of Larivere et al. (2013).

Matching between the author–name and gender–name lists resulted in the assignment of a female or male gender to 56.1% of distinct given names (e.g., Thomas, Elizabeth) and 59.5% of distinct full author names (e.g., Thomas Whitten, Elizabeth Barry). A significant portion of author names only provided initials (31.0%). Therefore, in terms of the percentage of authors that provided given name beyond initial(s), gender was assigned to 57.3% of distinct names and 83.0% of distinct full names. At the level of distinct papers and paper–authors (e.g., the sum of each author appearing on the byline of articles), the results are similar: 81.3% of papers had gender assigned to at least one of the authors and 65.2% of the author–paper combinations had a gender assigned. When authors with only initials are excluded, the percentage of author–paper combinations with gender assignment increases to 86.1%. (For full tables detailing these results, see Lariviere et al. (2013).)

To assess the accuracy of our analysis, we selected 1,000 records at random representing an individual author who had been categorized into each of the following five exclusive categories: initials, unknown, unisex, male, and female (for a total of 5,000 distinct records). These authors were associated with a specific country, institution, and, in some cases, an email address. This information was used to locate biographical information with pronouns or a photograph on the web that could be used to verify the accuracy of the categorization. The percentage of the random sample that could be gender identified varied by category (see Table 1) and was dependent on many variables, including the status of the author. For example, in the male category, many of the authors were technicians and staff members who lacked lengthy biographical information (which would contain pronouns) and photographs.

In previous studies, it was shown that women accounted for fewer than 30% of fractionalized authorships, whereas men represent slightly more than 70% (Lariviere et al. 2013). It has been suggested that this may underrepresent women as they may be disproportionately likely to use initials or have names that are unclassed (unknown or unisex). However, as shown in Table 1, we were able to identity nearly 83.9% of those in our sample who were using initials. Of those identified, 23.6% were identified as female and 76.4% were identified as male. This disproves any assumption that women may disproportionately rely on initials. Unknown names were 31.7% female and unisex were 20.9% female. Given this, we can assume that women are not underrepresented in our data. If anything, women may be slightly overrepresented: our analysis of names classed as female found that 13.3% of these names were incorrectly classed as female when they should have been classed as male. Only 1.7% of names were incorrectly classed as male when they should have been female.

3.5 Analysis

Indicators presented in this paper are based on the number of articles and review articles published by authors of each gender. For output, these numbers are based on fractional counting of papers: that is, each author is given $1/x$ count of the authorship where x represents the number of authors for which a gender could be assigned on the given paper. For each country, the proportion of the research output by female or male authors is aggregated

Table 1. Percentage male and female in each category

Category	# and % identified	# and % female (of identified)	# and %–– male (of identified)
Initials	839 (83.9%)	198 (23.6%)	641 (76.4%)
Unknown	890* (89.0%)	282 (31.7%)	607 (68.2%)
Unisex	540 (54.0%)	113 (20.9%)	427 (79.1%)
Male	605 (60.5%)	10 (1.7%)	595 (98.3%)
Female	830 (83.0%)	720 (86.7%)	110 (13.3%)

*Number here is not sum of male and female because one author self-identified as 'other'. They are, therefore, neither male, female, or unidentified

at country level. For a country with k papers associated with female authors and n papers associated with male authors, the female-to-male research output ratio $r_{F,M}$ in each country can be calculated as:

$$r_{F,M} = \frac{p_F}{p_M}$$

$$p_F = \sum_i^k \frac{a_F}{a_T}$$

$$p_M = \sum_i^n \frac{a_M}{a_T}$$

where p_F denotes the fractional counts of publications by female authors, and p_M denotes the fractional counts of publications by male authors, a_F is the number of male authors in publication i, a_T is the total number of authors in publication i, and there are k publications produced by this country. The gender disparity for each country can then be calculated as:

$$d_p = |1 - r_{F,M}|$$

It should be noted that, in the regression analysis, $1 - r_{F,M}$ is utilized on the x-axis in order to denote directionality for each country.

To gain a greater understanding of the relationship between female-to-male output in each country and overall contribution and potential to contribute, we normalized output by the country's population. Fig. 3 shows the relative distance of each country's output (normalized by population) to the median of the world output (y-axis) by the country's closeness to gender equality (i.e., d_p). A positive x value indicates that males out-produce females, negative value indicates females out-produce males in that country. The distribution of y is displayed in a boxplot by HDI/GII level on the right-hand-side of Fig. 3. It was calculated as:

$$y = \begin{cases} \sqrt{p - m_p}, \text{ when } p - m_p \geq 0 \\ (-1)\sqrt{|p - m_p|}, \text{ when } p - m_p < 0 \end{cases}$$

where p is the number of papers in each country normalized by its population size, and m_p is the median value p in all countries in the world (i.e., m_p = median($_P$)), which results in a median value of p: 247.7 by HDI, and 285.3 by GII. Countries with $p - m_p < 0$ were therefore displayed below zero. The x-axis is shows the closeness to gender equality which was calculated as:

$$x = 1 - r_{F,M}$$

The distribution of x is shown in the boxplot by HDI or GII level at the bottom of Fig 3.

The collaboration in each country is examined from two different perspectives: national and international. For each country, the international and national collaboration rate is calculated for females and males. The female international collaboration rate (FIR) is the percentage of internationally collaborative publications with female authors out of the total number of publications involving female authors in each country. The female national collaboration rate (FNR) is the percentage of domestically collaborative publications involving female authors out of the total number of publications involving female authors in each country. For instance, for Turkey, there are a total 32,018 publications with at least one female author, 3,512 of them were authored by female authors collaborating with other authors from outside of the country, and 30,516 of them were authored by female authors collaborating with others from within the country. Therefore, for Turkey: FIR = 3,512/32,018 = 10.97%, and FNR = 28,503/32,018 = 89.03%. Similar approaches were used in calculating the male international collaboration rate (MIR) and the male national collaboration rate (MNR).

This paper uses the difference between the FIRs and MIRs as one indication of gender disparity in terms of collaboration. For a specific country, the disparity in collaboration is calculated as:

$$d_c = |FIR - MIR|$$

The average of relative citations (ARC) is used as a proxy of impact in this paper. The ARC for each article is calculated as the article's number of citations divided by the average number of citations received by articles in the same discipline published that year, in order to reduce possible bias among disciplines. When the ARC is above one, a given article is cited above the world average for the same field. Conversely, an ARC below one means that the number of citations received is below the world average.To measure the disparities in citations between females and males, this paper created a disparity indicator in citation, d_i, based on the difference in citations among five types of papers:

$$d_i = \sqrt[2]{\begin{array}{l}(sf - sm)^2 + (nff - nfm)^2 + (nlf - nlm)^2 \\ +(iff - ifm)^2 + (ilf - ilm)^2\end{array}}$$

where sf is the citation for single-female-authored papers, and sm is the citation for single-male-authored papers; nff is the citation for nationally collaborated papers with female first author, and nfm the citation for nationally collaborated papers with a male first author; nlf is the citation for nationally collaborated papers with a female last author, and nlm the citation for nationally collaborated paper with a male last author; iff is the citations for internationally collaborated papers with a female first author, and ifm is the citation for internationally collaborated papers with a male first author; ilf is the citation for internationally collaborated papers with female last author, and ilm is the citation for internationally collaborated papers with a male last author; the d_i shows the distance in citations among females and males in these five-dimensional spaces.

The well-known limitations of bibliometrics apply to this analysis, as the WoS does not index all of the world's scholarly literature. This is more problematic for the social sciences and the humanities, where there is virtually no coverage of research output in media other than journal articles, and there is very limited coverage of research output in the form of articles written in languages other than English (Archambault et al. 2006; Larivière et al. 2006). Hence, the results obtained underestimate the results obtained in the social sciences and humanities, especially for countries for which English is not the main language.

Multiple linear regression analysis (using the ordinary least square method) was employed to explore the relationship between gender disparities in output (d_p), collaboration (d_c), and impact (d_i) with multiple indicators of HDI and GII. Indicators of HDI and GII are significantly correlated. To avoid multicollinearity, this paper models the relationship between the gender disparities of these three aspects and indicators of HDI and GII separately.

4. Results

4.1 Research output and productivity

The relationship among output, as measured through fractional authorships, gender, and human development was measured by examining the female-to-male output by quartile. As shown in Fig. 2, there is very little difference between the median ratios for countries ranked as very high, high, and medium countries according to the HDI, with high countries coming closest to gender parity. The female contribution in each of these countries is typically around 30%, matching the global average (Larivière et al.

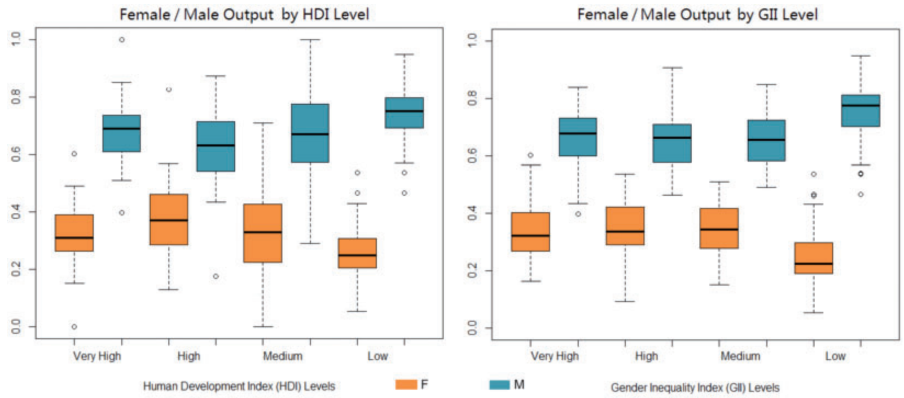


Figure 2. Female/male output by HDI and GII level (orange/lighter represents female output, blue/darker represents male output).

2013). However, the range of difference increases from very high to high and from high to medium (with medium having the greatest range of all categories). The largest discrepancy appears in the lowest developed countries, where the typical female contribution is closer to 20%. The stability across very high, high, and medium is even more pronounced in GII, with very little difference among these median rates. Countries in the low quartile have much lower female participation.

Fig. 3 presents the relationship between countries' productivity (papers/inhabitant) and gender parity. As can be seen, very high and low countries tended to be furthest from gender equality, with high and medium-developed countries trending closer to equality. Fig. 3 also demonstrates that highly developed countries tend to be more productive, with a few exceptions. Countries with a very high HDI are the next most productive, with a median significantly above that of medium and low countries. There was very little difference in typical productivity between medium and low countries, although the range for medium countries was much larger. Instead of plotting countries by closeness to gender equality in productivity (d_p introduced in Section 3), Fig. 3 locates countries considering the direction of gender inequality in each country. Countries with women out-producing men are below zero, and countries with men out-producing women are above zero on the x-axis. The plotting of countries by their distance to overall world productivity on the y-axis is similar.

Fig. 4 represents similar data, but calculated by GII level, rather than HDI level. A similar distinction was found in terms of GII, with productive countries being those with very high HDI and countries with high and medium GII levels trending closer to equality in production.

4.2 Collaboration

Larivière et al. (2013) provided evidence that women collaborate more domestically and less internationally than men in the same country. Fig. 5 reinforces this, but shows differences by HDI and GII ranking. It demonstrates, for the ten most productive countries in each quartile, the proportion of female/male collaborations that are international vs. national. For example, it shows that 22% of female collaboration in the USA is international, while 78% is the result of national collaboration. For each country, male (blue/darker) and female (orange/lighter) average international/national collaboration rates are presented. If female rates of collaboration exceed male rates in international/national collaboration, then the two dots are connected with a line colored in orange (lighter color). Otherwise, the two dots will be connected via a blue line (darker). The trailing orange/lighter dots on the left of each image show the domestic tendencies of female collaboration. However, Fig. 5 also demonstrates a shift towards higher degrees of international collaboration as the degree of development and gender equality decreases in the country. This suggests that countries with low levels of development seek international collaboration at a greater rate than highly developed countries, but that levels of development do not change the ratio of domestic/international collaboration between men and women in a single country. Similar results are found between HDI and GII country rankings.

4.3 Scientific impact

The ten most productive countries in each category were also investigated for relative impact, as measured through field-normalized citations, for men and women in these countries, according to authorship position. As shown in Figs 6 and 7, women receive fewer citations as single authors and in dominant author positions (i.e., first and last authors) irrespective of the type of collaboration (international/national), and there is a citation advantage to collaborative work and an added advantage to international collaborations. This supports the work of previous studies (Larivière et al. 2013), and provides evidence that these macro-level trends are observed in almost all countries. Furthermore, the citation advantage of 'very high' countries versus all other countries is highly visible

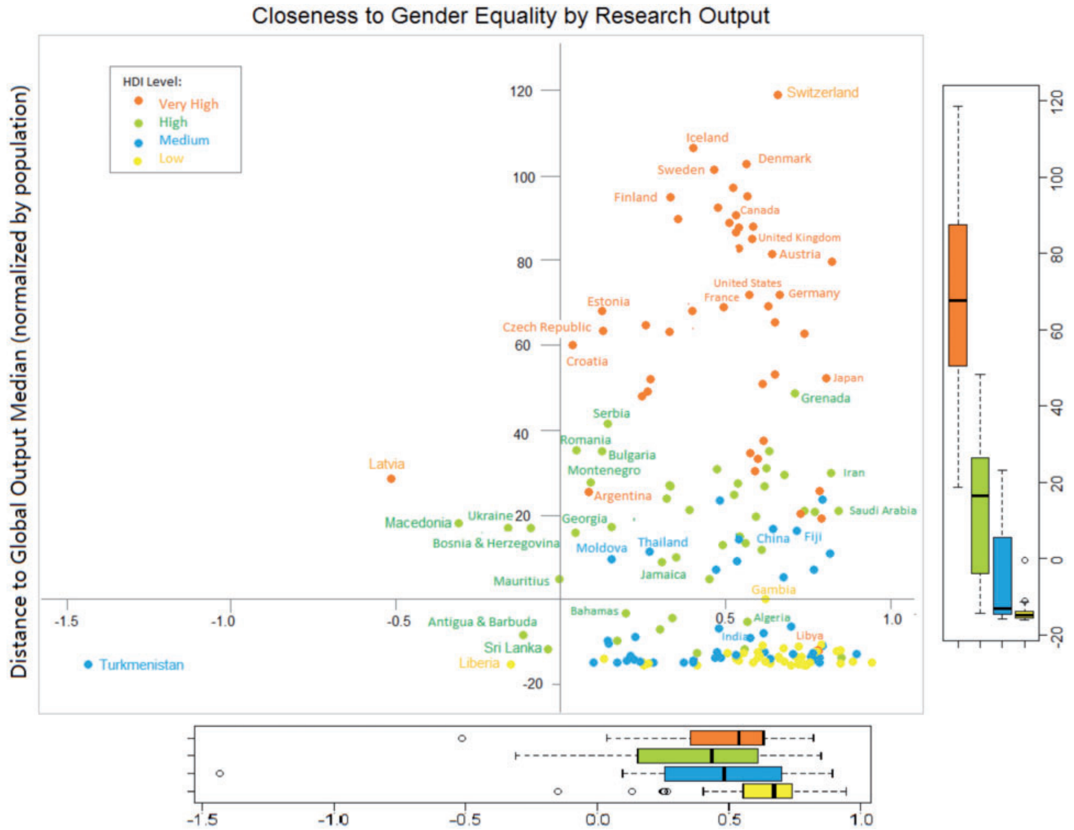


Figure 3. Productivity and gender equality in research output by HDI level.

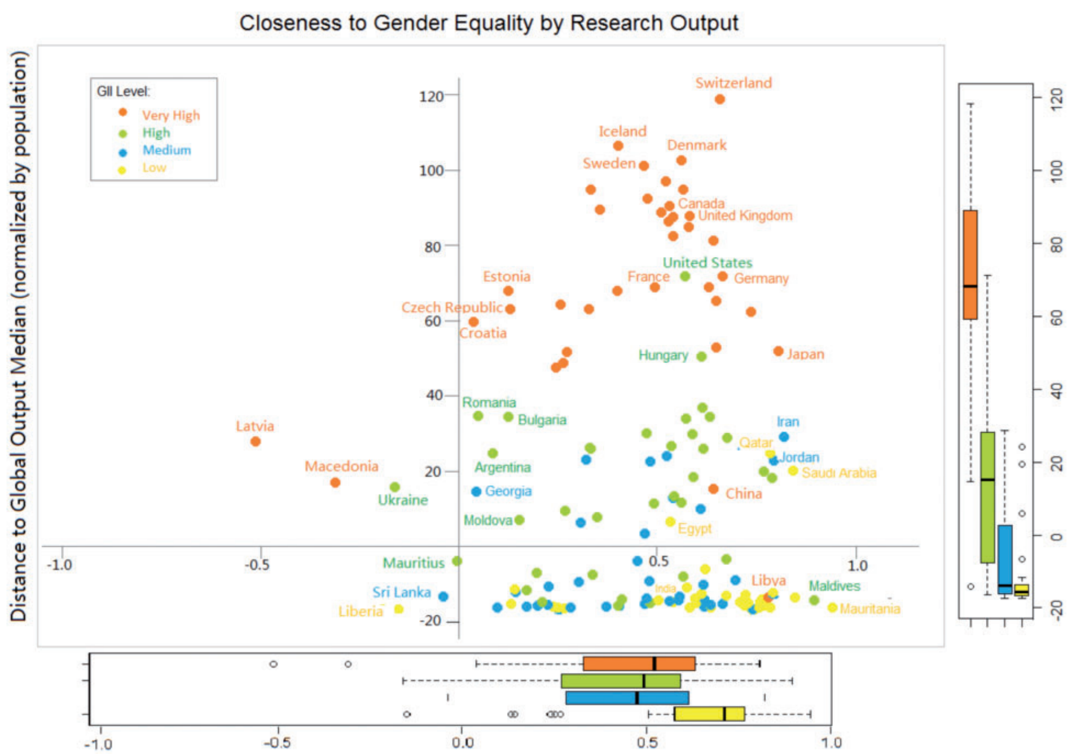


Figure 4. Productivity and gender equality in research output by GII level.

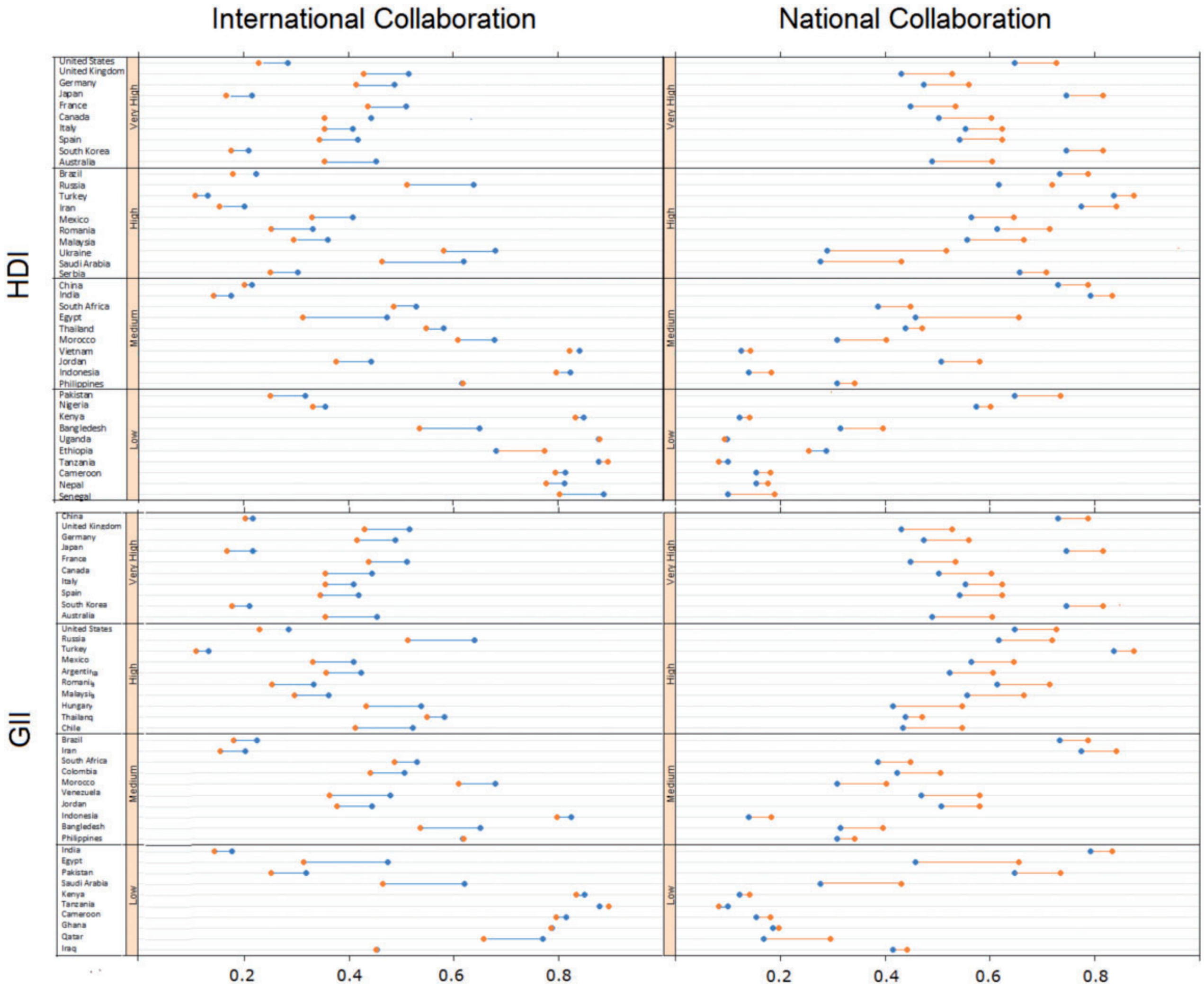


Figure 5. Female and male collaboration rates, ranked by top most productive countries in each quartile, for HDI and GII. (orange/lighter represents female collaboration rate, blue/darker represents male collaboration rate).

and the particular disadvantage of women in lower-developed countries and those countries with lower GII rankings. Figs. 6 and 7 show the average relative citation (see Section 3 for details).

4.4 Correlations and regressions

As mentioned earlier, HDI and GII are highly correlated. Table 2 presents the Spearman rho correlation for HDI,[2] GII, and the gender disparities in the three scientific indicators (i.e., productivity, collaboration, and impact) (see Section 3 for detailed methods of calculating d_p, d_c, and d_i). The number used for productivity is the closeness to the gender equality measurement, collaboration is calculated as the absolute value of the difference between FIRs and MIRs, and impact is disparity between female and male citation rates (incorporating five different authorship types). Therefore, for each of these indicators, a lower number indicates less gender disparity.

Table 2. Spearman's rho correlation by HDI, GII, d_p, d_c and d_i[3]

	HDI	GII	Productivity	Collaboration	Impact
HDI	1				
GII	0.88**	1			
Productivity	0.17*	0.28**	1		
Collaboration	−0.42**	−0.32**	0.07	1	
Impact	0.54**	0.49**	0.03	−0.22**	1

**Significant at 0.01 level; *significant at 0.05 level.

As shown in Table 2, HDI and GII are highly and significantly correlated. Productivity and impact are significantly and positively correlated with both GII and HDI, but collaboration is negatively correlated with HDI and GII. This indicates that the higher the HDI and GII rank, the larger the disparity between FIRs and MIRs.

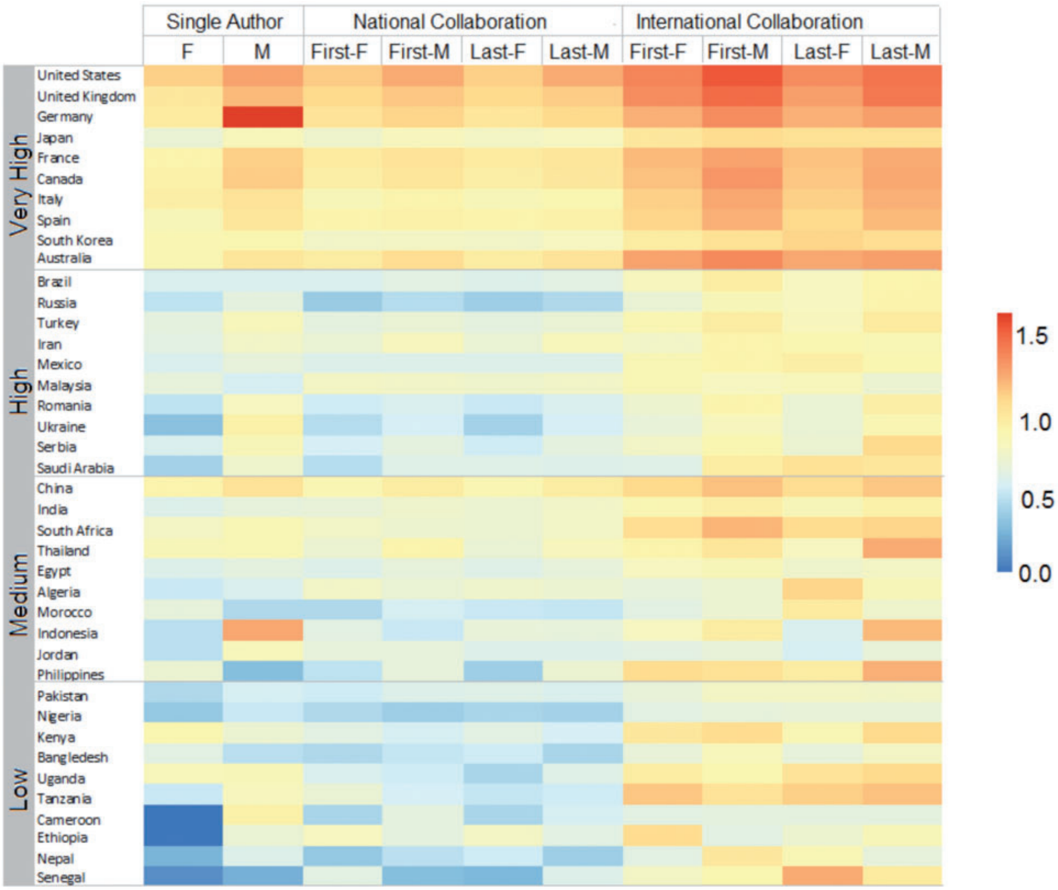


Figure 6. Average relative citations for top 10 countries (by output) in each HDI level.

4.4.1 HDI, productivity, collaboration, and impact. As mentioned earlier, the HDI value of each country is an overall indication of its level of human development, measured from three dimensions (health, education, and living standards) with four individual indicators (life expectancy at birth, mean years of schooling, expected mean years of schooling, and GNI per capita).[4] Multiple linear regression is used to investigate the contribution of a country's human development in the three dimensions to its gender disparities regarding scientific productivity, collaboration, and impact. Additionally, the relationship between productivity, collaboration, impact and HDI indicators by country's HDI level is modeled to examine whether gender disparities in productivity, collaboration, and impact vary by HDI level and whether the contributing factors vary.

The possible multicollinearity issue was considered in our initial modeling process. In regression analysis (see Table 3), there are multiple ways to diagnose multicollinearity, among which the variance inflation factor (VIF) and tolerance (TOL) are commonly used. The VIF estimates how much the variance of a coefficient is 'inflated' because of linear dependence on other independent variables. A variable whose VIF value is greater than four may merit further investigation. The TOL is an indication of the percentage of variance in the independent variable that cannot be accounted for by the other independent variables, hence very small values indicate that an independent variable is redundant, and values under 0.25 usually merit further investigation. We tested the multicollinearity in our initial model, and did not find severe issues among the four indicators of HDI (see Appendix 1 for details).

Additionally, possible outliers in the model were detected in order to avoid possible distortion in the analysis. Cook's distance was used to detect outliers. Cook's distance is commonly used to estimate the influence of a data point when performing least squares regression analysis. The cut-off point of Cook's distance for an outlier vary, and here we used Cook's $d > 4/(n-(k+1))$ as the criteria (where n is the number of subjects and k is the number of independent variables). Outlier countries regarding productivity, collaboration, and impact were removed in the subsequent regression analysis (see Appendix 2 for a list of countries detected as outliers).

Table 4 displays the regression results for productivity, collaboration, and impact with the four indicators of HDI separately for all countries and countries by HDI level. For gender disparities in productivity, all indicators with

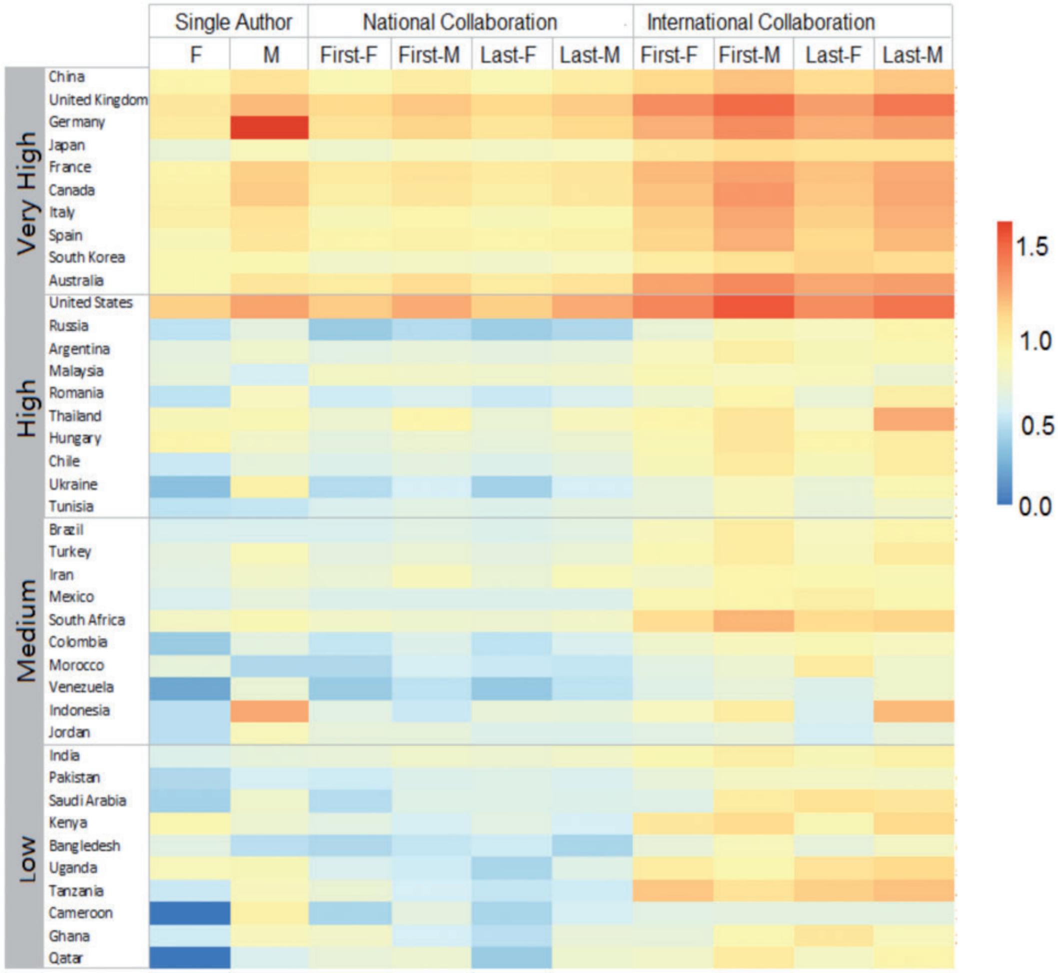


Figure 7. Average relative citations of top 10 countries with highest output in each GII level.

the exception of life expectance at birth matter significantly. All indicators are significant for collaboration. The only variable that contributes significantly to disparities in impact is the GNI per capita. R-square values for each of the three regressions are displayed. R-square is typically (but not exclusively) used to describe how much variation in the data can be explained by the proposed model. It is not surprising to find low R-square values for these regression analyses, as the factors contributing to the gender disparities in the three aspects are complicated and more than just these four indicators of HDI. Therefore, the low R-square value in this analysis is not unexpected.

Countries with different HDI levels vary in the contribution of the four HDI indicators to gender disparities in productivity, collaboration, and impact. For instance, life expectancy at birth, expected mean years of schooling, and GNI per capita contribute significantly to the gender disparities in productivity in countries having a very high HDI level, but what contributes significantly at high (and also low) HDI countries is the mean years of schooling. As for disparities in collaboration, the expected mean years of schooling seems to be a significant contributing factor in all countries except for high HDI countries. For countries of medium HDI level, no significant contributors to disparities in productivity were found. While for countries of high HDI level, no significant contributors of impact disparities were found.

Given that all of our measurements (output, collaboration, and impact) are degrees of disparity, it is reasonable that negative correlations are found with items such as education levels—implying that as the education levels in a country increase, gender disparities in these dimensions decrease.

4.4.2 GII, productivity, collaboration and impact. There are seven individual indicators that are used to measure the gender inequality index: maternal mortality ratio, adolescent fertility rate, female national parliament seat, female secondary education, male secondary education, female labor force rate and male labor force rate. In modeling the relationship between GII indicators and gender disparities in productivity, collaboration, and impact, we calculate the difference between females and males in secondary education and in the labor force

Table 3. Dependent and independent variables for HDI regression analysis

Model	Dependent variable	Independent variables	Notes
Model 1	Productivity, d_p	LifeExpBirth SchoolMeanYear ExpectedMeanSchoolYear GNIPerCapita	Modeling based on full list of countries Individual modeling on countries by their HDI levels
Model 2	Collaboration, d_c	LifeExpBirth SchoolMeanYear ExpectedMeanSchoolYear GNIPerCapita	Modeling based on tfull list of countries Individual modeling on countries by their HDI levels
Model 3	Impact, d_i	LifeExpBirth SchoolMeanYear ExpectedMeanSchoolYear GNIPerCapita	Modeling based on full list of countries; Individual modeling on countries by their HDI levels

Table 4. Regression result of HDI and productivity, collaboration and impact

	Independent variable	Very high	High	Medium	Low	All
Output	(Intercept)	−1.14E+00	7.55E-01	1.78E-02	1.06E+00***	8.22E-01***
	LifeExpBirth	2.60E-02*	4.44E-06	7.44E-03	−5.70E-03	1.12E-03
	SchoolMeanYear	1.08E-02	−6.52E-02*	3.79E-02	−3.02E-02.	−2.20E-02*
	ExpectedMeanSchoolYear	−4.09E-02*	1.05E-02	−2.47E-02	−1.35E-02	−2.28E-02*
	GNIPerCapita	4.11E-03*	8.09E-03	2.82E-06	1.17E-01**	6.18E-06***
	R-square	0.382	0.243	0.070	0.264	0.156
Collaboration	(Intercept)	−6.65E-01	4.65E-02	1.47E-01	5.70E-02	−1.35E-02
	LifeExpBirth	1.60E-02*	2.17E-04	9.45E-04	5.18E-04	1.83E-03.
	SchoolMeanYear	1.30E-02	−3.31E-04	8.83E-03*	3.18E-03	5.94E-03.
	ExpectedMeanSchoolYear	−3.88E-02**	−1.80E-03	−2.06E-02**	−6.20E-03*	−8.88E-03*
	GNIPerCapita	−1.53E-03	1.78E-03*	7.20E-04	−9.37E-03	9.86E-07.
	R-square	0.235	0.148	0.247	0.140	0.108
Impact	(Intercept)	5.03E+00**	9.95E-01	6.79E+00*	1.62E+00*	3.31E + 00**
	LifeExpBirth	−6.10E-02**	−1.68E-03	−6.10E-02.	−5.56E-03	−3.31E-02
	SchoolMeanYear	−5.58E-02	2.50E-02	1.00E-01	−1.71E-01*	−1.64E-05
	ExpectedMeanSchoolYear	3.62E-02	−3.13E-02	−2.63E-01	2.40E-02	−1.60E-02
	GNIPerCapita	8.26E-03*	−2.76E-03	1.80E-01*	8.30E-02	8.75E-07*
	R-square	0.234	0.024	0.283	0.165	0.159

Significant codes: 0'***'; 0.001'**'; 0.01'*'; 0.05'.'

rate as the two new indicators (see Table 5), constructing five indicators in the regression analysis, which eliminates possible multicollinearity problems among these indicators (see Appendix 3 for detailed VIF and TOL on the indicators). Cook's distance was also calculated to detect outliers (see Appendix 4 for the influence plot and outlier countries).

Similar to the regression analysis on HDI indicators, the impact of GII indicators on gender disparities in productivity, collaboration, and impact was also modeled for all countries and countries by GII level, as shown in Table 6. For all countries, the difference in female labor force rate and male labor force rate (i.e., FMLaborForce) matters for all three aspects of gender disparities considered here. Results also show that the FMLaborForce contributes significantly to the gender disparities in productivity in countries regardless of their GII level. No such factors can be identified for disparities in productivity and impact. The R-square values of models by each GII level are on average higher than that of models for all countries, which to some extend indicates that the regression analysis at individual GII level is more meaningful overall.

As with the HDI model, many significant correlations are negative, demonstrating that, for instance, when there are proportionally more women in the labor force, disparities decrease.

Table 5. Dependent and independent variables for HDI regression analysis

Original indicator	New indicators
Maternal mortality ratio Adolescent fertility rate Female national parliament seat	Maternal mortality ratio Adolescent fertility rate Female national parliament seat
Female secondary education Male secondary education	FM secondary education = female secondary education- male secondary education
Female labor force rate Male labor force rate	FM labor force = female labor force rate – male labor force rate

Table 6. Regression result of GII and productivity, collaboration and impact

	Independent variable	Very High	High	Medium	Low	All
Output	(Intercept)	5.22E-01*	2.82E-01*	1.14E-01	5.28E-01***	3.86E-01***
	MaterMorta	7.36E-03.	8.11E-04	8.16E-04*	8.69E-05	3.36E-04**
	AdoFert	−2.02E-02**	7.56E-04	1.96E-04	7.87E-04	−6.73E-04
	FemaleParliament	6.65E-04	−4.31E-03	−1.50E-03	−4.91E-03.	−9.11E-04
	FMSecEdu	1.55E-03	7.66E-03	−7.90E-03	−1.23E-03	−2.88E-03
	FMLaborForce	−5.34E-03.	−9.04E-03***	−7.89E-03*	−3.38E-03*	−4.19E-03***
	R-Square	0.349	0.392	0.359	0.241	0.165
Collaboration	(Intercept)	5.04E-02*	1.22E-01***	2.21E-02	9.66E-02***	6.80E-02***
	MaterMorta	−21E-03**	−5.22E-04	9.07E-05	−6.68E-06	1.78E-07
	AdoFert	6.23E-04	−5.95E-04	−1.45E-04	−4.44E-04**	−3.48E-04***
	FemaleParliament	4.38E-04	2.57E-04	−5.51E-04	−7.00E-04	−2.31E-04
	FMSecEdu	8.90E-05	3.44E-03*	1.02E-04	1.48E-03*	6.19E-04
	FMLaborForce	−5.36E-04	1.57E-04	−1.02E-03*	−5.51E-04.	−6.04E-04*
	R-Square	0.358	0.191	0.194	0.510	0.158
Impact	(Intercept)	3.52E-01	6.27E-01*	1.58E + 00*	1.58E + 00	1.07E + 00***
	MaterMorta	7.15E-03	8.95E-03*	−1.05E-03	3.26E-04	1.18E-03*
	AdoFert	6.26E-03	−8.53E-04	−5.71E-03	8.17E-04	3.33E-03
	FemaleParliament	1.18E-03	−1.95E-02*	3.54E-03	−2.67E-04	−1.62E-02*
	FMSecEdu	8.79E-03	1.25E-02	−5.44E-03	2.39E-02*	1.59E-02.
	FMLaborForce	9.30E-03	−1.41E-04	1.55E-02	1.46E-02*	9.14E-03.
	R-Square	0.101	0.292	0.090	0.375	0.185

Significant codes: 0‘***’; 0.001‘**’; 0.01‘*’; 0.05‘.’

5. Discussion

The discussion has been organized to provide a list of research questions, by section, followed by a discussion of the relevant results.

5.1 Scientific output

RQ1a. What is the relationship between gender parity in scientific output and levels of human development?

H1a. We would expect that levels of gender parity in scientific output and levels of human development would be related: that is, that gender parity would increase in countries with higher levels of human development.

RQ1b. What is the relationship between gender parity in scientific output and levels of gender equity?

H1b. We would expect that levels of gender parity in scientific output and levels of gender equity would be related; that is, that gender parity in scientific output would increase in countries with higher level of gender equity.

The first set of hypotheses for this study expected to find higher levels of gender parity in scientific output in countries that were more developed and those which were ranked higher in terms of gender equity according to country-level variables. We found mixed results in this regard. Countries ranked as ‘low’ in both the HDI and GII had the lowest contribution of female participation in science. However, there were very few differences to be found between the first three quartiles (very high, high, and medium), although the levels of variability increased across this spectrum (with high having the lowest variability). This suggests that, at the macro-level, HDI and GII rank are not sufficient to explain differences in the participation of women in science. This may also reinforce the notion that development is uneven, even within countries ranked as highly developed (Ynalvez and Shrum 2011), and the quartile rankings are not precise enough to capture distinctions in development.

> RQ2a. What is the relationship between a country's capacity to produce (i.e., population) and overall production, gender parity in scientific output, and levels of human development?

> H2a. We would expect that the most productive countries (normalized by population size), would tend to be those in highly developed countries and that these countries would have higher levels of gender parity in scientific output.

> RQ2b. What is the relationship between a country's capacity to produce and overall production, gender parity in scientific production, and levels of gender equity?

> H2b. We would expect that the most productive countries (normalized by population size), would tend to be those in countries with high degrees of gender equity and that these countries would also be those trending towards gender parity in scientific output.

We find that countries that are the most productive (normalized by population) also tend to be those with high levels of human development and those ranked highly by the GII. While there is a clear distinction between very high and other categories, the difference in productivity is diminished in the other quartiles. However, we do not see, as expected, a clear relationship between gender parity and countries highly ranked by HDI and GII. On average, countries ranked as high have greater gender parity than countries ranked as very high in HDI. Countries ranked as medium in GII come closest to gender parity in scientific output. This suggests that gender parity is not associated with the highest level of development, but is more likely to occur, on average, in middle-developed nations.

5.2. Collaboration

> RQ3a. What is the relationship between gendered participation in collaboration and levels of human development?

> H3a. We expect that this difference will be smaller in highly developed countries and more pronounced in lesser developed countries. Furthermore, we expect higher levels of international collaboration in highly developed countries.

> RQ3b. What is the relationship between gendered participation in collaboration and levels of gender equity?

> H3a. We expect that women's portfolios will be more domestic than their male compatriots (as shown in Larivière et al. 2013). However, we expect that this difference will be more pronounced in countries with low gender equity (i.e., low GII rankings).

We expected to find higher levels of international collaboration in highly developed countries and those ranked highly in the GII, given that collaboration does not seem to lead to increased productivity in developing countries (Ynalvez and Shrum 2011). However, our analysis demonstrated that highly developed countries favored national collaborations while lesser developed countries and those with lower levels of GII sought international collaboration at a greater rate. It is likely that this is because of the high level of resources concentrated in highly developed countries—there is more to be gained from a less developed country collaborating with a more developed country than vice versa.

5.3. Impact

> RQ4a. What is the relationship among gender, collaboration, and impact, by levels of human development?

> H4a. Previous research (Larivière et al. 2013) has shown that articles with women in dominant author positions are less cited than those with men in these positions. We expect that this difference will be smaller in more developed countries. We also expect to see higher levels of overall impact for women coming from more developed countries.

> RQ4b. What is the relationship among gender, collaboration, and impact, by levels of gender equity in a country?

> H4b. We expect to see higher impact for articles with women in dominant author positions in countries with high levels of gender equity and smaller differences between dominant male/female authors in countries with higher gender equity.

We anticipated that the impact of publications would also be highly related to the country's HDI and GII ranking, with higher impact associated with higher ranking. This was confirmed by our analysis: countries ranked very high in both HDI and GII also had higher levels of impact per paper and levels of impact decreased by quartile, with the exception of a few countries (e.g., China). In nearly all countries, however, women's impact remained lower than men's. There did not appear to be a significant difference in the relative difference between male and female citedness by levels of HDI and GII, suggesting that these indicators of human development are not sufficient to explain this difference.

5.4 Relationship among output, collaboration, and impact

> RQ5a. What is the relationship between gender parity in scientific output, collaboration, and impact and each HDI indicator, by levels of development?

H5a. We would anticipate that those indicators most related to education would be the most related to gender parity in science. We would expect that these issues would be constant across levels of development.

RQ5b. What is the relationship between gender parity in scientific output, collaboration, and impact and each GII indicator, by indicator rank?

H5b. We would anticipate that those indicators most related to female participation in the labor market would be most related to the gender parity in scientific output. We would expect that these issues would be constant across levels of gender equity.

Although significant, there was a fairly low correlation between the gender parity in terms of output, and HDI and GII ranks, suggesting a weak relationship between these. A negative relationship was found between HDI and GII ranks and collaboration, demonstrating that the disparity between FIRs and MIRs increases as the HDI and GII rank decreases. This reinforces the finding by Ynalvez and Shrum (2011) of a male advantage in the number of collaborative projects in developing countries. This may be explained by the gendered difference in collaboration motivation: with men taking 'instrumental' and 'experience' strategies to form collaborations, while women and men tended to be motivated by 'mentoring' (Bozeman and Gaughan 2011).

The model for the HDI indicators found that all indicators except life expectancy at birth were significant in explaining gender differences in output. Expected schooling was the most significant contributor to collaboration and GNI per capita was the only significant variable contributing to differences in impact. This disagrees with our hypotheses, which proposed that educational variables would be most significant. We also hypothesized that indicators would be stable across ranks. This was found to be false. Variations were found by rank. For example, in 'very high' countries, life expectancy at birth was the most significant indicator for impact differences; in 'low' countries, mean years of education was the only significant indicator for impact differences. This demonstrates the complexity of identifying a single variable in explaining gender disparities in scientific output.

This complexity was reinforced in the model for GII. Only two indicators were significant in explaining disparities in output for the aggregate country set: maternal mortality rates and the ratio of women to men in the labor force. The finding in regards to the labor force reinforce the findings of Hunt et al. (2013), who found that women's underrepresentation in engineering and jobs related to patenting explained much of the gender gap in patenting. Adolescent fertility rates and female participation in the labor force were significant in terms of collaboration. All indicators except adolescent fertility rates were significant in explaining gender disparities in impact. As with the HDI model, large differences were seen across quartiles.

5.5 Limitations

The exploratory nature of this study introduced limitations in the generalizability of the work. For example, it could be argued that a number of additional variables could have been reasonably added to the models (e.g., levels of religion, internet use, salary information, childcare etc.). However, HDI and GII were used given that they are reasonably stable, cover a large number of countries, and have been employed and validated in a number of studies. These indices were therefore the most appropriate for a comprehensive study at the global level. Future work could build upon this foundation and incorporate other comprehensive and valid country-level indicators.

6. Conclusions

This study provided a large-scale analysis exploring the relationship between human development and gender equity indicators with scientometric measures. The study found that countries ranked in the lowest quartile for human development and gender equity had the lowest levels of female contribution to scientific output, but found little difference in the higher quartiles, suggesting that levels of development and gender equity are imprecise predictors of gender equality in science.

The findings regarding collaboration are promising in terms of future research. Women have been shown to be more likely to engage in interdisciplinary research (van Rijnsoever and Hessels 2011). However, it has also been suggested that women are risk averse and suffer from lack of established professional networks (Abreu and Grinevich 2013). Therefore, as universities reorganize into centers promoting collaborative research and 'grand challenges' (Su 2014), they might consider the composition of the teams—both in terms of gender inclusivity and gender balance in types of tasks and specializations (to avoid the construction of 'antagonistic subgroups' (Faems and Subramanian 2013)). Future research should examine the ways in which gender disparities could be addressed through policies that promote collaboration, particularly between developed and less developed nations.

Heightened awareness of gender disparities in science is also necessary given the high degree of mobility in the scientific workforce. For instance, in the US, one-third of science and engineering faculty members are foreign-born (Sabharwal 2011). Highly developed countries are also typical loci for foreign-born students. Institutions, research centers, and collaborative teams should therefore be aware that cultural values around gender and scientific work may impact the way in which scholars approach and engage in research. Institutional training for those in the position to train and mentor foreign-born doctoral and post-doctoral students is imperative.

Further investigation is also needed in regards to the negative relationship between the wealth of the country and the gender disparity in impact. Our results suggest that gender disparities in science lessen in wealthier countries. There are many potential explanations for this. One is that in a low resource country, men receive more support in scientific work; whereas in wealthy countries, both sexes receive equal resources. Further analysis is necessary to clarify the nature of this relationship. Regardless, this finding demonstrates the relationship between wealth and the ability to exert influence in scientific discourse. This has clear policy implications and suggests an obligation on the part of the scientific community to find ways to globalize wealth in the international science community, thereby making visible the contributions of previously marginalized voices in the scientific community. Many national funding bodies are beginning to incentivize international projects (e.g., The Digging into Data Challenge focuses on using big data approaches to solve problems in the humanities and social sciences. It combines funding from several research funders across the globe and requires that research teams are international and interdisciplinary in composition. http://diggingintodata.org/), but more could be done, particularly in expanding these incentives to include scholars from less wealthy nations.

Merton (1973) argued that 'careers should be open to talent'. However, our analysis shows systemic gender disparities across countries, regardless of levels of human development. There is no single solution for these disparities. Factors highly related to gender disparities in a highly developed country are not those at work in a less developed country. These particularistic factors suggest that solutions to achieve equality must be tailored to the particular context and that blanket solutions may not be effective. Effective initiatives in one country are not necessarily appropriate for application in another. Furthermore, the scientific community must begin to think about the global implications of science policy and continually address whether all voices are receiving an opportunity to equitably participate in the global scientific conversation.

Funding

This work was supported by the Science of Science Innovation and Policy program of the National Science Foundation (grant number 1158670).

Acknowledgements

An earlier version of this paper was presented at the Atlanta Conference on Science and Innovation Policy, held 26–28 September, 2013 in Atlanta, Georgia at the Georgia Tech Global LEarning Center and received valuable comments from the audience. The authors would also like to thank Andrew Tsou for his work in proofreading and formatting the references.

Notes

1. The suffixes (-ova, -ina, -eva, -aia, -aya) are typically associated with females. The suffixes (-ov, -in, -ev, -ky, -kii, -kiy, -yi, -ny, -oy, -oi, except -tsoi and -tsoy) are typically associated with males.
2. The Human Development Report by the UNDP reports the overall value of the Human Development Index, as well as the values for each of the five specific indicators of HDI at different dimensions. The Spearman correlation here is based on the overall value and rank of HDI. Similar for GII used here for the overall Spearman rho.
3. It should be noted that only countries with both HDI and GII indicators available from the UNDP are analyzed for the overall correlation.
4. The technical notes from UNDP provide a detailed description of each indicator <http://hdr.undp.org/en/2013-report> accessed 01 March 2015.

References

Abreau, M. and Grinevich, V. (2013) 'The nature of academic entrepreneurship in the UK: Widening the focus on entrepreneurial activities', *Research Policy*, 42: 408–22.

Anonymous. (2013) 'Science for all', *Nature*, 495/7439: 5.

Archambault, É., Vignola-Gagne, E., Côté, G., Larivière, V. and Gingras, Y. (2006) 'Benchmarking scientific output in the social sciences and humanities. The limits of existing databases', *Scientometrics*, 68: 329–42.

Bordons, M., Morillo, F., Fernandez, M. T. and Gomez, I. (2003) 'One step further in the production of bibliometric indicators at the micro level: Differences by gender and professional category of scientists', *Scientometrics*, 57: 159–73.

Bozeman, B., Dietz, J. and Gaughan, M. (2001) 'Scientific and technical human capital: An alternative model for research evaluation', *International Journal of Technology Management*, 22: 636–55.

—— and Gaughan, M. (2011) 'How do men and women differ in research collaborations? An analysis of collaborative motives and strategies of academic researchers', *Research Policy*, 40: 1393–402.

Ceci, S. J. and Williams, W. M. (2011) 'Understanding current causes of women's underrepresentation in science', *Proceedings of the National Academy of Science*, 108: 3157–62.

Cole, J. R. and Zuckerman, H. (1984) 'The productivity puzzle: Persistence and change in patterns of publication of men and women scientists', *Advances in Motivation and Achievement*, 2: 217–58.

Council of Canadian Academies. (2012), 'Strengthening Canada's research capacity: The gender dimension: The expert panel on women in university research'.

Ding, W. W. (2006) 'Gender differences in patenting in the academic life sciences', *Science*, 313: 665–7.

—— and Choi, E. (2011) 'Divergent paths to commercial science: A comparison of scientists' founding and advising activities', *Research Policy*, 40: 69–80.

Etzkowitz, H., Kemelgor, C. and Uzzi, B. (2000) *Athena Unbound: The Advancement of Women in Science and Technology*. New York: CUP.

——, Fuchs, S., Gupta, N., Kemelgor, C. and Ranga, M. (2008) 'The coming gender revolution in science'. In: Hackett, E. J., Amsterdamska, O., Lynch, M. E. and Wajcman, J. (eds) *The Handbook of Science and Technology Studies*, pp. 403–28. Cambridge, MA: MIT Press.

European Commission. (2007), *Renumeration of Researchers in the Public and Private Sectors,* Service Contract REM 01 <http://www.eui.eu/Documents/MWP/AcademicCareers/SalaryComparisonEUreport2007.pdf> accessed 02 March 2015.

Faems, D. and Subramanian, A. M. (2013) 'R&D manpower and technological performance: The impact of demographic and task-related diversity', *Research Policy*, 42: 1624–33.

Frietsch, R., Haller, I., Funken-Vrohlings, M. and Grupp, H. (2009) 'Gender-specific patterns in patenting and publishing', *Research Policy*, 38: 590–9.

Fox, M. F. (2005) 'Gender, family characteristics, and publication productivity among scientists', *Social Studies of Science*, 35: 131–50.

Ganzi, A., Sugimoto, C. R. and Didegah, F. (2012) 'Mapping world scientific collaboration: Authors, institutions, and countries', *Journal of the American Society for Information Science and Technology*, 63: 323–35.

Gonzalez-Brambila, C. and Veloso, F. M. (2007) 'The determinants of research output and impact: A study of Mexican researchers', *Research Policy*, 36: 1035–51.

Haeussler, C. and Colyvas, J. A. (2011) 'Breaking the ivory tower: Academic entrepreneurship in the life sciences in UK and Germany', *Research Policy*, 40: 41–54.

Hagmann, M. (1999) 'EU confronts the gender gap', *Science*, 286/5445: 1659.

Holden, C. (2001) 'General contentment masks gender gap in first AAAS salary and job survey', *Science*, 294: 396–411.

Human Development Report. (2013a) *Gender Inequality Index (GII)* <http://hdr.undp.org/en/statistics/gii/> accessed 13 August 2014.

—— (2013b) *Human Development Index (HDI)* <http://hdr.undp.org/en/statistics/hdi/> accessed 13 August 2014.

Hunt, J., Garant, J.-P., Herman, H. and Munroe, D. J. (2013) 'Why are women underrepresented amongst patentees?', *Research Policy*, 42: 831–43.

International Labour Organization. (2013) *Women and Technology—The Attitude Gap*, <http://www.ilo.org/global/about-the-ilo/newsroom/features/WCMS_195857/lang–en/index.htm> accessed 13 August 2014.

Larivière, V., Archambault, E., Gingras, Y. and Vignola-Gagne, E. (2006) 'The place of serials in referencing practices: Comparing natural sciences and engineering with social sciences and humanities', *Journal of the American Society for Information Science and Technology*, 57: 997–1004.

——, Ni, C., Gingras, Y., Cronin, B. and Sugimoto, C. R. (2013) 'Global gender disparities in science', *Nature*, 504/7479: 211–3.

——, Sugimoto, C., Tsou, A. and Gingras, Y. (2014) 'Team size matters: Collaboration and scientific impact since 1990', *Journal of the Association for Information Science and Technology*, (Unpublished).

——, Vignola Gagné, E., Villeneuve, C., Gélinas, P. and Gingras, Y. (2011) 'Sex differences in research funding, productivity, and impact: An analysis of Quebec university professors', *Scientometrics*, 87: 483–98.

Leahey, E. (2006) 'Gender differences in productivity–Research specialization as a missing link', *Gender & Society*, 20(6): 754–80.

Ley, T. J. and Hamilton, B. H. (2008) 'The gender gap in NIH grant applications', *Science*, 322: 1472–4.

Long, J. S. (1992) 'Measures of sex differences in scientific productivity', *Social Forces*, 71: 159–78.

—— and Fox, M. F. (1995) 'Scientific careers: Universalism and particularism', *Annual Review of Sociology*, 21: 45–71.

Mauleón, E. and Bordons, M. (2006) 'Productivity, impact and publication habits by gender in the area of materials science', *Scientometrics*, 66: 199–218.

McCook, A. (2013) 'Barred from the boardroom: A broken business model', *Nature*, 495: 25–7.

Merton, R. K. (1973) *The Sociology of Science*. Chicago, IL: University of Chicago Press.

Moed, H. F. (1996) 'Differences in the construction of SCI based bibliometric indicators among various producers: A first overview', *Scientometrics*, 35: 177–91.

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J. and Handelsman, J. (2012) 'Science faculty's subtle gender biases favor male students', *Proceedings of the National Academy of Sciences*, 109: 16474–9.

Mühlenbruch, B. and Jochimsen, M. A. (2013) 'Only wholesale reform will bring equality', *Nature*, 495: 40–2.

OECD. (2012) *Education at a Glance 2012*. Paris: OECD.

Ostergaard, C. R., Timmermans, B. and Kristinsson, K. (2011) 'Does a different view create something new? The effect of employee diversity on innovation', *Research Policy*, 40: 500–9.

Paul-Hus, A., Bouview, R. L., Ni, C., Sugimoto, C. R. *et al.* (accepted) 'Forty years of gender disparities in Russian science: A historical bibliometric analysis', *Scientometrics*, doi:10.1007/s11192-014-1386-4.

Peñas, C. S. and Willett, P. (2006) 'Brief communication: Gender differences in publication and citation counts in librarianship and information science research', *Journal of Information Science*, 32: 480–5.

Pezzoni, M., Sterzi, V. and Lissoni, F. (2012) 'Career progress in centralized academic systems: Social capital and institutions in France and Italy', *Research Policy*, 41: 704–19.

Provasnik, S., Kastberg, D., Ferraro, D., Lemanski, N., Roey, S. and Jenkins, F. (2012) Highlights From TIMSS 2011: Mathematics and Science Achievement of U.S. Fourth- and Eighth-Grade Students in an International Context (NCES 2013-009 Revised). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, DC.

Prpic, K. (2002) 'Gender and productivity differentials in science', *Scientometrics*, 55: 27–58.

Sabharwal, M. (2011) 'Job satisfaction patterns of scientists and engineering by status of birth', *Research Policy*, 40: 853–63.

Schubert, A. and Braun, T. (1986) 'Relative indicators and relational charts for comparative assessment of publication output and citation impact', *Scientometrics*, 9: 281–91.

Seifert, T. A. and Umbach, P. D. (2008) 'The effects of faculty demographic characteristics and disciplinary context on dimensions of job satisfaction', *Research in Higher Education*, 49: 357–81.

Shen, H. (2013) 'Mind the gender gap: Inequality quantified', *Nature*, 495: 22–4.

Sonnert, G. and Holton, G. (1995) *Gender Differences in Science Careers: The Project Access Study*. New Brunswick, NJ: Rutgers University Press.

Su, X. (2014) 'Academic scientists' affiliation with university research centers: Selection dynamics', *Research Policy*, 43: 382–90.

UN Development Programme. (2013) *Frequently Asked Questions (FAQs) About the Human Development Index (HDI)* <http://hdr.undp.org/en/media/FAQs_2011_HDI.pdf> accessed 02 March 2015.

Van Arensbergen, P., van der Weijden, I. and van den Besselaar, P. (2012) 'Gender differences in scientific productivity: A persisting phenomenon?', *Scientometrics*, 93: 857–68.

Van Rijnsoever, F. J. and Hessels, L. K. (2011) 'Factors associated with disciplinary and interdisciplinary research collaboration', *Research Policy*, 40: 463–72.

Wellcome Trust. (2013) *Risks and Rewards: How PhD Students Choose Their Careers*. London: Wellcome Trust.

West, J. D., Jacquet, J., King, M. M., Correll, S. J. and Bergstrom, C. T. (2013) 'The role of gender in scholarly authorship', *PLoS ONE*, 8/7: e66212.

Xie, Y. and Shauman, K. A. (1998) 'Sex differences in research productivity: New evidence about an old puzzle', *American Sociological Review*, 63: 847–70.

—— and —— (2003) *Women in Science: Career Processes and Outcomes*. Cambridge, MA: Harvard University Press.

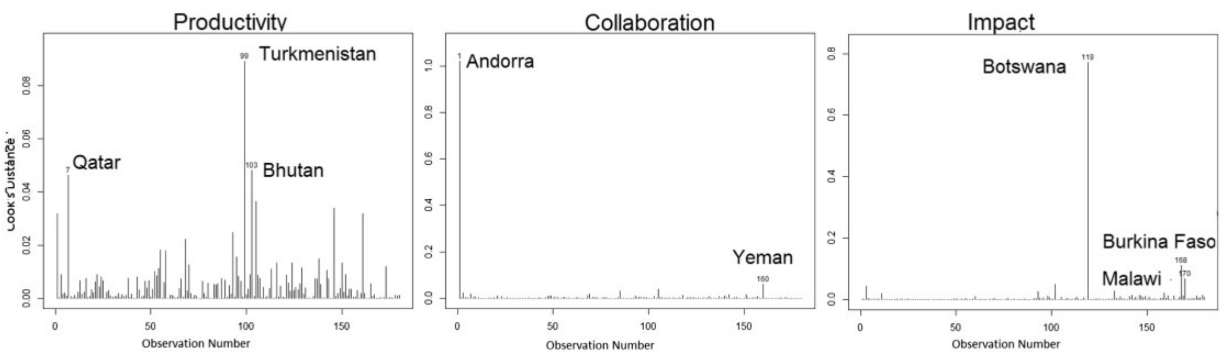
Ynalvez, M. A. and Shrum, W. M. (2011) 'Professional networks, scientific collaboration, and publication productivity in resource-constrained research institutions in a developing country', *Research Policy*, 40: 204–16.

Zuckerman, H. (1991) 'The careers of men and women scientists: A review of current research'. In: Zuckerman, H., Cole, J. R. and Bruer, J. T. (eds) *The Outer Circle. Women in the Scientific Community*, pp. 27–57. New York: W. W. Norton.

Appendix

Appendix 1. Tests of multicollinearity for HDI variables

	LifeExpBirth	SchoolMeanYear	ExpectedMeanSchoolYear	GNIPerCapita
VIF	2.866	3.256	3.843	1.656
TOL	0.349	0.307	0.260	0.604

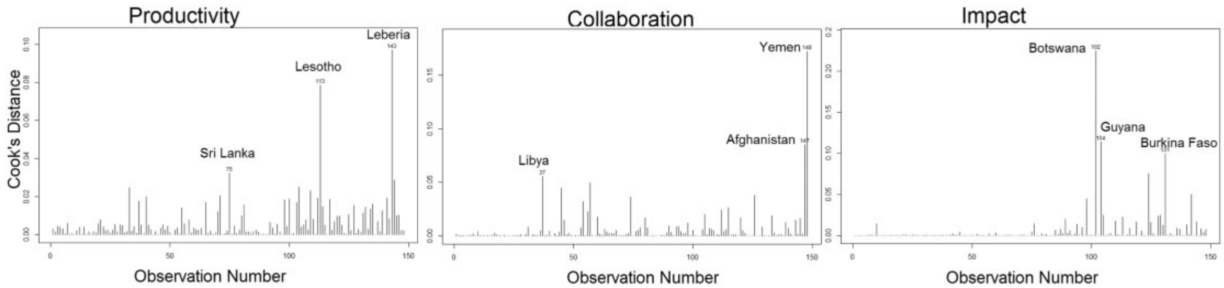


Appendix 2. Outlier countries detected with Cook's distance: HDI.

Appendix 3. Tests of multicollinearity for GII variables. (a) VIF and TOL for original GII indicators; (b) VIF and TOL for GII indicators after transformation

	MaterMorta	AdoFert	FemaleParliament	FSecEdu	MSecEdu	FLabForce	MLabForce
VIF	2.661	2.161	1.090	16.878	16.323	1.395	1.684
TOL	0.376	0.463	0.918	0.059	0.061	0.717	0.594

	MaterMorta	AdoFert	FemaleParliament	FMSecEdu	FMLabForce
VIF	2.098	1.880	1.088	1.181	1.138
TOL	0.47664	0.532	0.919	0.847	0.878



Appendix 4. Outlier countries detected with Cook's distance: GII.

Appendix 5. Research questions, hypotheses, and support

Identifier	Research question	Hypothesis	Support
RQ1a	What is the relationship between gender parity in scientific output and levels of human development?	We would expect that levels of gender parity in scientific output and levels of human development would be related; that is, that gender parity would increase in countries with higher levels of human development.	Mixed. Countries ranked 'low' in HDI had the lowest contribution of female participation in science. However, there were few differences found among the other three quartiles, with 'high' having greater female output than 'very high'.
RQ1b	What is the relationship between gender parity in scientific output and levels of gender equity?	We would expect that levels of gender parity in scientific output and levels of gender equity would be related; that is, that gender parity in scientific output would increase in countries with higher level of gender equity.	Mixed. Countries ranked 'low' in GII had the lowest contribution of female participation in science. However, there were few differences found among the other three quartiles.
RQ2a	What is the relationship between a country's capacity to produce (i.e. population) and overall production, gender parity in scientific output, and levels of human development?	We would expect that the most productive countries (normalized by population size), would tend to be those in highly developed countries and that these countries would have higher levels of gender parity in scientific output.	Not supported. Highly productive countries (normalized by population size) tend to be those ranked highly in HDI. However, this is not matched by gender parity: 'high' countries have greater gender parity than countries ranked as 'very high' in HDI.
RQ2b	What is the relationship between a country's capacity to produce (i.e. population) and overall production, gender parity in scientific production, and levels of gender equity?	We would expect that the most productive countries (normalized by population size), would tend to be those in countries with high degrees of gender equity and that these countries would also be these trending towards gender parity in scientific output.	Not supported. Highly productive countries (normalized by population) tend to be those ranked highly in GII. However, this is not matched by gender parity. Countries ranked as 'medium' in GII come closest to gender parity in scientific output.
RQ3a	What is the relationship between gendered participation in collaboration and levels of human development?	We expect that this difference will be smaller in highly developed countries and more pronounced in lesser developed countries. Furthermore, we expect higher levels of international collaboration in highly developed countries.	Not supported. Countries of low development seek international collaboration at a greater rate than highly developed countries, but levels of development do not change the ratio of domestic/international collaboration between men and women.
RQ3b	What is the relationship between gendered participation in collaboration and levels of gender equity?	We expect that women's portfolios will be more domestic than their male compatriots (as shown in Larivière et al. 2013); however, we expect that this difference will be more pronounced in countries with low gender equity (i.e. low GII rankings).	Not supported. Countries with low gender equality seek international collaboration at a greater rate than countries with higher gender equality, but levels of equality do not change the ratio of domestic/international collaboration between men and women.
RQ4a			

(continued)

Appendix 5. Continued

Identifier	Research question	Hypothesis	Support
	What is the relationship among gender, collaboration, and impact, by levels of human development?	Previous research (Larivière et al. 2013) has shown that articles with women in dominant author positions are less cited than those with men in these positions. We expect that this difference will be smaller in more developed countries. We also expect to see higher levels of overall impact for women coming from more developed countries.	Supported. Women in highly developed countries had a clear citation advantage. Previously supported claims about women in dominant author positions and in collaborations held.
RQ4b	What is the relationship among gender, collaboration, and impact, by levels of gender equity in a country?	We expect to see higher impact for articles with women in dominant author positions in countries with high levels of gender equity and smaller differences between dominant male/female authors in countries with higher gender equity.	Supported. Women in countries ranked very high in gender equality had a clear citation advantage. Previously supported claims about women in dominant author positions and in collaborations held.
RQ5a	What is the relationship between gender parity in scientific output, collaboration, and impact and each HDI indicator, by levels of development?	We would anticipate that those indicators most related to education would be the most related to gender parity in science. We would expect that these issues would be constant across levels of development.	Mixed. Education indicators were significant in explaining output and collaboration, but not impact (GNIP was the only indicator that contributed significantly in the model). Not supported. Countries of different HDI levels varied in the contribution of the four HDI indicators to gender disparities in productivity, collaboration, and impact.
RQ5b	What is the relationship between gender parity in scientific output, collaboration, and impact and each GII indicator, by indicator rank?	We would anticipate that those indicators most related to female participation in the labor market would be most related to the gender parity in scientific output. We would expect that these issues would be constant across levels of gender equity.	Mixed. Female participation in the labor force was significant in explaining output and collaboration, but not in explaining impact. Mixed. Countries of different GII levels varied in the contribution of the indicators to gender disparities in productivity, collaboration, and impact.