Titre: Which gender gap? Factors affecting researchers' scientific impact in science and medicine
Title:

Auteurs: Catherine Beaudry, & Vincent Larivière
Authors:

Date: 2016

Type: Article de revue / Article

Référence: Beaudry, C., & Larivière, V. (2016). Which gender gap? Factors affecting researchers' scientific impact in science and medicine. Research Policy, 45(9), 1790-1817. https://doi.org/10.1016/j.respol.2016.05.009
Citation:

Titre de la revue: Research Policy (vol. 45, no. 9)
Journal Title:

Maison d'édition: Elsevier
Publisher:

URL officiel: https://doi.org/10.1016/j.respol.2016.05.009
Official URL:

Mention légale:
Legal notice:

*Research Highlights

Which gender gap? Factors affecting researchers' scientific impact in science and medicine

Highlights

- Academic women publish less and receive fewer citations in health and NSE fields
- Female academics are less successful at raising smaller amounts of funds in NSE fields
- For equal amounts of funding and of publications, NSE women are equally cited
- In journals of similar impact factors, female health scientists are less cited
- And, individuals publishing with a higher proportion of female authors are less cited

Which gender gap? Factors affecting researchers' scientific impact in science and medicine

Catherine Beaudry

Polytechnique Montréal
Centre Interuniversitaire de Recherche sur la Science et la Technologie

Vincent Larivière

Université de Montréal
Centre Interuniversitaire de Recherche sur la Science et la Technologie

Abstract

The article examines whether scientific production, research funding, Impact Factor of journals and size of collaborative teams have an influence on the propensity to receive more citations, and whether the influence of these factors differs across genders. Using a very complete database of funding, scientific papers and citations compiled at the individual researchers' level, we estimate panel data regressions on the discipline-normalised citation rates of individual academics in Quebec. Our results show that although most of the indicators examined have a positive influence on the relative citation rate, when it comes to gender differences, not having enough public funding and raising private funding appear slightly detrimental for women in the health sciences. In addition, when women collaborate with the same number of co-authors as men, or target similar Impact Factor journals, their articles are less cited then those of their male colleagues. Almost no gender effect is found in the natural sciences and engineering where women are still a minority. Our results worryingly show that academics who publish with a larger proportion of female co-authors are less cited. Furthermore, when targeting similar Impact factor journals, researchers who collaborate with a higher proportion of female co-authors are consistently less cited in both the health and NSE fields than if they were publishing with a male dominated group of co-authors.

Keywords: Gender, scientific production, citations, research funding, Impact Factor

1 Introduction

A recent Nature paper (Larivière et al., 2013) confirms that women are lagging behind in terms of worldwide scientific production and in terms of citations, taking into account the authors' ranking (first or last), countries, collaborative practices as well as the citation density of various disciplines. It therefore seems that the glass ceiling is still very much present, despite more than a decade of specific policies aimed at supporting women in science. As Xie and Shauman (1998) state, "Women scientists publish fewer papers than men because women are less likely than men to have the personal characteristics, structural positions, and facilitating resources that are conducive to publication" (1998:863). On these inequalities noted regarding access to research funding and equipment, Larivière et al. (2011) showed that in Quebec, women have raised less research funds than men and that their funding is less diversified, especially in the middle of their careers. The authors suggested that the smaller global scientific production of women is likely to be linked to the fact that women receive less funding than men, but as the authors state: "the data can only establish the correlation and not a causal relationships between these two findings" (2011:491). Although the literature on scientific production is extensive and covers several decades (see, among others, Cole and Zuckerman, 1984; Xie and Shauman, 2003; Zuckerman, 1991), few papers have been published on the subject of what resources, structural positions, teams of collaborators are necessary to improve the impact and quality of articles published by women.

This paper aims to provide a more complete portrait of the performance of women—including both the inputs and outputs of research—, taking the province of Quebec, identified by Larivière et al. (2013) as one of the Canadian provinces closest to achieving gender parity, as an example. With women accounting for 14.7% of Quebec researchers in the natural sciences and engineering fields, and 26.8% of Quebec researchers in the health fields, one could argue that this still remains far from gender parity. Similarly, while women represent more than half of the students at the bachelor level (the first university degree in Quebec), their proportion decreases dramatically after graduation and very few venture into academia. Similar trends are observed for other Canadian provinces. In fact, the highest the academic rank, the lowest is the proportion of women in academia. Although we acknowledge the rarity of women in science in Quebec and their slightly inferior scientific performance, both in terms of funding received and published outputs, our goal is to try to elucidate where the discrepancies are (in terms of funding, scientific production, Impact factor and co-authorship), to explain the differences in impact (using the data available).

A large part of the literature on women in science tends to be bibliometric based. In this paper, we build on this literature and use classic bibliometric indicators as dependent and explanatory variables in econometric models. Using panel data to account for the evolution of the various attributes, we are able to establish the incidence of all of these factors on scientific impact, something that bibliometric methods alone cannot address. The paper also differs from the main sociology of science literature that considers socio-demographic factors such as marriage and children to explain the lesser performance of academic women. These factors, although important are not taken into account in this article, as it was not possible to obtain such information through available data sources. Our approach is therefore somewhat exploratory because we do not have access to data related to maternity leave, parenthood, academic rank, and so on.

Our results show that the visibility accrued by a greater number of articles published, regardless of author rank, publishing in higher impact factor journals and collaborating with a greater number of co-authors all have positive effect on the relative citation rate. In contrast, funding has a mitigated effect. Regarding gender differences, not having enough public funding and raising private funding is detrimental for female health scientists. In addition, when women collaborate with the same number of co-authors as men, or target similar Impact Factor journals, their articles are less cited then those of their male colleagues. Almost no gender differences are found in the natural sciences and in engineering where women are still a minority. More worrying is the fact that our results show that researchers who publish with a larger proportion of female co-authors are less cited. Furthermore, when targeting similar Impact factor journals, researchers who collaborate with a higher proportion of female co-authors are consistently less cited in both the health and NSE fields then if they were co-authoring with a lesser proportion of women. This last result is a worrying trend that suggests systematic gender discrimination.

The remainder of the paper is organised as follows: Section 2 presents the theoretical framework and the resulting hypotheses; Section 3 describes the data and explains the research methodology; Section 4 briefly discusses the descriptive statistics of the main variables used in the regression models that are presented and analysed in Section 5; Section 6 examines the results in light of the proposed hypotheses; Finally, Section 7 discusses the implication of the results and concludes.

2 Theoretical framework

Many scholars have examined the gender differences in research output and scientific impact. Despite their different methods, disciplines and countries on which they focus, these studies generally show that women publish less than their male colleagues (Fox, 2005; Hesli and Lee, 2011; Kyvik and Teigen, 1996; Long, 1992; Zuckerman, 1991); a phenomenon that Cole and Zuckerman (1984) refer to as “the productivity puzzle”. For instance, a 20-30% difference in terms of research production has been measured in favour of men (Prpić, 2002; Xie and Shauman, 1998; 2003), but the gap is closing as the number of women in science increases (Xie and Shauman, 1998; Abramo et al., 2009). In Croatia, however, the productivity gap is increasing (Prpić, 2002). It is generally accepted that this lower scientific productivity is widespread and observed across countries, although it varies across disciplines (Larivière et al., 2013) and is somewhat smaller than what is often portrayed in the literature. For instance, Turner and Mairesse (2005) suggested that female physicists publish on average 0.9 articles less than men, while Gonzalez-Brambila and Veloso (2007) found a difference of 0.07 publications in favour of Mexican male academics. In this latter study, the largest gap was found in the health sciences (0.25 articles) and in physics (0.20 articles). In Canada, Nakhaie (2002) found that the factors that explain why Canadian women publish less than men are seniority, discipline, type of institution and time devoted to research.

Less information is known regarding the citation record of female academics when compared to men, and the evidence presented is rather inconclusive, mainly because of the various methods used as well as the discipline and country of focus. The scientific impact of women’s papers is either similar (Gonzalez-B. and Veloso, 2007; Lewison, 2001; Long and Fox, 1995; Mauleón and Bordons, 2006), superior (Bordons et al., 2003; Long, 1992) or lower (Knobloch-Westerwick and Glynn, 2013; Larivière et al., 2013; Maliniak et al. 2013; Peñas and Willett, 2006) than that of their male colleagues. Gonzalez-Brambila and Veloso (2007) highlighted disciplinary differences in the impact gap, and found that Mexican female natural scientists and health scientists receive 0.05 and 0.14 fewer citations than their male colleagues, while in the social sciences and humanities as well as in engineering, female scientists receive slightly more citations than men (0.02 and 0.04 citations respectively). Other authors found that it takes more time for women to receive their maximal number of citations (Ward, Gast and Grant, 1992), which may explain the differences if the number of citations is calculated up to a specific number of years after publication. Despite the mixed evidence, our first hypothesis stems from this last observation:

H1 Female academics are generally less cited than men.

The first factor that may explain why women are less cited is their lower scientific production. Aksnes et al. (2011) showed that gender differences observed in terms of scientific impact (measured by the number of citations) is attributable to gender differences in scientific productivity (measured by the number of publications). Larivière et al. (2013) found that gender affects the visibility of research output and demonstrated that women in dominant authorship positions (e.g. first or last author) receive fewer citations than their male colleagues. The marginal increase in citation grows with the increase in publication output and because men have more publications, they can benefit more from this advantage, and hence obtain more citations (Aksnes et al., 2011). As women are less productive – and thus visible to the scientific community – they tend to be less cited; a phenomenon that one could call the cumulative disadvantage of women or Matilda effect (Rossiter, 1993). Long (1992) similarly argued that the "smaller number of citations received by females results from their fewer publications, not from the quality of their publications" (1992:159).

In the very few disciplines where men and women are equally prolific, as in dendrochronology (Copenheaver et al., 2010) or academic surgery (Housri et al., 2008), the citation rate of both genders is similar. In other disciplines, such as librarianship and information science, however, even though men contribute to a greater number of papers, their work is not more cited than that of women (Peñas and Willett, 2006). This supports the often-invoked hypothesis that, in research, women focus more on quality than quantity (Sonnert and Holton, 1995). Symonds et al. (2006) even found that in a sample of evolutionary biology and ecology scientists, men tend to go for quantity of publications while women prefer quality of scientific publications and hence are more cited when controlling for the quantity of articles. In light of the evidence presented, our second hypothesis reflects the fact that less productive scientists, because they are less visible or perceived as such, will obtain fewer citations per paper published. In addition, as in most disciplines (Pontille, 2004) the authorship position matters a great deal in some circles, we will modulate this hypothesis by the number of articles published according to the position of the individual in the author list. Women who co-author the same number of papers should therefore be less cited than their male colleagues. Our second hypothesis therefore reads as follows:

H2a Academics who publish a larger number of publications (i) in general, (ii) as first author, (iii) as last author, or (iv) as middle author will be more cited than their less prolific colleagues.

H2b Gender will have a **negative moderating effect** on the relationship between scientific impact and scientific production (i) in general, (ii) as first author, (iii) as last author, or (iv) as middle author, i.e. female academics who publish a larger number of publications will be less cited than men with an equal scientific production.

The second factor that may explain why women are less cited than men is their lower level of research funding. A smaller proportion of women benefit from research funds (Stack, 2004), but both men and women receive an amount of grant proportional to the number of submitted proposals at NIH and NSF (Fox, 1991). Apart from many unpublished bibliometric research reports, little evidence exists as to the influence of research funding on the scientific impact of publications (two recent exceptions being Fortin and Currie, 2013, and Mirnezami et al., 2015). For instance, Mirnezami et al. (2015) showed that only in the natural sciences does funding have a positive and significant influence on the citation rate. What little evidence there is, focuses on the impact of research funding on scientific production. Stack (2004) as well as Xie and Shauman (1998) showed that federal support in the form of grants has a positive impact on scientific production[1]. Beaudry and Allaoui (2012) showed that a larger amount of funding in the form of grants has a positive impact on the number of papers published by an individual nanotechnology scientist. With greater scientific production generally comes more visibility, which in turn should increase the number of citations obtained. We therefore suggest that more funding should also influence the citation rate. It is not obvious that because women are less funded, they should receive fewer citations. For an equivalent amount of dollars raised in research funding, both men and women may exhibit similar citation rates, as they do in most disciplines with the exception of natural sciences (Mirnezami et al., 2015). Nevertheless, we suspect that women with a similar amount of funding will attract a smaller number of citations. Our third hypothesis therefore goes as follows:

H3a Academics who raise more (i) public funding, (ii) private funding, (iii) not-for-profit funding, will be more cited than their colleagues who have raised lesser amounts of research funds.

H3b Gender will have a **negative moderating effect** on the relationship between scientific impact and the amount of (i) public funding, (ii) private funding, (iii) not-for-profit funding, i.e. female academics will be less cited than their male colleagues who have raised similar amounts of research funds.

The notoriety of high Impact Factor journals offers a greater visibility to scientists, which in turn should increase the number of citations received, hence contributing to a somewhat positive feedback loop, or Matthew Effect (Larivière and Gingras, 2010). Because papers published in journals with higher Impact Factors, or for that matter because journals with higher Impact Factors publish articles that are more cited, we anticipate a strong and positive relationship between the number of citations and the Impact Factor of the journal. Following the often-invoked argument that women prefer 'quality' to quantity of articles, female scientists may then target better journals. Because they may concentrate their publications in better journals, their average citation rate may be higher than that of men. Housri et al. (2008) for instance found that women in the academic surgery publish in journals with higher Impact Factors then their male colleagues. In contrast, Bordons et al. (2003) found no significant difference between men and women in terms of the Impact Factor of the journals in which Spanish research council scientists in natural resources and chemistry publish. This brings us to our fourth hypothesis:

H4a Academics who publish in journals with higher Impact Factors will be more cited than their colleagues who target lesser Impact Factor journals.

[1] These studies use a dummy variable taking the value 1 if the scientist has a grant from the federal government and 0 otherwise, and as such cannot measure productivity.

H4b Gender will have a **positive moderating effect** on the relationship between scientific impact and the journals' Impact Factor, i.e. female academics who publish in journals with similar Impact Factor will be more cited than their male colleagues.

Finally, a number of studies argue that networking and collaborating is beneficial to both men and women, hence reinforcing the visibility argument evoked above. As Copenheaven et al. (2010) suggests, collaborating with male co-authors brings the work of female co-authors to their attention. The fact that most papers are now written in collaboration may contribute to reducing the gender differences in citations. Long (1990) argued that opportunities for women to collaborate were significantly less than those of men when women have young children. Kyvik and Teigen (1996) further added that childcare and lack of research collaboration were the main obstacles to increasing productivity. We would therefore expect that because women work in smaller and more localised teams their research may have a lesser impact. Our last hypothesis is therefore:

H5a Academics who collaborate with a greater number of scientists will be more cited than their colleagues who publish with a similar number of co-authors.

H5b Gender will have a **negative moderating effect** on the relationship between scientific impact and the size of the co-authoring team, i.e. female academics who collaborate with a greater number of scientists will be less cited than their male colleagues who publish with a similar number of co-authors.

Our five hypotheses can be summarised in Figure 1. The direct effects for hypotheses 1, 2a, 3a, 4a and 5a are represented by solid lines while the gender moderating effects of hypotheses 2b, 3b, 4b and 5b are represented by dashed lines.

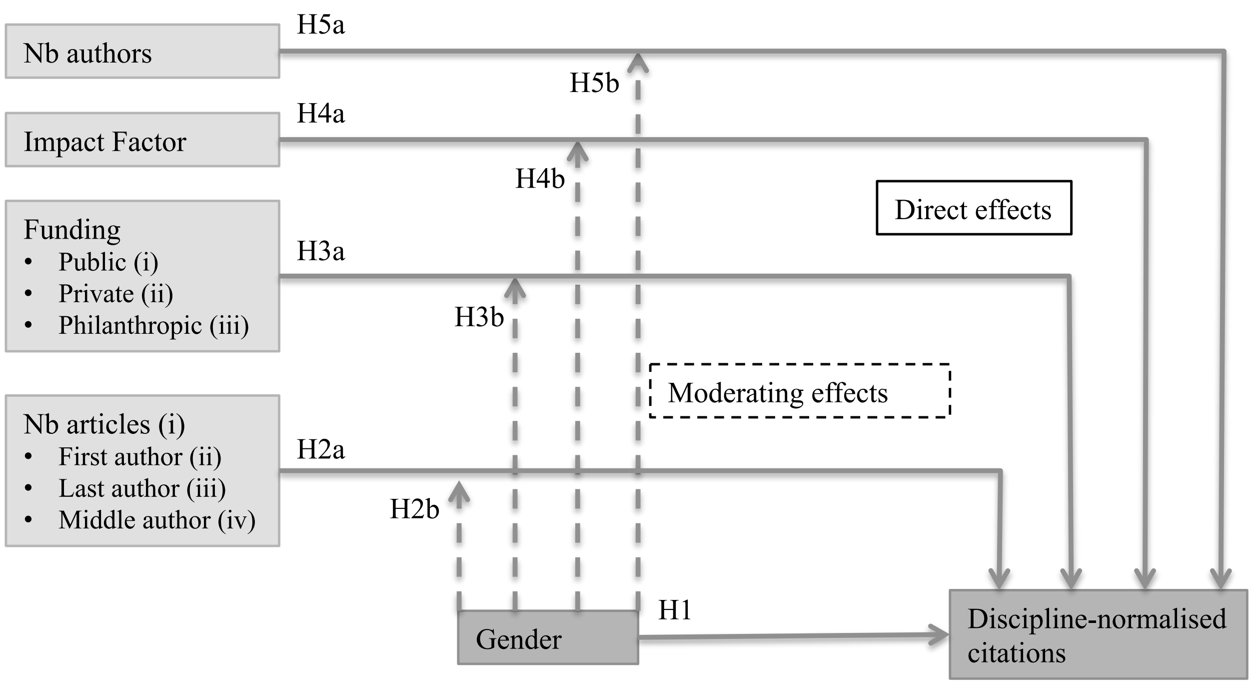


Figure 1 – Relationships between the hypotheses: Direct and gender-moderating effects between scientific impact (discipline-normalised citations) and the four categories of indicators

At this point it is important to emphasise that it is not individual researchers who are cited but the individual publications, which are the result of the efforts of teams of researchers. Most articles published by women are also published with a number of male colleagues and/or male students. It is therefore entirely possible that the number of citations obtained by an article be driven by the notoriety of only one of the authors of the team, which then benefits the entire co-authoring group. In light of this, citation analysis performed at the individual level is always based on the overall publication record of the individual. What this paper examines is whether the publications of women in a particular year, to which men have also contributed, have a higher impact than those of men, to which some women may also have contributed. In other words, we do not compare single gender articles.

3 Data and methodology

Two data sources are required for this study: data on scientific output and on funding. The first source of information is the Thomson Reuters Web of Science (WoS) database that lists scientific publications of a widely recognized set of journals (about 12,000 in 2013). For the second source of information, we are fortunate in Quebec to have access to a very comprehensive database of university funding, the University Research Information System ("*Système d'information sur la recherche universitaire*" or SIRU). This database provides information on all university accounts held by academics in the province on a yearly basis. As each project is attributed a different university account, we are able to distinguish grants from contracts, public funding from private funding, operational costs from infrastructure costs, provincial and Canadian sources from foreign sources, and so on. In addition, all interuniversity transfers are accounted for, which implies that collaborative grants are divided into real amounts (as opposed to averages based on the total amounts divided by the number of co-PIs) according to the funds that were truly transferred from one institution to another. The only drawback so far in the database stems from the fact that we are not yet able to identify the principal investigator (PI) for each grant and from the assumption that we make that the amount held in each university account is divided equally between the co-applicants listed for each account held in the same university. In other words, we are not able to distinguish 'within' university transfers, as such mechanisms do not exist. In addition, we cannot directly attribute each grant to specific articles, hence the need to aggregate the yearly information at the individual scientist level.

The *Observatoire des sciences et des technologies* (OST) in Quebec has disambiguated and uniquely identified every academic in Quebec and provides a comprehensive database of their scientific output and funding (see Larivière et al., 2011). The traditional homonymy and synonymy problems that normally plague all bibliometric databases have thus been resolved prior to us gaining access to the data and have given rise to a vast number of publications in bibliometrics and scientometrics.

With these data, we are able to construct a number of variables to characterize scientific output and research funding for the period 2000-2012. In a manner similar to Nakhaie (2002), we adopt a short-term input-output framework, as opposed to a total career output framework, and will account for changes over the years using panel data, hence all variables are measured on a yearly basis. We also follow his recommendation to the effect that "one has to include a large number of the covariates in a multivariate analysis in order to fully account for gender differences in publication" (2002: 156). Our dependent variable counts the number of citations up to 10 years following the publication year of each article relative to the average citation rate of the papers published worldwide in the same discipline during the same year (**normCit10**[2]). For this calculation, the US National Science Foundation classification of journals into 143 disciplines and specialties is used. This normalised measure allows the comparison between disciplines without having to introduce dummy variables for each of the disciplines if we were to simply count the raw number of citations per article (**nbCit10**) or even the fractional number of citations (i.e. divided by the number of authors – **fracCit10**). We have nonetheless estimated our models on all three dependent variables for the sake of comparativeness with prior research[3].

The variable of interest is obviously the gender of the scientist (H1), which we model using a dummy variable (**dFemale**) taking the value 1 if the scientist is a woman and 0 otherwise. Because part of our story also relates to the critical mass of women in the field, we also add to the models the proportion of female authors per article (**shareWomen**). This variable is available on a wide scale from 2006 onwards because of a change in the reporting of author's names in the WoS[4].

Because a more prolific author is quite likely to have more visibility (H2), we add the number of articles published in a given year (**nbArticles**)[5] as an explanatory variable. To account for the fact that the order of the author list may provide a better reflection of the importance of each author, we propose to use the number of articles as first author (**nbArtFirst**), last author (**nbArtLast**) and as middle author (**nbArtMiddle**) as an alternative to the simple counting of articles regardless of author ranking[6]. Single-author articles are counted solely as first-author articles, two-author articles are counted as one first author and one last author so as not to overinflate the publication rates of individuals with small authorship papers.

To take into consideration the importance of collaboration (H5), we also add the number of authors per paper and average the value per researcher per year (**avgAuthors**), i.e. over all the papers published by an individual in a given year. The reason for introducing such a measure is two-fold: first, it gives an idea of the underlying collaboration necessary to produce scientific

[2] Regressions with **normCit10** suffered from a strong size effect that we have corrected by taking the natural logarithm of the variable. The regressions will therefore be estimated on ln(**normCit10**).
[3] Only the regression results for the discipline-normalised citations will be presented in the paper. A short discussion regarding the results obtained by using variations of the dependent variable is provided in Appendix C.
[4] We were able to accurately assess the gender of authors for about 36% of the articles listed in our database for the years 2006-2012.
[5] In order to take into account co-authorship, it is common practice in bibliometrics to fractionally count the number of papers of an individual or, in other words, to count the number of papers divided by the number of contributors on the author list (**fracArticles**). For instance, a four-author publication counts for 0.25 article for each author. Although this yields a more accurate picture of the workload or productivity of an individual scientist, it somewhat dilutes the visibility that accrues with a greater number of publications and with a greater number of co-authors, which in turn should both influence the citation rate. To disentangle the two effects (and avoid the obvious influence of the number of authors on the fractional count of articles) necessitates the use of the raw number of articles – rather than the fractional count – and of the number of authors as explanatory variables.
[6] To test what the concerns raised in the previous footnote, we have also estimated the regressions using the fractional number of articles published in a given year (**fracArticles**) as an explanatory variable, in addition to the fractional number of first-author articles (**fracArtFirst**), of last-author articles (**fracArtLast**) and of middle-author articles (**fracArtMiddle**). As anticipated, the results obtained are much less significant and robust than to those presented in the paper.

articles and second, more authors evolving in one's own scientific networks should provide a greater visibility to an article, which may yield a greater number of citations[7].

In addition, the prestige of specific journals may induce a greater visibility and yield a greater number of citations. We account for the "quality" of the journal by introducing the 5-year Impact Factor of the journal in which an individual has published a specific paper in a given year (H4), averaging over all the papers published by an individual in that year (**ImpactFact5**)[8].

In terms of funding variables (H3), we classify each funded project according to the sources of its funding: public, private and what can be construed as philanthropic or not-for-profit funding. For each of these categories, we distinguish the funds dedicated to operational costs from those aimed at infrastructure and therefore not directly contributing on a regular basis to the production of scientific articles (infrastructure grants are often one-off funds). In order to smooth out any sudden rise in funding from a given category, we calculate a three-year moving average of the amount of public funding for operational costs (**avgPubFundO3**), of the amount of private funding for operational costs (**avgPrivFundO3**) and of the amount of philanthropic funding for operational costs (**avgPhilFundO3**)[9].

The database is built as an unbalanced panel providing data for the years 2000 to 2012 for each individual scientist. Because our dependent variable has been normalised, and is thus continuous, we can use ordinary least squares regressions for panel data (i.e. the procedure *xtreg* in Stata)[10,11]. In addition, we also estimated our models with the number of articles as a dependent variable, simply to compare the behaviour of Quebec female scientists with that found in the prior literature. We estimate the models using the dependent variable in the simple count format

[7] Similarly, we have tested a second variable counting the number of affiliations listed on the paper averaged per researcher per year (**nbAffiliations**). This second measure can be considered a proxy for inter-institution collaboration. Once again, the rationale is that a greater number of affiliations should provide an increased visibility to a paper. This variable is very highly correlated with the number of authors. For this reason and because the results are very similar to those obtained with the number of authors, they are not be presented in this paper.

[8] The 5-year Impact Factor is generally calculated for the citations received during the five years prior to the current year of publication of the articles examined in our study.

[9] All monetary values have been deflated by the consumer price index and are therefore presented and analysed as constant Canadian dollars of 2002.

[10] During the course of the study, we suspected that our model may suffer from endogeneity due to the fact that scientific production influences the capacity to raise funds and in return, more funds provide greater resources to produce more scientific papers. To correct for potential endogeneity, we used instrumental variables and instrumented for the average amount of public funding (and hence used the procedure *xtivreg* in Stata). Endogeneity tests performed in non-panel regressions accounting for the non-independence of observations for the same individual (*estat endog* in Stata) and Hausman specification tests consistently rejected that there was presence of endogeneity in our models (even though the instruments were valid – verified by the *xtoverid* procedure). In addition, we also tested the use of Hausman–Taylor estimators for error-components models (*xthtaylor* in Stata) that allow estimating models where covariates are correlated with the unobserved individual-level random effects. These models allow the inclusion of both time varying and time invariant endogenous regressors. Unfortunately, the models were plagued with weak instruments that preclude their usage (they were not correlated enough with the other explanatory variables nor with the error term of the regression). The results presented in this paper will therefore not need to account for possible endogeneity.

[11] We also compare the regression results using two commonly used dependent variables, the number of citations (**nbCit10**) which requires the use of Poisson or negative binomial regression models (because of over-dispersion, we used negative binomial regressions with the procedure *xtnbreg*) and the natural logarithm of the fractional citation count (**fracCit10**), which can also be estimated using OLS for panel data. These regression results are discussed in Appendix C.

(**nbArticles**). For this purpose, we use negative binomial regressions on the number of articles using *xtnbreg*[12] for panel data.

Because individual-level effects are not appropriately modeled by random-effect models, all our regressions were estimated using fixed-effects models[13], which preclude the use of gender (**dFemale**) as a stand alone explanatory variable, a constraint the we circumvented by interacting gender with all other explanatory variables to try to disentangle the moderating effect of gender on the relationship between the explanatory variables and the discipline-normalised citation rate[14]. This comment is valid for all the regression results presented in this paper.

Finally, it has been shown that women often work in universities with a lesser research intensity (Sonnert and Holton, 1995; Xie and Shauman, 1998). And when they work in universities with high research intensity, women occupy lower academic ranks than men (Fox, 1991; Leahey, 2007; Sonnert and Holton, 1995). We therefore expect the university environment to have an impact on the citation rate[15]. To account for any time effects, we also add year dummy variables (d2001 to d2012).

4 Descriptive statistics

Once the observations for which one of the variables (with the exception of **shareWomen**) is missing are removed, our sample comprises 1,578 Health scientists and 1,734 NSE scientists over a period of 12 years (corresponding to 10,258 and 10,630 observations respectively), of which 473 are female health scientists (resulting in 2,749 observations) and 256 are female NSE scientists (corresponding to 1,558 observations)[16]. In our sample a greater proportion of women work in the health fields (26.8%) than in the natural sciences and engineering (14.7%). The descriptive statistics of the sample are presented in Table A1 to Table A4 (in Appendix A).

Comparing the overall characteristics of men and women, we find that men contribute to papers that are relatively more cited, hence validating our first hypothesis, produce more papers, occupy more often the last-author rank and the middle-author rank, target higher Impact Factor journals and raise more funds from public, private and philanthropic sources. Mean comparison tests show that the differences between men and women are all statistically significant with the exception of philanthropic funding and Impact Factor in the NSE fields. In contrast, women are more often first author on their papers (only statistically significant in the health fields) but obtain lower relative citation counts. One would not be inclined to think that women aim for quality and not necessarily quantity (Duch et al., 2012). Otherwise, these results are very much

[12] Appendix C discusses the results estimated OLS using regressions where the dependent variable is in fractional count format (**fracArticles**).
[13] Hausman specification tests systematically rejected the random effect models in favour of the fixed effect models.
[14] To lessen the impact of multicollinearity when using interactive variables, we have standardized the variables using the z-score. We have also considered centering using the grand mean and the mean per cluster/group of departments and the results are similar.
[15] In order to account for the university influence, we introduced university dummy variables for each Quebec university and then interacted them with gender. While some of the university dummy variables were significant, none of the interactive variables showed any significance. Furthermore, as our models were estimated using fixed effects rather than random effects, these university dummy variables had to be omitted from the models as too few academics changed university affiliation throughout the period examined.
[16] Given the well-known limitations of bibliometrics for the analysis of the social sciences and the humanities (Archambault et al., 2006; Larivière et al., 2006), these were excluded from the analysis.

in line with most of the literature on women in academia (see country-level data in Larivière et al., 2013, as well as results on authorship position found in Larivière et al., 2011).

The Impact Factor difference between men and women is only statistically significant in the health fields, presumably because women are too few, proportionately, in the NSE fields. As most of these articles are written collaboratively, one would not necessarily find many articles written solely by women[17]. Furthermore, the proportion of women amongst the list of authors in the health fields has increased from about 38% in 2006 to slightly more than 40% towards the end of our sample while in the NSE fields, the proportion of women went from less than 22% to more than 25% over the same period. Over a slightly shorter period, the number of female professors in Quebec universities increased from 30.3% in 2006 to 33.0% in 2009 (Naudillon and Noureddine, 2013), an augmentation that roughly matches their importance amongst author lists.

The funding gap observed between men and women in the health fields regarding the average amount of public funding raised by Quebec academics is rather large, more than 35,000$ per annum on average. In contrast, the difference between men and women in the NSE fields is considerably less, about 12,000$ per annum in favour of men. This alone cannot fully explain why women generate fewer citations than their male colleagues. The private sector may contribute to this discrepancy since women raise much smaller amounts of funds from private sources: The amount raised by women is reduced fourfold in the health fields and almost twofold in the NSE fields. The success rates of funding applications may have a role to play in explaining gender disparities in science. We contacted the principal federal and provincial granting councils to obtain the success rates of men and women in various program competitions. At the provincial level, the Fonds de recherche du Québec – Santé (Health) as well as Nature et Technologies (NSE) report a difference in success rates in favour of men (3,45% more for men the health fields and 7.15% in the NSE fields for the last two years available), but the difference is not significant at the 5% level mainly because of the small sample size. At the federal level, i.e. for the entire country, the Natural Science and Engineering Research Council of Canada (NSERC) reports significant differences in success rates in favour of men (2.9% for discovery grants and 1.2% for strategic projects). In contrast, for the Canadian Institute of Health research (CIHR) – the Canadian health research council – the success rate of awards is in favour of women (1.89% more women succeed in raising funds)[18].

None of the abovementioned factors can uniquely explain the poor performance, comparatively, of female health and NSE scientists. We must therefore turn to regression analyses to take into consideration all these factors together to try to identify the factors that are the most important towards improving one's citation rate and to see how gender moderate these relationships.

[17] In our sample, only about 6.9% of articles are published by a female-only team in the health fields, while in the NSE fields, this proportion drops to about 5.2%.

[18] We used proportion comparison tests on the average annual success rates to various programs graciously provided by the Fonds de recherche du Québec – Santé (FRQS), the Fonds de recherche du Québec – Nature et Technologies (FRQNT), the Natural Science and Engineering Research Council of Canada (NSERC) and the Canadian Institutes of Health Research (CIHR). Note that it was not possible for the granting councils to determine whether the difference between the funding demanded and granted was larger for men or for women; as far as we know, this data is not kept in their records.

5 Regression results

5.1 Scientific production

Before presenting the regression results for the impact of various factors on the discipline-normalised citation rate, we first examine whether some of these variables influence the production rate according to gender. Table 1 and Table 2 present the marginal effects of the corresponding regression results found in Appendix D (Table D1 and Table D2, respectively). For each variable, the top part of the table presents the marginal effects of increasing the value of the said variable by one unit (Appendix B explains the calculations of the different marginal effects accounting for the interactive variables). The Men rows present the marginal effects of increasing the explanatory variable by one unit (the units are defined in Appendix B) while keeping all other variables at the mean and **dFemale** at 0. The Women rows present the marginal effects of increasing the explanatory variable by the same unit while keeping all other variables at the mean and **dFemale** at 1. In this latter case, if the interactive term between **dFemale** and the explanatory variable is not significant, the marginal effect for women should not be considered different as that of men. Further down the table (for the Women vs men rows), for each interactive variable of **dFemale** with the other explanatory variables, both tables present the marginal effects of being a woman with respect of being a man (comparing **dFemale** = 1 with **dFemale** = 0) while keeping all other variables constant at the mean. This latter marginal effect, better corresponds to the sign of the coefficient of the interactive variable.

[*Put* Table 1 ***approximately here*]**

[*Put* Table 2 ***approximately here*]**

Not surprisingly, scientists with more public and philanthropic funding publish more, and those who target high Impact Factor journals do to. For instance, targeting journals with an Impact Factor higher than the average by 0.5 is generally associated with 0.07 to 0.09 more articles in the health fields and with roughly 0.04 more articles in the NSE fields. The importance of public funding was to be expected as most of it is targeted at operational costs and dedicated to paying graduate students. Adding 100,000$ per annum to one's public research funding results in about 0.07 more articles for health scientists (column Art-H-1) and about 0.05 more articles for natural scientists and engineers (column Art-NSE-1), or about 0.04 articles for men and 0.16 articles for women in the health fields (see column Art-H-6). For an equal amount of funding, female health scientists publish about 0.36 more articles than their male colleagues. Although the production of women in the health fields is less than that of men, their marginal productivity is approximately four times higher. Complementary to public funding, not-for-profit funding also plays an important role: 30,000$ per annum of additional philanthropic funding generally yields 0.11 to 0.13 more articles for both the health and NSE fields. In the NSE fields however, the only noticeable difference is that women appear more prolific when funded from the private sector: when women benefit from an additional 30,000$ per annum of philanthropic funding, their production increases by about 0.28 articles (column NSE-5); and compared to men with similar amounts of philanthropic funding, women publish 0.39 more articles. Women therefore seem to be able to "produce more with less".

In addition, natural scientists and engineers who collaborate with a larger number of co-authors publish more; an additional co-author resulting in 0.005 more articles on average. In the health sciences, only for women does an additional co-author has an impact, yielding 0.059 more articles on average. For an identical co-authorship size, women will contribute to 0.76 more articles than their male counterparts.

Our results suggest that most of the scientific production differences observed between genders in the literature can be explained by different levels of funding obtained (Fox, 2005; Hesli and Lee, 2011; Kyvik and Teigen, 1996; Long, 1992; Nakhaie, 2002; Prpić 2002; Xie and Shauman, 1998 and 2003; Zuckerman, 1991). Let us now turn to the core of this paper, i.e. the factors that are associated with the "quality" or impact of the scientific production of women.

5.2 Scientific impact of women vs men

Let us now turn to the main subject of this paper, the factors that explain the discipline-normalised number of citations. The descriptive statistics of our sample show that Quebec women generally contribute to papers that are less cited than Quebec men. The various factors that may influence this relative citation rate are examined in the paragraphs below. Among the factors under scrutiny, we examine the amount of research funds raised, thereby allowing more or less research to be performed, the number of publications, as greater visibility may attract more citations, the number of first-, middle- and last-author rankings, hence signalling the 'importance' of an author, the Impact Factor of the journal targeted, and the size of the teams involved. Let us take each category of explanatory variables interacted with the gender variables in turn. The first results are organised in 4 tables: Table 3 presents the marginal effects of the regression results of Table D3 (Appendix D) for the health fields and Table 4 presents the marginal effects of the regression results of Table D4 (Appendix D) for the NSE fields. With the exception of the first column (H-1 and NSE-1), all the other columns include a row for both Men, Women as well as Women vs men[19].

[*Put* Table 3 *approximately here*]

[*Put* Table 4 *approximately here*]

In regards to the discipline-normalised number of citations (**normCit10**), funding plays a much less fundamental role in explaining the number of relative citations than the number of articles. It is not because one receives greater amounts of funding that one will be more cited[20]. Quite the contrary in fact: the coefficient of public funding is negative and significant, but in the health fields only. In this case, an increase of 100,000$ per annum in public funding reduces the number of relative citations obtained by about 0.1 (the marginal effects vary between -0.06 and -0.11). Because the negative effect is weakly significant for the health fields, non significant for the NSE fields, and the interactive term with **dFemale** is not significant, we investigated whether a non-linear effect would better fit the model (in columns H-7sq in Table 3 and NSE-7sq in Table 4). Figure 2 compares the quadratic effect of the average amount of public funding on the discipline-normalised citation rate for both men and women in the health fields. The positive linear effect combined with the negative quadratic effect implies an inverted U-shaped relationship, hence supporting Arora's et al. (1998) decreasing returns to funding on quality-adjusted publications (a mix of number of publications and citations), for men. This suggests that

[19] The marginal effects are presented in the same way as in Table 1 and Table 2.

[20] Our unit of observation is averaged at the scientist-year level, and thus all citation measures are averages for all the articles published by this scientist in a given year.

when health scientists benefit from less than 109,026$ per annum[21] in public funding (which is the case for 60,25% of male health scientists), their relative citation rate increases with more funding, while beyond that threshold, the citation rate slowly declines thereafter. Better-publicly-funded health scientists are thus more cited than their colleagues, but only up to a point. In contrast, for women, the moderating effect of dFemale reverses the inverted U-shaped curve into a U-shaped curve. In this case, the citation rate of women declines with increased funding until they reach an average annual intake of public funding of about 92,581$ per annum (which is the case for 64,10% of female health scientists). Hence the majority of men are located in the increasing part of the curve whereas women reside in the decreasing part of the curve: for equal amounts of public funding in the health fields, women systematically receive fewer citations than their male colleagues. We suspect that this result may be due to seniority effects, which, in turn, might lead to a decrease in visibility for women[22]. As shown by Larivière et al. (2011), female university professors in Québec are, on average, younger than male university professors, which makes them less likely to be heads of laboratories, and, hence, benefit from the larger visibility—and thus citations—associated with such positions.

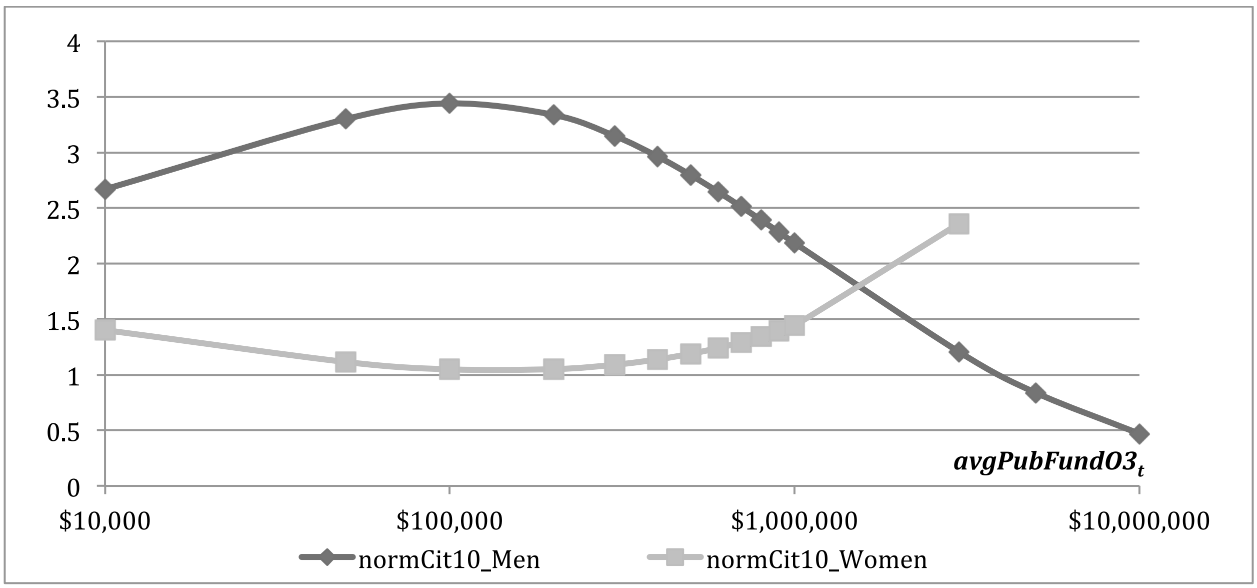


Figure 2 – Predicted discipline-normalised citation according to the quadratic effect of the average public funding for men and women in the health fields (the x axis corresponds to the natural logarithm of the z-score of the average amount of public funding raised over the past three years)

---

[21] To minimise multicollinearity problems, all continuous explanatory variables were transformed using the z-scores and then by the natural logarithm to normalise the variables as much as possible. The formula to calculate the maximum discipline-normalised citation rate according to varying value of public funding (**avgPubFundO3**) is therefore given by $\left(e^{-\beta_1/2\beta_2} - 3\right) \times StdErr + Mean$ where β_1 and β_2 are the coefficients of the linear and quadratic terms respectively. The minimum value of the curve for women is calculated using the same equation where β_3 and β_4, the coefficients of the linear and quadratic interactive terms, are added to β_1 and β_2 respectively $\left(e^{-(\beta_1+\beta_3)/2(\beta_2+\beta_4)} - 3\right) \times StdErr + Mean$.

[22] We included age as an explanatory variable to account for 'seniority' but the coefficients were never significant. We would need information relative to academic rank and regaring directorship of laboratories and research groups to test this suggestion.

At first sight (when interactions with **dFemale** are not included in the models), neither private funding nor philanthropic funding has any impact on the capacity of scientists to attract citations in the health fields. Similarly, none of the funding sources have any incidence on the relative citation rate in the NSE fields. There is however a clear detrimental effect associated with private funding for female health scientists (in columns H-2, H-7 and H-7alt in Table 3). The augmentation of both public and private funding is therefore associated with a negative impact on the relative citation rate of women in the health sciences. In other words, when women in this field raise an equivalent amount of private funding, they are relatively less cited by about 1.16[23] discipline--normalised citations[24]. In addition, raising 30,000$ per annum of additional private funding results in about 0.21 more relative citations for men, but in roughly 0.23 less relative citations for women.

In the NSE fields, only philanthropic funding has a positive effect on female natural scientists and engineers' citations: for the same amount of philanthropic funding, women increase their relative citation rate by about 0.6 compared to men, and if women raise 30,000$ per annum of additional funds from philanthropic sources, their relative citation rate increases by about 0.8 compared to their male colleagues. For men, the marginal effect is non-significant. Raising funds from not-for-profit organisations therefore appears to provide the necessary visibility that women require for their work to get noticed. This might be due to the increasing tendency of women to focus on topics that have a direct impact on society (Cockburn 1988; Collin 1986; Witz 1992), and funding from such organisations might allow female researchers to perform research on topics in which they excel.

What matters more than funding, which after all in not visible outside university accounts with the exception of federal funds, is the visibility granted by the publication of articles (see columns H-4 in Table 3 and NSE-4 in Table 4)[25], one more article published increases the relative citation rate of both men and women[26] by 0.086 in the health fields and by 0.054 in the NSE fields.

Splitting the number of articles according to the rank of each individual in the author list, yields positive and significant coefficients for all three ranks, first-, middle- (only in the NSE fields) and last-author, in the NSE fields (in all columns except H-4 and NSE-4), the first author having the most marginal impact – one more article as first author increasing the number of relative citations by about 0.27 in the health fields and by about 0.25 in the NSE fields[27].

---

[23] Unless otherwise specified, the marginal effects reported in the text are those corresponding to models H-7alt and NSE-7).

[24] This result accounts for differences between disciplines, as our dependent variable is discipline-normalised. In addition, this result is also robust to the exclusion of the two most female dominated disciplines in the health fields: nursing and rehabilitation therapy, which each account for about 300 observations.

[25] Introducing a quadratic term for the number of articles does not yield a significant coefficient suggesting that there is no limit to the impact of publishing. Non-linear effects were systematically tested but were never significant.

[26] All else being equal, women in the health and NSE fields with the same number of publications are equally cited as their male colleagues (the coefficient of the interactive variables between **dFemale** and **nbArticles** is non-significant).

[27] We also added the square of the number of first-, last- and middle-author articles but none of the quadratic terms yielded a significant coefficient. The impact of scientific production on the relative citations is therefore linear (in the log-log transformation). In addition, replacing the numbers of articles according to rank by their proportion in addition to the number of articles in general gave the same results (these results are available in an unpublished appendix available from the authors).

While no significant impact is found for women according to author rank in the NSE fields, in the health fields, occupying the middle rank is clearly detrimental to the relative citation rate of women. Female health scientists who publish more articles as middle-ranked author are relatively less cited than their male colleagues (by about 1.16 relative citations) with an equivalent number of middle-author articles. In the health fields, publishing one more article as middle-author will increase the number of discipline-normalised citations by about 0.15 for men, but only by roughly 0.006 for women.

Publishing in higher Impact Factor journals what that matters most in terms of the number of discipline-normalised citations: an increment of 0.5 in the average Impact Factor of the journals targeted by an individual (and her team) yields approximately 4.7 additional relative citations in the health fields and roughly 0.93 more relative citations in the NSE fields. Further bad news for women health scientists is suggested by the negative coefficient of the interaction between **dFemale** and **ImpactFact5** (columns H-5 and H-7alt in Table 3). It seems that when women publish in journals with similar Impact Factor as men, their relative citation record drops by about 2 compared with their male counterparts. In column H-7 (where both the interactions of gender with the number of authors and with the Impact factor are present), we lose the significance of the interaction between gender and Impact Factor. We have therefore tested the addition of a triple interaction between gender, Impact Factor and the average number of authors per article (not shown in the paper): The two double interactions with gender become positive and significant and the triple interaction becomes negative and significant. In other words, women in the health fields publishing with a similar number of authors and in similar Impact Factor journals are relatively less cited than their male colleagues. In the second part of this section, we address this issue by introducing the proportion of female authors in the model.

In accordance to the rationale that with publishing a greater number of articles an individual scholar is more visible, having a more numerous author list may also improve the visibility of a paper and hence attract more citations. Adding one more co-author to the co-author list, also contributes to improving ones' relative citation record, by about 0.26 in the health fields, but only by about 0.02 in the NSE fields. The important difference noted in the marginal effects between the two fields stems from the fact that the average number of co-authors on NSE publications is generally (and sometimes substantially) larger than in the health fields.

Given the same resources, the same visibility from publishing, targeting the same journals and working with a similar number of co-authors, our results do not give credence to the hypothesis that women scientists and engineers are less capable of publishing high impact research (the only significant interactive variable with **dFemale** is the amount of philanthropic funding and it yields a positive and significant coefficient). The same cannot be said for female health scientists. Women are still a minority in the NSE fields, which suggests that there may be a critical mass effect at play. Because women in the NSE fields represent only about 15% of academics in Quebec compared to roughly 27% in the health fields, we suspect that we are not able to appropriately measure the gender effect in the NSE fields as women are drowned amongst a larger number of male natural scientists and engineers. Moreover, in the health fields, this indicator is not specific enough to measure gender effects and clouds the relationship between gender and the Impact Factor. The proportion of women in the author list may provide a means by which to disentangle the impact of gender from that of the size of the author list.

5.3 Scientific impact according to the proportion of female authors

The introduction of the proportion of women amongst the author list averaged over the number of articles published by an individual in a given year (**shareWomen**) and of its interaction with the other explanatory variables allows to partially account for the composition of the author list. While we do not compare the citations obtained by articles published only by men to those published by an only female authorship, the average proportion of female authors with whom an individual publishes helps getting closer to measuring the collective impact of women on the relative citation rate. Because the proportion of women in the author list is only available from 2006 and only a proportion of the articles list the complete first names, adding this variable to the regression results in a reduction in sample size[28]. The regressions presented in the rest of the article are estimated on this reduced sample.[29]

These results on the 2006-2012 sample are presented in six tables: Table 5 shows the marginal effects of the regression results presented in Table D5, where we simply add the average proportion of women authors (**shareWomen**) in the articles published by an individual in a given year and its interaction with gender for a reduced number of models; Table 6 and Table 7 show the marginal effects of the regression results presented in Table D6 and Table D7 for the health and NSE fields respectively, where **dFemale** is replaced by **shareWomen** as the variable with which all the explanatory variables are interacted.

In Table D5, the coefficient of **shareWomen** is negative and significant while its interaction with **dFemale** is positive, giving an overall positive coefficient for women in the health fields. In terms of marginal effects (see Table 5), male health scientists who co-author articles with 25% more women reduce their relative citation rate by 0.25 (see column H-7b) while female health scientists only reduce their relative citation rate by 0.11. This result suggests that despite the fact that articles published with a higher proportion of female authors are less cited, the effect is less detrimental for women than for their male colleagues. For the same proportion of female co-authors, women will increase their relative citation rate by about 0.01. This suggests that men are more reticent at citing the work of women (Klonbloch-Westerwick and Glynn, 2013). This is not

[28] The introduction of the variable **shareWomen** in the models hence reduces the sample size slightly as well as the number of years available for our regressions. In the health fields, we had 1,578 individuals (10,258 observations) for the period 2000-2012, for the period 2006-2012, the sample is reduced to 1,482 individuals (6,420 observations) which is further reduced to 1,462 individuals (6,075 observations) when **shareWomen** is added to the sample. In the NSE fields, the 2000-2012 sample comprises of 1,734 individuals (10,630 observations), for the period 2006-2012, the sample is reduced to 1,615 individuals (6,657 observations), which is further reduced to 1,524 individuals (5,798 observations) when **shareWomen** is added to the sample.

[29] We performed a series of tests to compare the results obtained from estimating the model on the entire sample (2000-2012), and on two reduced samples for the first (2000-2006) and second (2006-2012) half periods. These tests were performed with the *suest* procedure in Stata using simple OLS regressions while controlling for the non-independence of the observations across the same individual (*suest* does not allow panel regressions). We find no significant difference between the regressions estimated on the first half period (2000-2006) and on the entire sample (2000-2012). Although we find a significant difference between the regression results obtained for the 2000-2012 and 2006-2012 samples, the vast majority of coefficients keep the same sign and only a few change significance level. For instance, in contrast to what was found for the 2000-2012 sample, for the period 2006-2012, we find no effect of public funding for women in both the health and NSE fields (neither the linear nor the quadratic term are significant). For the period 2006-2012, given the same amount of public research funds, women perform as well as men in terms of discipline-normalised citation rate. These results are available from the authors in an unpublished appendix.

the case in the NSE fields, where the proportion of female authors is much smaller (and neither variable is significant).

[*Put* Table 5 *approximately here*]

In this more recent sample, the results are sensibly the same with a few exceptions that we discuss here. Philanthropic funding now reduces a woman's relative citation rate in the health fields (by about 1.1 for an increase of 30,000$ per annum), but increases a woman's relative citations rate in the NSE fields (by about 1.5 for an increase of 30,000$ per annum). In the health fields, the first-author rank now provides an advantage to women over their male colleagues: while men increase their relative citation rate by 0.24 with one more article published as first author, women increase their relative citation rate by 0.33; for a similar production, women increase their relative citation rate by 0.79 compared to their male colleagues. As a comparison, we estimated the same model on the first part of the sample only (2000-2006) and found that the interactive coefficient between gender and the number of first-author articles is not significant (which is congruent with the results of Table 3 for the 2000-2012 sample). This would therefore appear to be a more recent phenomenon. Furthermore, the impact of occupying the middle author is now detrimental to female health scientists, decreasing their relative citation rate by 0.016 for an increase in production of 1 more article as middle author, and decreasing their relative citation rate by 0.91 compared to men with a similar production rate.

Similarly as before, across all fields, targeting journals with higher Impact Factors and collaborating with a greater number of co-authors are the common factors that contribute to improving one's citation rate. Once we account for the proportion of women in the author list of the articles to which an individual contributes, we lose the significance of both the interactive terms between gender and Impact Factor as well as the number of authors. This would tend to show that when women publish in the same journals, they get the same level of citations. The same can be said for the average number of authors per article – no gender difference is observed in this regard. For the early part of the sample (2000-2006), we obtain similar results as for the entire sample. The proportion of female authors appears to take the weight of these two variables and to moderate their relationship with the dependent variable.

Let us now turn to the last set of marginal effects presented in Table 6 and Table 7 (the regressions results are presented in Table D6 and Table D7), where the gender interactive term is replaced by the proportion of female co-authors. Health scientists who publish with a higher number (proportionately) of women, receive more philanthropic funding or are ranked amongst the middle authors, are less cited than average (see the coefficients of the interactive variables between **shareWomen** and these two indicators in Table D6)[30]. This is reminiscent of the results presented above.

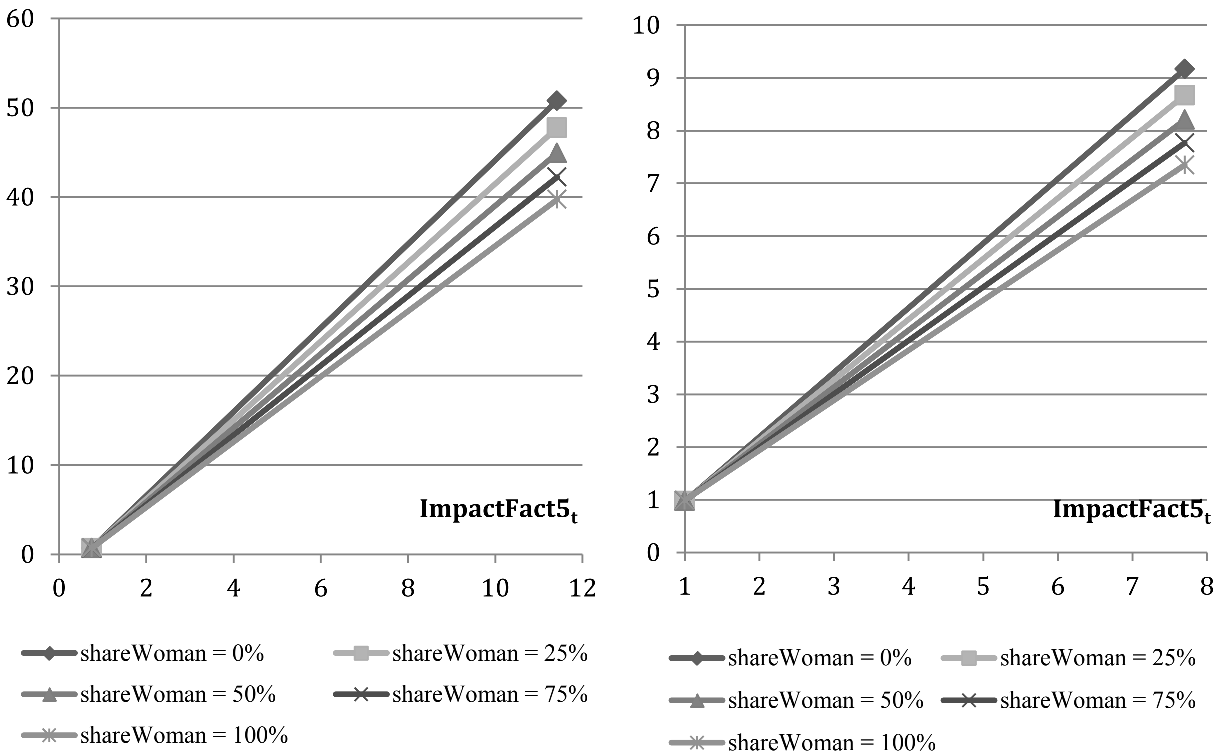
[*Put* Table 6 *approximately here*]

[*Put* Table 7 *approximately here*]

The influence of the five-year Impact Factor of journals is where the role of the proportion of women amongst the author list highlights a worrying trend. Not only does it matter to be a woman (and indeed the interactive term between **dFemale** and **ImpactFact5** was significant in Table 3) but to be surrounded by a higher proportion of women authors also has a detrimental

[30] All interactions between continuous variables were systematically examined graphically to ensure that the tendencies reported are correct. These graphs are not reported here as they would not add to the narrative.

effect on one's relative citation rate. For a similar Impact Factor, researchers who co-author articles with a higher proportion of women (larger **shareWomen**) are less cited than individuals who generally collaborate with a smaller proportion of women. This is observed for both the health and NSE fields. Figure 3 illustrates the predicted value of the discipline-normalised citation rates according to the Impact Factor of the journals targeted for various proportions of female co-authors. For both the health and NSE fields, the graphs clearly show the declining number of relative citations as the proportion of female co-authors increases while keeping the Impact Factor constant. The overall coefficient is still positive (the sum of both coefficients), suggesting that articles written by a greater proportion of women benefit less from the visibility provided by high impact journals[31]; a finding that can be interpreted in the context of the Matilda effect (Rossiter, 1993). Either women are systematically discriminated against, or they concentrate on different types of research that is less cited. The preference of women for less specialised research, which is often more difficult to publish, is addressed by Rhoten and Pfirman (2007) and by Leahey (2006).



(a) **(b)**

Figure 3 – Predicted values of the discipline-normalised citation rate according to Impact Factor for varying proportions of the number of women co-authors for (a) the health fields and (b) the NSE fields

[31] To investigate whether this could be due to the inclusion of nursing, or rehabilitation therapy, disciplines dominated by female scientists, we removed these disciplines from the regressions, but the results remains the same. We also interacted the variables **shareWomen**, ImpactFact5 and dNursing/dRehabilitation and the results are also congruent.

6 Discussion

At the beginning of this paper, we set out to validate five hypotheses on the gender gap in science. Here, we discuss each of these hypotheses prior to presenting the general conclusion. Table 8 presents a summary of the validation of each hypothesis for both fields. The first hypothesis was simply the observation that Quebec academic women are statistically less cited then their male counterparts. H1 is validated for both the health and NSE fields.

Table 8 – Summary of the validation of the hypotheses

	Hypothesis	Health		NSE	
		Men (a)	**Women (b) – Gender moderating effect)**	**Men (a)**	**Women (b) – Gender moderating effect)**
H1	Gender		**yes (negative impact)**		**yes (negative impact)**
H2	i – articles	**yes (positive)**	no (N.S.)	**yes (positive)**	no (N.S.)
	ii – First-author articles	**yes (positive)**	no (N.S.)	**yes (positive)**	no (N.S.)
	iii – Last-author articles	no (N.S.)	no (N.S.)	**yes (positive)**	no (N.S.)
	iv – Middle-author articles	**yes (positive)**	**yes (negative)**	**yes (positive)**	no (N.S.)
H3	i – Public funding	**~ (Inverted-U)**	**~ (U-shaped)**	no (N.S.)	no (N.S.)
	ii – Private funding	**yes (positive)**	**yes (negative)**	no (N.S.)	no (N.S.)
	iii – Philanthropic funding	no (N.S.)	no (N.S.)	no (N.S.)	**no (positive)**
H4	Impact Factor	**yes (positive)**	**yes (negative)**	**yes (positive)**	no (N.S.)
H5	Number of co-authors	**yes (positive)**	**yes (negative)**	**yes (positive)**	no (N.S.)

The second hypothesis (H2a) aimed to validate the argument that the accrued visibility ensuing from a greater number of publications is associated with a higher citation rate. H2a is generally validated with the exception of H2a(iii) for the health fields. We had anticipated that gender would have a negative moderating effect on this positive relationship (H2b) in that papers from more productive female scientists should receive fewer citations and because women publish less, they should be less cited as well. There is a mitigated support for this hypothesis: only H2b(iv) is validated for health fields as our results show that women who publish a greater number of articles as middle-authors are less cited than their male colleagues. For other author rankings, women and men perform equally well, thereby refuting most of hypothesis H2b(i, ii, and iii). Given the same level of scientific production and visibility as first, last or middle author, NSE women receive a similar relative number of citations. In the second half of the period examined (2006-2012), we also show that women who publish a higher number of first-author articles obtain more citations in the health fields. These results find echo in the findings of Housri et al. (2008). Although we cannot directly prove this claim, we suspect that part of the explanation for this result is that female PhD students (who later enter academia) who publish with their supervisor will do so as first author and thus benefit from the notoriety of their supervisor. The publication norms in the laboratory-based fields of health and NSE are typically to have the head of the group as last author, and the researcher responsible for most of the work as the first author (Pontille, 2004), these observations suggest that lab directors and to some extent PhD supervisors more often occupy the last author rank. In Quebec, the ratio graduate student to university professor has increased by 29% between 1998 and 2010 (FQPPU, 2015). Furthermore, in 2008, 47% of doctoral students in Quebec were women, a proportion that has been steadily augmenting for a number of years now. As a consequence, the augmentation of graduate students by about 50% (a great proportion of which are women) over the period examined linked to the fact that PhD supervisors are more often than not the last authors, contributes to explaining this positive result for women first authors in more recent years.

Our third hypothesis assessed the influence of funding. In the health fields, public and private funds appear to have an impact on the propensity to be more cited than average, while philanthropic funding has no impact on the relative citation rate. For public funding, the relationship follows and inverted U-shaped curve which imply that for about 60% of male health scientists (located along the increasing part of the curve), H3a(i) is validated whereas for the other 40%, this hypothesis is rejected. H3b(i) is also validated for 64% of female health scientists who are located on the decreasing part of the U-shaped curve representing the relationship between the discipline-normalised citation rate and public funding, and rejected for the other 36%. Similarly, in terms of impact on their relative citation rate, male health scientists benefit from more private funding while female health scientists who receive equal amounts of private funding as their male counterparts obtain fewer relative citations, thus supporting H3a(ii) and H3b(ii). In contrast, female natural scientists and engineers who raise the same amount of not-for-profit funds improve their normalised citation record compared with men – H3b(iii) predicted the opposite. No other impact of funding is observed in the NSE fields. Considering that women are slightly less successful than men in raising public funds in the NSE fields, the absence of a difference in relative citation rates in the NSE fields would tend to indicate that women somehow make up for their smaller amounts of public funds. An alternative explanation is that because citations are difficult to attribute to an individual, the small number of female academics in the NSE fields and the ensuing small proportion of women in the author list of articles cloud our measures of impact (in comparison to more than a quarter of Quebec academics being women in the health fields).

Our fourth hypothesis examined the influence of the Impact Factor of journals in which scientists publish. The average 5-year Impact Factor of journals has a direct effect on the citation rate of individuals who publish in those journals, hence validating H4a. Contrarily to all expectations, however, it is not in the NSE fields—where they account for a very small proportion of the researchers—that women are less cited given an equal Impact Factor of the journal, but in health fields. Our results cannot validate or refute H4b in the NSE fields but clearly reject the positive moderating effect of gender in the health fields. After the social sciences and humanities fields, the health fields are where women are the most present. Even when removing from the analysis the disciplines traditionally occupied by women, such as nursing and rehabilitation therapy disciplines, the results are similar. Is it possible that in promoting women in science for a great number of years now, we have neglected women in the health sciences? Further reflection is needed regarding the state of female health scientists and more importantly, the perception that their male colleagues have of their work. In this respect, it might be interesting to look at the provenance of citations: are female researchers more likely to cite other female colleagues' work, or are referencing practices the same for both genders?

Our last hypothesis examined the teams with which scientists publish. Once again, our results echo the general wisdom that dictates that a wider visibility provided by a larger author base has a positive impact on the propensity to attract citations, and thus validating H5a. While the picture is similar for both men and women in the NSE fields (refuting H5b), for women in the health fields, the impact of a larger team is less than that of their male colleagues, hence validating H5b for the health fields. It would therefore appear that collaboration, in the health fields, remains an obstacle for women (Kyvik and Teigen, 1996). International collaboration, as shown by Larivière et al. (2013), is likely to play a role here.

7 Conclusion

This study therefore raises the question of the impact of a critical mass within a discipline or a field. We compared a domain where women account for 14.7% of researchers (NSE) with one where they account for 26.8% (health). With such a small proportion of women in the former, it is difficult to disentangle the role of each individual author in attracting citations to a paper. In the health fields, because women account for a comparatively larger proportion of researchers, we suspected that their lower relative citation rate was possibly due to gender discrimination.

Our study has compiled the proportion of female scientists per article for about a third of the articles to which Quebec scientists have contributed. With this information, we have shown what we could not measure at the individual level, i.e. that articles (on average and at the individual level) with a greater proportion of female authors get less cited than articles with a male dominated author list, controlling for the amount of funding received, the number of articles published by an individual, the Impact Factor of the journal, and so on. For the health fields, we replicated the results that were highlighted with gender taken at the individual level: a larger proportion of women in the author list exhibits a negative moderating effect implying that a greater amount of philanthropic funding, publishing a greater number of articles as a middle-author, targeting higher Impact Factor journals and publishing with a greater number of co-authors all reduce the relative citation rate of articles with a larger share of female authors. These results thus confirm our previous results: women in the health fields seem to be discriminated against in terms of discipline-normalised citation rates, and in addition, papers that have greater proportion of women as co-authors are less cited. This latter systematic gender bias is observed for both the health and NSE fields.

This is undoubtedly the most important finding of this paper: Once we have accounted for differences in the amount of funding at the disposal of individual scientists, the number of articles they publish, the Impact Factor of the journals they target and the specific number of co-authors with whom they collaborate (which have all lower values for women), writing with a higher proportion of female co-authors systematically decreases one’s relative citation rate. For the top cited journals in the health fields, going from an all male team to an all female team reduces the number of relative citations by more than 10 and by about 2 in the NSE fields – these are non-negligible numbers because they are discipline normalised rather than raw citation counts. When we embarked on this research, we were convinced that we would find that once we account for most of the gender differences mentioned above, we would find that women perform equally well. We did not anticipate such a strong bias against articles co-authored by a larger proportion of women. In light of the evidence presented in the article, science policy should continue to encourage women to enter a research career, without achieving a more gender balanced research environment, these gender biases will remain. In addition, more research is required to truly measure the societal impact of one’s research.

A number of other avenues for future research are open. While our results hint at possible gender discrimination, to be certain, one would have to rule out other factors that may result from conscious choices by women themselves. One such factor is the greater tendency for women to focus on less specialised and possibly more multidisciplinary research. It has been suggested by Leahey (2006) that women specialise less than men and that this fact hinders their capacity to get published and cited. We have nonetheless shown that men and women that are equally prolific receive similar relative citation rates. This specialisation argument may have repercussions on the choice of collaborators and on the constitution of research teams.

Introducing social network analysis indicators in the regressions to provide a richer analysis of the structure/network of collaborations is an obvious avenue to pursue. For instance, Maliniak et al. (2013) showed that articles co-authored by women are less central that those of men. Beaudry and Allaoui (2012) found a strong effect of the position of individual researchers in the co-publication network. Women, who often devote more time to teaching and administrative duties than men (Barzebat, 2006; Bellas and Toutkoushian, 1999; DesRoches et al., 2010; Xie et Shauman, 1998) to the detriment of research activities, may have less time to devote to maintaining the necessary links of an efficient collaborative team, and may as a consequence occupy more peripheral positions within the collaboration network.

With the increasing importance of altmetrics (Bowman, 2015; Haustein et al., 2016; Piwowar, 2013; Priem, 2014) and of the need to market one's research on the conference circuit and via social media (Van Noorden, 2014), time and the disposable income to go to conferences, to pay someone to feed information towards increasing one's social media visibility and to some extent to do research marketing may be more limited to women. Women may not have shown as much interest in this regard as their male colleagues. In addition to the methodologies developed to measure a researcher's contribution to social media, future research could keep track of the number of versions of the same paper, using Google scholar for instance, to assess the extent to which the research has been presented at conferences and thus reached a wider audience.

In addition to the 'choice' indicators highlighted in the previous paragraphs (type of research, choice of collaborators, research marketing, and so on), more research is needed to understand the reasons that explain why (Quebec) women are less successful at raising public research funds in the NSE fields, more successful in the heath fields but raising smaller amounts, while they contribute to research that is less cited in both fields. In the natural sciences and engineering, our study clearly shows that given the same opportunities, women perform equally well as men, both in terms of scientific production and impact. As we did with the proportion of female co-authors, it would be interesting to study the impact of the collective fundraising performance of research teams behind scientific articles. With funding data increasingly available in the acknowledgements section of WoS-indexed papers, this type of research may soon become a reality.

This research has a number of limitations, the most obvious being the sample chosen. Larivière et al. (2013) mentioned Quebec as one of the North-American regions closest to achieving gender parity in science. The picture presented in this paper may not reflect the realities of other regions or countries. That being said, given that our analysis solely focuses on 'internationalised' research fields (NSE and health), which are not affected by language-bias (Archambault et al., 2006), findings of this analysis are likely to be relevant for the rest of Canada, as well as to other developed countries. The second is the fact that not all Québec academics are included in the analysis, those for which we could not find the gender were excluded from the study. It is however quite unlikely that those for which we did not have the gender behaved differently from those for which we have this information. Third, we could only find the gender information for a portion of the articles to which Quebec scientists have contributed and this reduced our sample size by half the number of years. In future research, having a more accurate picture would require the disambiguation and gender assignation of all co-authors of the articles to which Quebec scientists have contributed. Fourth, this research is at the confluent of bibliometrics and econometrics, more information on socio-demographic attributes and on the collaborative aspect of science is missing from this study. For instance, there are many other factors that have been shown to affect the gender gap in science; factors which cannot be addressed here. These factors include seniority (Kyvik, 1990; Larivière, 2011), which has been found to have a positive effect on researchers' scientific output and citations[32].

Other factors that are not included in our analysis relate to familial organisation and children. Many studies have shown that women perform a greater share of household work which, in turn, might leave them with less time to perform research activities (Sax et al., 2002). Other authors also found that having children was associated with a lower productivity (Long, 1990, Kyvik, 1990; Kyvik and Teigen, 1996; Stack, 2004). In contrast, Fox (2005) has provided a more nuanced portrait of the situation: women with older children, as well as women with pre-school children, were almost as productive as men, while women with elementary and secondary age children formed the most unproductive group. Along these lines, marriage was found to be positively linked with productivity (Long and Fox, 1995; Xie and Shauman, 1998). Access to top positions can also be linked to such factors. For instance, Ginther and Kahn (2006) provided evidence that the lower proportion of women in tenure track positions was entirely explained by fertility decisions, and that differences in academic promotion (tenure or full professorship) disappeared once one controls demographic, family, employer and productivity variables. Further research should therefore combine survey data with bibliometric methods across a spectrum of fields, countries, and institutions, in order to obtain a more precise portrait of the situation and assess the factors that are most likely to affect the trends observed. Only this systemic mixed method analysis of gender discrepancies will be able to truly disentangle gender discrimination (in access to funding, to journals, to PhD students, etc.) from choices (related to career, children, type of research, research marketing, multidisciplinarity, etc.).

8 References

Abramo, G., D'Angelo, C. A., Di Costa, F., 2009. Research collaboration and productivity: is there correlation? Higher Education 57(2), 155–171.

Aksnes, D. W., Rorstad, K., Piro, F., Sivertsen, G., 2011. Are female researchers less cited? A large-scale study of Norwegian scientists. Journal of the American Society for Information Science and Technology 62(4), 628–636.

Archambault, É., Vignola-Gagne, É., Côté, G., Larivière, V., Gingras, Y., 2006. Benchmarking scientific output in the social sciences and humanities: The limits of existing databases. Scientometrics 68(3), 329–342.

Arora, A., David, P. A., Gambardella, A., 1998. Reputation and competence in publicly funded science: estimating the effects on research group productivity. Annales d'Economie et de Statistique 49/50, 163-198.

Barzebat, D. A., 2006. Gender differences in research patterns among PhD economists. Journal of Economic Education 37(3), 359–375.

[32] Cole and Zuckerman (1984) as well as Cole (1979) showed that age has only a minor influence on scientific performance, once other variables, such as rank and type of institution, are controlled for.

Beaudry, C., Allaoui, S., 2012. Impact of public and private research funding on scientific production: The case of nanotechnology. Research Policy 41(9), 1589–1606.

Bellas, M. L., Toutkoushian, R. K., 1999. Faculty time allocations and research productivity: Gender, race and family effects. The Review of Higher Education 22(4), 367–390.

Bordons, M., Morillo, F., Fernandez, M. T., 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(2), 159–173.

Bowman, T. D. (2015, July). Investigating the use of affordances and framing techniques by scholars to manage personal and professional impressions on Twitter (Dissertation). Indiana University, Bloomington, IN, USA.

Cole J.R. 1979. Fair Science : Women in the scientific community. New York : Columbia University Press.

Cole, J. R., Zuckerman, H. 1984. The productivity puzzle: Persistence and change in patterns of publication of men and women scientists. In P. Maehr and M. W. Steinkamp (Eds.), Advances in motivation and achievement, Vol. 2, pp. 217–258. Greenwich, CT: JAI Press.

Copenheaver, C. A., Goldbeck, K., Cherubini, P., 2010. Lack of gender bias in citation rates of publications by dendrochronologists: What is unique about this discipline? Tree-Ring Research 66(2), 127–133.

Cockburn, C. 1988. Machinery of dominance: Women, men and technical know-how. Boston: Northeastern University Press.

Collin, J. 1986. La dynamique des rapports de sexes à l'université, 1940–1980. Histoire sociale—Social History, 19(38), 365–385.

Davenport, E., Snyder, H., 1995. Who cites women? Whom do women cite?An exploration of gender and scholarly citation in sociology. Journal of Documentation 51(4), 404–410.

DesRoches, C. M., Zinner, D. E., Sowmya, R. R., Iezzoni, L. I., Campbell, E. G., 2010. Activities, productivity, and compensation of men and women in the life sciences. Academic Medicine 85(4), 631–639.

Duch, J., Zeng, X. H. T., Sales-Pardo, M., Radicchi, F., Otis, S., Woodruff, T. K., Nunes Amaral, L.A., 2012. The possible role of resource requirements and academic career-choice risk on gender differences in publication rate and impact. PLoS ONE 7(12), e51332.

Fortin, J.-M., Currie, D. J., 2013. Big Science vs. Little Science: How Scientific Impact Scales with Funding. PLoS ONE 8(6), e65263.

Fox, M. F., 2005. Gender, family characteristics, and publication productivity among scientists. Social Studies of Science 35(1), 131–150.

FQPPU, 2015. État du financement de la recherche universitaire au Québec, Rapport du Comité sur le financement des universities (soon to be published on the http://www.fqppu.org/ website)

Ginther, D. K., Kahn, S., 2006. Does science promote women? Evidence from academia 1973-2001. National Bureau of Economic Research.

Gonzalez-Brambila, C., Veloso F. M., 2007. The determinants of research output and impact: A study of Mexican researchers. Research Policy 36(7), 1035–1051.

Haustein, S., Bowman, T. D., Costas, R. (2016). Interpreting "altmetrics": viewing acts on social media through the lens of citation and social theories. In C. R. Sugimoto (Ed.), Theories of Informetrics and Scholarly Communication. A Festschrift in Honor of Blaise Cronin (pp. 372–405). Berlin: De Gruyter.

Hesli, V.L., Lee, J.M., 2011. Faculty Research Productivity: Why Do Some of Our Colleagues Publish More than Others? PS: Political Science and Politics 44(02), 393–408.

Housri, N., Cheung, M. C., Koniaris, L. G., Zimmers, T. A., 2008. Scientific impact of women in academic surgery. Journal of Surgical Research 148(1), 13–16.

Knobloch-Westerwick, S., Glynn, C. J. (2013). The Matilda effect—Role congruity effects on scholarly communication: A citation analysis of Communication Research and Journal of Communication articles. Communication Research, 40(1), 3–26.

Kyvik S. 1990. Motherhood and Scientific Productivity. Social Studies of Science. 20(1),149-160.

Kyvik, S., 1995. Are big university departments better than small ones? Higher Education 30(3), 295–304.

Kyvik, S., Teigen, M., 1996. Child care, research Collaboration, and gender differences in scientific productivity. Science, Technology and Human Values 21(1), 54–71.

Larivière, V., Archambault, É., Gingras, Y., Vignola-Gagné, É., 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(8), 997–1004.

Larivière, V., Gingras, Y., 2010. The Impact Factor's Matthew effect: a natural experiment in bibliometrics. Journal of the American Society for Information Science and Technology 61(2), 424–427

Larivière, V., Ni, C. Q., Gingras, Y., Cronin, B., Sugimoto, C. R. (2013). Global gender disparities in science. Nature 504(7479), 211–213.

Larivière, V., Vignola-Gagné, É, Villeneuve, C., Gélinas, P., Gingras, Y., 2011. Sex differences in research funding, productivity and impact: an analysis of Québec university professors. Scientometrics 87(3), 483–498.

Leahey, E., 2006. Gender differences in productivity research specialization as a missing link. Gender & Society 20(6), 754–780.

Lewison, G., 2001. The quantity and quality of female researchers: A bibliometric study of Iceland. Scientometrics 52(1), 29–43.

Long, J. S., 1990. The origins of sex differences in science. Social Forces 68(4), 1297–1316.

Long, J. S., 1992. Measures of sex differences in scientific productivity. Social Forces 71(1), 159–178.

Long, J.S., Fox, M.F., 1995. Scientific Careers: Universalism and Particularism. Annual Review of Sociology 21, 45–71.

Maliniak, D., Powers, R., Walter, B. F. (2013). The gender citation gap in international relations. International Organization, 67(04), 889-922.

Mauleón, E., Bordons, M., 2006. Productivity, impact and publication habits by gender in the area of Materials Science. Scientometrics 66(1), 199–218.

Mirnezami, S. R., Beaudry, C., Larivière, V. (2015). What determines researchers' scientific impact? A case study of Quebec researchers. *Science and Public Policy*, scv038.

Nakhaie, M. R., 2002. Gender Differences in Publication among University Professors in Canada. Canadian Review of Sociology/Revue canadienne de sociologie 39(2), 151–179.

Naudillon, F., Noureddine, N., 2013. Femmes et pouvoirs dans les universités québécoises, FQPPU, 182 p.

Peñas, C. S., Willett, P., 2006. Gender differences in publication and citation counts in librarianship and information science research. Journal of Information Science 32(5), 480–485.

Piwowar, H. (2013). Altmetrics: Value all research products. Nature, 493(7431), 159-159.

Pontille, D., 2004. La signature scientifique: Une sociologie pragmatique de l'attribution. Paris : CNRS Éditions.

Priem, J. (2014). Altmetrics. Beyond bibliometrics: harnessing multidimensional indicators of scholarly impact, 263-288.

Prpić K., 2002. Gender and productivity differentials in science. Scientometrics 55(1), 27–58.

Rhoten, D., Pfirman, S. (2007). Women in interdisciplinary science: Exploring preferences and consequences. Research policy, 36(1), 56-75.

Rossiter, M. W. (1993). The Matthew Matilda effect in science. Social Studies of Science 23(2), 325–341.

Sax, L.J., Serra Hagedorn, L., Arredondo, M., and Dicrisi III, F.A. 2002. Faculty Research Productivity : Exploring the Role of Gender and Family-Related Factors. Research in Higher Education. 43(4), 423-446.

Sonnert, G., Holton, G., 1995. Gender differences in science careers: the project access study. New Brunswick, N.J.: Rutgers University Press. 187 pages.

Stack, S., 2004. Gender, Children and Research Productivity. Research in Higher Education 45(8), 891–920.

Summers, L. H., 2005. Remarks at NBER conference on diversifying the science and engineering workforce. January 14, 2005.

Symonds, M. R. E., Gemmell, N. J., Braisher, T. L., Gorringe, K. L., Elgar, M. A., 2006. Gender differences in publication output: Towards an unbiased metric of research performance. PloS One 1(1), e127.

Turner, L., Mairesse, J., 2005. Individual Productivity Differences in Public Research: How important are non-individual determinants? An Econometric Study of French Physicists' publications and citations (1986-1997). Centre National de la Recherche Scientifique.

Van Noorden, R. (2014). Online collaboration: Scientists and the social network. Nature 512(7513), 126-129.

Ward, K. B., Gast, J., Grant, L., 1992. Visibility and Dissemination of Women's and Men's Sociological Scholarship. Social Problems 39(3), 291–298.

Witz, A. 1992. Professions and patriarchy. New York: Routledge.

Xie, Y., Shauman, K. A., 1998. Sex differences in research productivity: New evidence about an old puzzle. American Sociological Review 63(6), 847–870.

Xie Y., Shauman K. A., 2003. Women in science. Career processes and outcomes. Cambridge, MA: Harvard University Press. 318 pages.

Zuckerman H., 1991. The careers of men and women scientists : A review of current research. In Zuckerman H., J. R. Cole and J. T. Bruer (Eds). The Outer Circle. Women in the Scientific Community, pp. 27–57. New York: W W Norton & Company.

Acknowledgements

We acknowledge funding from the Canada research chair program. We are indebted to the three referees and to the editor for their extremely pertinent comments. We are grateful for the help of Carl St-Pierre for his advice on statistics and of Seyed Reza Mirnezami for his clever Stata programming. We would also like to acknowledge the helpful comments on a previous version of the paper from the participants of the DRUID and EU-SPRI conferences. All remaining errors are our own.

Marginal effects of the results

Table 1– Marginal effects of the regression results – Number of articles – Health fields 2000-2012 (Table D1)

Variables	Marginal effects	Art-H-1	Art-H-2	Art-H-3	Art-H-4	Art-H-5	Art-H-6
ln(avgPubFundO3$_t$)	**Men**[a]	**0.0664***	**0.0414***	**0.0663***	**0.0676***	**0.0413***	**0.0433***
	Women[b]		**0.1630**ii[d]			**0.1630**ii	**0.1553**ii
ln(avgPrivFundO3$_t$)	**Men**	0.0407	0.0423	0.0407	0.0391	0.0422	0.0447
	Women		-0.0072			-0.0069	-0.0323
ln(avgPhilFundO3$_t$)	**Men**	**0.1146*****	**0.1311*****	**0.1146*****	**0.1132*****	**0.1311*****	**0.1323*****
	Women		0.0518			0.0522	0.0338
ln(ImpactFact5$_t$)	**Men**	**0.0718****	**0.0736****	**0.0746***	**0.0668****	**0.0755***	**0.0896****
	Women			0.0634		0.0680	0.0058
ln(avgAuthors$_t$)	**Men**	0.0088	0.0087	0.0089	-0.0064	0.0088	-0.0081
	Women				**0.0514**iii		**0.0585**iii
dFemale x ln(avgPubFundO3$_t$)	**Women vs men**[c]		**0.3872****			**0.3874****	**0.3566****
dFemale x ln(avgPrivFundO3$_t$)	**Women vs men**		-0.1036			-0.1028	-0.1609
dFemale x ln(avgPhilFundO3$_t$)	**Women vs men**		-0.2122			-0.2110	**-0.2632***
dFemale x ln(ImpactFact5$_t$)	**Women vs men**			-0.0123		-0.0082	-0.0912
dFemale x ln(avgAuthors$_t$)	**Women vs men**				**0.6591*****		**0.7595*****

Notes: [a]The marginal effects for Men are calculated while varying only the explanatory variable and keeping all other variables at the mean. [b]The marginal effects for Women are calculated while varying only the explanatory variable while keeping all other variables at the mean and dFemale = 1. [c]The marginal effects of Women vs Men are calculated while varying only dFemale from 1 to 0 and keeping all other variables at the mean. [d]The significance of the coefficients of the interactive variables are reported in the upper part of the table for the Women rows (iii, ii, i represent levels of significance at the 1%, 5% and 10% levels respectively). ***, **, * show significance at the 1%, 5% and 10% level respectively. Natural logarithm of the z-scores used for all continuous variables.

Table 2 – Marginal effects of the regression results – Number of articles – NSE fields 2000-2012 (Table D2)

Variables	Marginal effects	Art-NSE-1	Art-NSE-2	Art-NSE-3	Art-NSE-4	Art-NSE-5	Art-NSE-6
ln(avgPubFundO3$_t$)	**Men**[a]	**0.0499****	**0.0500****	**0.0498****	**0.0496****	**0.0502****	**0.0507****
	Women[b]		0.0207			0.0193	0.0135
ln(avgPrivFundO3$_t$)	**Men**	0.0388	0.0249	0.0387	0.0408	0.0252	0.0259
	Women		0.2502			0.2422	0.2446
ln(avgPhilFundO3$_t$)	**Men**	**0.1347*****	**0.1194****	**0.1350*****	**0.1346*****	**0.1202****	**0.1216****
	Women		**0.2888**i[d]			**0.2830**i	**0.2679**i
ln(ImpactFact5$_t$)	**Men**	**0.0440****	**0.0439****	0.0338	**0.0445****	0.0379	0.0387
	Women			0.1018		0.0770	0.0737
ln(avgAuthors$_t$)	**Men**	**0.0046*****	**0.0046*****	**0.0046*****	**0.0042*****	**0.0046*****	**0.0044*****
	Women				0.0062		-0.1033
dFemale x ln(avgPubFundO3$_t$)	**Women vs men**[c]		-0.0812			-0.0856	0.0054
dFemale x ln(avgPrivFundO3$_t$)	**Women vs men**		**0.4039***			**0.3888***	**0.3919***
dFemale x ln(avgPhilFundO3$_t$)	**Women vs men**		0.1817			0.1747	0.1570
dFemale x ln(ImpactFact5$_t$)	**Women vs men**			0.0910		0.0524	0.0469
dFemale x ln(avgAuthors$_t$)	**Women vs men**				0.2269		0.1130

Notes: [a]The marginal effects for Men are calculated while varying only the explanatory variable and keeping all other variables at the mean. [b]The marginal effects for Women are calculated while varying only the explanatory variable while keeping all other variables at the mean and dFemale = 1. [c]The marginal effects of Women vs Men are calculated while varying only dFemale from 1 to 0 and keeping all other variables at the mean. [d]The significance of the coefficients of the interactive variables are reported in the upper part of the table for the Women rows (iii, ii, i represent levels of significance at the 1%, 5% and 10% levels respectively). ***, **, * show significance at the 1%, 5% and 10% level respectively. Natural logarithm of the z-scores used for all continuous variables.

Table 3 – Marginal effects of the regression results – Normalised citations – Health fields 2000-2012 (Table D3)

Variables	Marginal effects	H-1	H-2	H-3	H-4	H-5	H-6	H-7	H-7alt	H-7sq[a]
ln(avgPubFundO3$_t$)	**Men[b]**	**-0.0667** *	**-0.0817** *	**-0.0672** *	**-0.0707** *	**-0.0713** *	**-0.0878** *	**-0.1053** *	-0.0749	**-0.1635** *i
	Women[c]		0.0021	-0.0626	-0.0666	-0.0617	-0.0324	-0.0018	-0.0020	**0.0252** **i
ln(avgPrivFundO3$_t$)	**Men**	0.0046	0.1007	0.0097	0.0010	0.0073	0.0195	**0.1396** *	**0.2077** ***	0.2199
	Women		**-0.3095** iii[e]	0.0090	0.0009	0.0063	0.0072	**-0.1395** i	**-0.2303** iii	-0.2165
ln(avgPhilFundO3$_t$)	**Men**	0.0328	0.0451	0.0309	0.0245	0.0342	0.0423	0.0566	0.0547	0.1116
	Women		-0.0106	0.0287	0.0230	0.0297	0.0156	-0.0073	-0.0041	-0.0740
ln(nbArtFirst$_t$)	**Men**	**0.2335** ***	**0.2559** ***	**0.1860** ***		**0.2408** ***	**0.3111** ***	**0.2778** ***	**0.2699** ***	**0.2697** ***
	Women		0.1664	0.3602		0.2086	0.1149	0.1254	0.2564	0.1443
ln(nbArtLast$_t$)	**Men**	0.0005	-0.0005	-0.0067		0.0008	0.0040	0.0010	-0.0086	-0.0003
	Women		-0.0003	0.0197		0.0007	0.0015	0.0015	0.0005	0.0025
ln(nbArtMiddle$_t$)	**Men**	**0.1113** ***	**0.1224** ***	**0.1443** ***		**0.1142** ***	**0.1458** ***	**0.1860** ***	**0.1532** ***	**0.1750** ***
	Women		0.0796	**0.0255** iii		0.0989	0.0538	**0.0193** i	**0.0060** iii	**0.0220** i
ln(nbArticles$_t$)	**Men**				**0.0858** ***					
	Women				0.0649					
ln(ImpactFact5$_t$)	**Men**	**3.6439** ***	**4.0284** ***	**3.7112** ***	**3.5464** ***	**4.0171** ***	**4.8259** ***	**5.3996** ***	**4.6961** ***	**5.1673** ***
	Women		2.6195	3.4571	3.3396	**2.7700** iii	1.7823	1.2280	**1.9173** iii	1.4191
ln(avgAuthors$_t$)	**Men**	**0.2218** ***	**0.2480** ***	**0.2264** ***	**0.2217** ***	**0.2340** ***	**0.3900** ***	**0.4145** ***	**0.2605** ***	**0.3997** ***
	Women		0.1612	0.2109	0.2088	0.2027	**0.0307** iii	**0.0328** iii	0.1296	**0.0398** iii
dFemale x ln(avgPubFundO3$_t$)	**Women vs men[d]**		0.2881					0.3320	0.2377	**-1.1131** *
dFemale x ln(avgPrivFundO3$_t$)	**Women vs men**		**-1.0368** ***					**-1.2528** ***	**-1.1646** ***	-0.5823
dFemale x ln(avgPhilFundO3$_t$)	**Women vs men**		-0.1586					-0.2207	-0.1625	1.3757
dFemale x ln(nbArtFirst$_t$)	**Women vs men**			0.3041				0.3291	0.3669	0.3075
dFemale x ln(nbArtLast$_t$)	**Women vs men**			0.1127				0.0202	0.0388	0.0342
dFemale x ln(nbArtMiddle$_t$)	**Women vs men**			**-0.5174** ***				**-0.5021** *	**-0.6267** ***	**-0.4702** *
dFemale x ln(nbArticles$_t$)	**Women vs men**				-0.1415					
dFemale x ln(ImpactFact5$_t$)	**Women vs men**					**-0.3475** ***		-0.2295	**-0.3556** ***	-0.2154
dFemale x ln(avgAuthors$_t$)	**Women vs men**						**-2.0692** ***	**-2.0361** ***		**-1.9337** ***

Notes: [a]The * and ** below represent the level of significance of ln(*avgPubFundO3$_t$*) as a linear and quadratic effect, i and ii below represent the level of significance of dFemale x ln(*avgPubFundO3$_t$*) as an interactive linear and quadratic effect. In addition, in this particular regression the squared terms of the three different types of funding are added to the model. [b]The marginal effects for Men are calculated while varying only the explanatory variable and keeping all other variables at the mean. [c]The marginal effects for Women are calculated while varying only the explanatory variable while keeping all other variables at the mean and dFemale = 1. [d]The marginal effects of Women vs Men are calculated while varying only dFemale from 1 to 0 and keeping all other variables at the mean. [e]The significance of the coefficients of the interactive variables are reported in the upper part of the table for the Women rows (iii, ii, i represent levels of significance at the 1%, 5% and 10% levels respectively).

Table 4 – Marginal effects of the regression results – Normalised citations – NSE fields 2000-2012 (Table D4)

Variables	Marginal effects	NSE-1	NSE-2	NSE-3	NSE-4	NSE-5	NSE-6	NSE-7	NSE-7sq[a]
ln(avgPubFundO3$_t$)	**Men**[b]	-0.0149	-0.0026	-0.0148	-0.0193	-0.0150	-0.0151	-0.0022	-0.0066
	Women[c]		-0.1713	-0.0170	-0.0214	-0.0151	-0.0127	-0.1692	-0.0941
ln(avgPrivFundO3$_t$)	**Men**	-0.0451	-0.0546	-0.0443	-0.0458	-0.0451	-0.0470	-0.0548	-0.0428
	Women		0.2468	-0.0508	-0.0509	-0.0457	-0.0395	0.2224	-0.0106
ln(avgPhilFundO3$_t$)	**Men**	0.0963	0.0377	0.0939	0.0826	0.0961	0.0994	0.0398	-0.0009
	Women		**0.8742** ii[e]	0.1077	0.0916	0.0974	0.0837	**0.8157** ii	0.6883
ln(nbArtFirst$_t$)	**Men**	**0.2539** ***	**0.2445** ***	**0.2567** ***		**0.2534** ***	**0.2608** ***	**0.2542** ***	**0.2523** ***
	Women		0.3320	0.2493		0.2567	0.2196	0.2775	0.2814
ln(nbArtLast$_t$)	**Men**	**0.0398** ***	**0.0379** ***	**0.0310** **		**0.0397** ***	**0.0404** ***	**0.0310** **	**0.0304** **
	Women		0.0514	0.1179		0.0402	0.0340	0.1203	0.1247
ln(nbArtMiddle$_t$)	**Men**	**0.0476** ***	**0.0456** ***	**0.0458** ***		**0.0475** ***	**0.0490** ***	**0.0444** ***	**0.0437** ***
	Women		0.0619	0.0655		0.0481	0.0412	0.0792	0.0823
ln(nbArticles$_t$)	**Men**				**0.0538** ***				
	Women				0.0889				
ln(ImpactFact5$_t$)	**Men**	**0.9688** ***	**0.9294** ***	**0.9514** ***	**0.9073** ***	**0.9620** ***	**0.9944** ***	**0.9317** ***	**0.9253** ***
	Women		1.2619	1.0915	1.0071	1.0054	0.8372	1.2578	1.2995
ln(avgAuthors$_t$)	**Men**	**0.0167** ***	**0.0160** ***	**0.0164** ***	**0.0156** ***	**0.0166** ***	**0.0175** ***	**0.0165** ***	**0.0164** ***
	Women		0.0217	0.0189	0.0173	0.0169	0.0120	0.0182	0.0186
dFemale x ln(avgPubFundO3$_t$)	**Women vs men**[d]		-0.3250					-0.3332	-0.7377
dFemale x ln(avgPrivFundO3$_t$)	**Women vs men**		0.4561					0.4320	1.1372
dFemale x ln(avgPhilFundO3$_t$)	**Women vs men**		**0.6447** **					**0.6189** **	0.7821
dFemale x ln(nbArtFirst$_t$)	**Women vs men**			-0.0500				-0.0538	-0.0576
dFemale x ln(nbArtLast$_t$)	**Women vs men**			0.3051				0.2565	0.2584
dFemale x ln(nbArtMiddle$_t$)	**Women vs men**			0.0500				0.0712	0.0746
dFemale x ln(nbArticles$_t$)	**Women vs men**				0.2170				
dFemale x ln(ImpactFact5$_t$)	**Women vs men**					0.0278		0.0258	0.0275
dFemale x ln(avgAuthors$_t$)	**Women vs men**						-0.3402	-0.2805	-0.2865

Notes: [a]In this particular regression the squared terms of the three different types of funding are added to the model. [b]The marginal effects for Men are calculated while varying only the explanatory variable and keeping all other variables at the mean. [c]The marginal effects for Women are calculated while varying only the explanatory variable while keeping all other variables at the mean and dFemale = 1. [d]The marginal effects of Women vs Men are calculated while varying only dFemale from 1 to 0 and keeping all other variables at the mean. [e]The significance of the coefficients of the interactive variables are reported in the upper part of the table for the Women rows (iii, ii, i represent levels of significance at the 1%, 5% and 10% levels respectively).

Table 5 – Marginal effects of the regression results – Normalised citations (adding shareWomen) – Health and NSE fields 2006-2012 (Table D5)

Variables	Marginal effects	H-1b	H-7b	H-10b	NSE-1b	NSE-7b	NSE-10b
ln(avgPubFundO3$_t$)	**Men**[a]	**-0.1702** ***	**-0.2018** **	**-0.1844** **	-0.1068	-0.1096	-0.0969
	Women[b]	-0.1792	1.5852	1.4915	-0.1070	-3.3780	-2.8567
ln(avgPrivFundO3$_t$)	**Men**	0.1210	**0.4046** **	**0.3715** **	-0.0084	-0.0185	-0.0175
	Women	0.1214	**-0.1558** ii	**-0.1469** ii[d]	-0.0084	0.3532	0.3134
ln(avgPhilFundO3$_t$)	**Men**	0.0534	0.1421	0.1236	-0.0427	-0.0980	-0.0891
	Women	0.0536	**-0.1320** ii	**-0.1186** ii	-0.0427	**1.5229** ii	**1.3459** ii
ln(nbArtFirst$_t$)	**Men**	**0.2978** ***	**0.2402** **		**0.2899** ***	**0.2460** ***	
	Women	0.2989	**0.3299** ii		0.2899	0.7691	
ln(nbArtLast$_t$)	**Men**	0.0177	0.0044		**0.0420** **	**0.0324** **	
	Women	0.0177	0.0298		0.0420	0.1891	
ln(nbArtMiddle$_t$)	**Men**	**0.0743** ***	**0.1428** ***		**0.0597** ***	**0.0533** ***	
	Women	0.0746	**-0.0164** iii		0.0597	0.1312	
ln(nbArticles$_t$)	**Men**			**0.0878** ***			**0.0607** ***
	Women			0.0278			0.2174
propNbArtFirst$_t$	**Men**			**0.3001** *			**0.0408** ***
	Women			**0.3481** i			0.1242
propNbArtMiddle$_t$	**Men**			**0.0723** ***			**0.0209** **
	Women			**-0.0141** ii			**-0.0848** ii
ln(ImpactFact5$_t$)	**Men**	**3.5886** ***	**4.4825** ***	**4.1236** ***	**1.0492** ***	**0.9099** ***	**0.7979** ***
	Women	0.3547	0.1889	0.1590	0.0266	0.1175	0.1305
ln(avgAuthors$_t$)	**Men**	**0.3534** ***	**0.4441** ***	**0.4048** ***	**0.0266** ***	**0.0231** ***	**0.0198** ***
	Women	0.3547	0.1837	0.1669	0.0266	0.0627	0.0596
shareWomen$_t$	**Men**	**-0.1946** **	**-0.2525** **	**-0.2310** **	0.0228	0.0189	0.0129
	Women	**-0.1953** ii	**-0.1070** i	**-0.0957** i	0.0228	0.0516	0.0343
dFemale x shareWomen$_t$	**Women vs men**[c]	**0.0089** **	**0.0095** *	**0.0088** *	0.0002	0.0002	0.0002
dFemale x ln(avgPubFundO3$_t$)	**Women vs men**		-0.1320	-0.1214		0.1137	0.1075
dFemale x ln(avgPrivFundO3$_t$)	**Women vs men**		**-0.8846** **	**-0.8316** **		0.4485	0.4120
dFemale x ln(avgPhilFundO3$_t$)	**Women vs men**		**-1.0592** **	**-0.9619** **		**1.4591** **	**1.3313** **
dFemale x ln(nbArtFirst$_t$)	**Women vs men**		**0.7885** **			0.0562	
dFemale x ln(nbArtLast$_t$)	**Women vs men**		0.2969			0.2642	
dFemale x ln(nbArtMiddle$_t$)	**Women vs men**		**-0.9058** ***			-0.0265	
dFemale x ln(nbArticles$_t$)	**Women vs men**			-0.1959			0.1704
dFemale x propNbArtFirst$_t$	**Women vs men**			**0.3019** *			0.0212
dFemale x propNbArtMiddle$_t$	**Women vs men**			**-0.2794** **			**-0.1776** **
dFemale x ln(ImpactFact5$_t$)	**Women vs men**		0.0015	-0.0197		0.0343	0.0504
dFemale x ln(avgAuthors$_t$)	**Women vs men**		-0.1138	-0.0214		-0.0077	0.1352

Notes: [a]The marginal effects for Men are calculated while varying only the explanatory variable and keeping all other variables at the mean. [b]The marginal effects for Women are calculated while varying only the explanatory variable while keeping all other variables at the mean and dFemale = 1. [c]The marginal effects of Women vs Men are calculated while varying only dFemale from 1 to 0 and keeping all other variables at the mean. [d]The significance of the coefficients of the interactive variables are reported in the upper part of the table for the Women rows (iii, ii, i represent levels of significance at the 1%, 5% and 10% levels respectively).

Table 6 – Marginal effects of the regression results – Normalised citations (adding shareWomen and its interactions) – Health fields 2006-2012 (Table D6)

Variables	Marginal effects	H-2c[i]	H-3c[i]	H-4c[ii]	H-5c[ii]	H-6c[i]	H-7c[i]	H-7c alt[i]	H-7csq[i]
ln(avgPubFundO3$_t$)	**Men[a]**	**-0.1708** ***	**-0.1626** **	**-0.1655** ***	**-0.1739** ***	**-0.1743** ***	**-0.1707** ***	**-0.1684** ***	-0.1762
	Women[b]	-0.1714	-0.1631	-0.1661	-0.1745	-0.1749	-0.1712	-0.1689	-0.1767
ln(avgPrivFundO3$_t$)	**Men**	0.0771	0.1118	0.1156	0.1121	0.1239	0.0767	0.0730	0.0561
	Women	0.0773	0.1121	0.1160	0.1125	0.1243	0.0770	0.0732	0.0563
ln(avgPhilFundO3$_t$)	**Men**	**0.0133** sss[c]	0.0484	0.0395	0.0566	0.0520	**0.0143** sss	**0.0154** ss	-0.0429
	Women	0.0123	0.0485	0.0397	0.0568	0.0522	0.0168	0.0170	-0.0357
ln(nbArtFirst$_t$)	**Men**	**0.2952** ***	**0.3071** ***		**0.3002** ***	**0.2988** ***	**0.3056** ***	**0.3060** ***	**0.3097** ***
	Women	0.2962	0.3184		0.3012	0.2998	0.3169	0.3180	0.3208
ln(nbArtLast$_t$)	**Men**	0.0164	0.0204		0.0182	0.0159	0.0193	0.0204	0.0169
	Women	0.0165	0.0204		0.0183	0.0159	0.0194	0.0205	0.0170
ln(nbArtMiddle$_t$)	**Men**	**0.0744** ***	**0.0646** ***sss		**0.0740** ***	**0.0757** ***	**0.0668** ***sss	**0.0656** ***sss	**0.0682** ***sss
	Women	0.0746	0.0648		0.0743	0.0760	0.0670	0.0658	0.0684
ln(nbArticles$_t$)	**Men**			**0.0694** ***					
	Women			0.0697					
ln(ImpactFact5$_t$)	**Men**	**3.5350** ***	**3.5532** ***	**3.4150** ***	**3.5343** ***ss	**3.5843** ***	**3.4756** ***	**3.4718** ***s	**3.4977** ***
	Women	3.5466	3.5651	3.4279	3.5467	3.5965	3.4860	3.4823	3.5086
ln(avgAuthors$_t$)	**Men**	**0.3471** ***	**0.3529** ***	**0.3363** ***	**0.3529** ***	**0.3350** ***ss	**0.3392** ***	**0.3472** ***	**0.3407** ***
	Women	0.3482	0.3541	0.3376	0.3541	0.3362	0.3403	0.3482	0.3418
shareWomen$_t$	**Men**	1.3890 *	0.1451	0.2199	0.3890	**3.9990** **	**4.2245** *	**1.9045** **	-1.0683
	Women	1.9184	0.5108	0.6262	0.8100	4.9358	5.0613	2.4573	-0.8941
dFemale x shareWomen$_t$	**Men**	**-0.5392** *	**-0.5863** *	**-0.5333** **	**-0.2620** **	**-0.2094** *	**-0.7314** *	**-0.7374** *	**-0.5648** *
	Women	-0.2802	-0.3251	-0.2483	0.0614	0.1135	-0.5225	-0.5255	-0.3247

Notes: [i,ii] indicate the level of significance of the dFemale x shareWomen interactive variable which is used to calculate the Women rows. [a]The marginal effects for Men are calculated while varying only the explanatory variable and keeping all other variables at the mean. [b]The marginal effects for Women are calculated while varying only the explanatory variable while keeping all other variables at the mean and dFemale = 1. The significance of the coefficients of the interactive variables with shareWomen are reported in the Men rows (sss, ss, s represent levels of significance at the 1%, 5% and 10% levels respectively).

Table 7 – Marginal effects of the regression results – Normalised citations (adding shareWomen and its interactions) – NSE fields 2006-2012 (Table D7)

Variables	Marginal effects	NSE-2c	NSE-3c	NSE-4c	NSE-5c	NSE-6c	NSE-7c	NSE-7csq
$\ln(avgPubFundO3_t)$	**Men**[a]	-0.1024	-0.1091	-0.1038	-0.1045	-0.1070	-0.1020	-0.1863
	Women[b]	-0.1024	-0.1091	-0.1038	-0.1045	-0.1070	-0.1020	-0.1863
$\ln(avgPrivFundO3_t)$	**Men**	-0.0137	-0.0070	-0.0079	-0.0072	-0.0084	-0.0119	-0.0246
	Women	-0.0137	-0.0070	-0.0079	-0.0072	-0.0084	-0.0119	-0.0246
$\ln(avgPhilFundO3_t)$	**Men**	-0.0292	-0.0432	-0.0440	-0.0437	-0.0427	-0.0298	**-0.4009** *
	Women	-0.0279	-0.0432	-0.0441	-0.0437	-0.0427	-0.0284	-0.4002
$\ln(nbArtFirst_t)$	**Men**	**0.2905** ***	**0.2825** ***		**0.2912** ***	**0.2899** ***	**0.2846** ***	**0.2986** ***
	Women	0.2905	0.2832		0.2912	0.2900	0.2853	0.2993
$\ln(nbArtLast_t)$	**Men**	**0.0415** **	**0.0413** **		**0.0425** ***	**0.0420** **	**0.0413** **	**0.0424** **
	Women	0.0415	0.0413		0.0425	0.0420	0.0413	0.0424
$\ln(nbArtMiddle_t)$	**Men**	**0.0598** ***	**0.0591** ***		**0.0596** ***	**0.0597** ***	**0.0590** ***	**0.0612** ***
	Women	0.0598	0.0591		0.0596	0.0597	0.0590	0.0612
$\ln(nbArticles_t)$	**Men**			**0.0728** ***				
	Women			0.0728				
$\ln(ImpactFact5_t)$	**Men**	**1.0529** ***	**1.0465** ***	**0.9253** ***	**1.0511** ***ss[c]	**1.0494** ***	**1.0524** ***ss	**1.1021** ***ss
	Women	1.0530	1.0466	0.9254	1.0512	1.0495	1.0525	1.1022
$\ln(avgAuthors_t)$	**Men**	**0.0266** ***	**0.0266** ***	**0.0234** ***	**0.0267** ***	**0.0265** ***	**0.0266** ***	**0.0279** ***
	Women	0.0266	0.0266	0.0235	0.0267	0.0265	0.0266	0.0279
$shareWomen_t$	**Men**	-0.2737	0.3985	0.2573	**0.3447** **	-0.0126	0.1577	1.2675
	Women	-0.3154	0.3172	0.1853	0.2511	-0.0720	0.0801	1.1497
dFemale x $shareWomen_t$	**Men**	0.1733	-0.1943	-0.0664	-0.1340	0.0137	-0.2138	-0.2471
	Women	0.1217	-0.2568	-0.1276	-0.2098	-0.0464	-0.2791	-0.3145

Notes: [a]The marginal effects for Men are calculated while varying only the explanatory variable and keeping all other variables at the mean. [b]The marginal effects for Women are calculated while varying only the explanatory variable while keeping all other variables at the mean and dFemale = 1. [c]The significance of the coefficients of the interactive variables with shareWomen are reported in the Men rows (sss, ss, s represent levels of significance at the 1%, 5% and 10% levels respectively).

Appendices

Appendix A – Descriptive statistics

Table A1 – Descriptive statistics – 2000-2012

Variables	mean	std. dev.	min	max	mean	std. dev.	min	max	Sig[a]	mean	std. dev.	min	max
HEALTH	**Women (N = 2749, n = 473)**				**Men (N = 7509, n = 1105)**					**Total (N =10258, n = 1578)**			
normCit10	1.2177	(1.8983)	0	53.1700	1.3439	(1.8825)	0	56.1210	***	1.3101	(1.8875)	0	56.1210
nbCit10	45.8931	(74.9289)	0	840	77.5821	(161.9706)	0	4106	***	69.0899	(144.5839)	0	4106
fracCit10	3.0806	(5.1212)	0	80.5000	3.5901	(5.9190)	0	112.0833	***	3.4536	(5.7203)	0	112.0833
nbArticles$_t$	3.0709	(2.6898)	1	22	3.8817	(3.7868)	1	67	***	3.6645	(3.5445)	1	67
nbArtFirst$_t$	0.4205	(0.7842)	0	8	0.3859	(0.8022)	0	9	***	0.3952	(0.7975)	0	9
nbArtLast$_t$	1.0582	(1.5312)	0	13	1.4511	(1.9889)	0	44	***	1.3458	(1.8852)	0	44
nbArtMiddle$_t$	1.5922	(1.8105)	0	15	2.0447	(2.4858)	0	35	***	1.9235	(2.3327)	0	35
fracArticles$_t$[b]	0.6635	(0.6272)	0.0027	5.3333	0.8331	(0.8289)	0.0037	8.9000	***	0.7877	(0.7836)	0.0027	8.9000
fracArtFirst$_t$	0.1375	(0.3149)	0	5	0.1582	(0.4078)	0	6	***	0.1527	(0.3852)	0	6
fracArtLast$_t$	0.2677	(0.4105)	0	4.1500	0.3621	(0.5026)	0	6.9000	***	0.3368	(0.4815)	0	6.9000
fracArtMiddle$_t$	0.2584	(0.3036)	0	2.7595	0.3128	(0.3632)	0	3.8569	***	0.2982	(0.3491)	0	3.8569
avgAuthors$_t$	6.8942	(11.1498)	1	374	7.1202	(12.8243)	1	668.5000	***	7.0596	(12.3976)	1	668.5000
ImpactFact5$_t$	1.1424	(0.7259)	0.0220	11.4170	1.2227	(0.7384)	0.0240	12.4760	***	1.2012	(0.7359)	0.0220	12.4760
avgPubFundO3$_t$	$100,654.8	($176,734.4)	$0	$3,299,431.0	$158,415.5	($373,102.4)	$0	$10,100,000	***	$142,936.4	($333,045.8)	$0	$10,100,000
avgPrivFundO3$_t$	$10,989.8	($43,893.2)	$0	$797,889.8	$38,137.7	($156,724.1)	$0	$6,865,308	***	$30,862.5	($136,529.0)	$0	$6,865,308
avgPhilFundO3$_t$	$16,990.0	($43,872.9)	$0	$703,228.9	$39,422.9	($184,722.3)	$0	$6,604,800	***	$33,411.2	($159,973.6)	$0	$6,604,800
dFemale										0.2680	(0.4429)	0	1
NSE	**Women (N = 1558, n = 256)**				**Men (N = 9072, n = 1478)**					**Total (N =10630, n = 1734)**			
normCit10	0.9815	(1.0590)	0	8.6740	1.1065	(1.7520)	0	74.5750	**	1.0881	(1.6691)	0	74.5750
nbCit10	29.6797	(62.9372)	0	988	39.7280	(101.5445)	0	1670	**	38.2552	(96.9161)	0	1670
fracCit10	2.4771	(3.5502)	0	48.1944	2.5852	(4.8590)	0	94.5000		2.5693	(4.6901)	0	94.5000
nbArticles$_t$	2.9660	(2.9290)	1	48	3.6488	(4.3967)	1	85	***	3.5487	(4.2205)	1	85
nbArtFirst$_t$	0.4089	(0.7920)	0	6	0.4278	(0.9165)	0	17		0.4250	(0.8993)	0	17
nbArtLast$_t$	1.1829	(1.3961)	0	13	1.6095	(2.1781)	0	25	***	1.5469	(2.0874)	0	25
nbArtMiddle$_t$	1.3742	(2.6353)	0	48	1.6116	(3.7034)	0	84	***	1.5768	(3.5678)	0	84
fracArticles$_t$	0.8417	(0.6760)	0.0055	6.5000	1.0680	(1.0228)	0.0006	12.2000	***	1.0348	(0.9829)	0.0006	12.2000
fracArtFirst$_t$	0.1733	(0.3850)	0	4	0.2109	(0.5007)	0	7.5000		0.2054	(0.4857)	0	7.5000
fracArtLast$_t$	0.3984	(0.5087)	0	6.5000	0.5479	(0.7582)	0	8.2444	***	0.5259	(0.7289)	0	8.2444
fracArtMiddle$_t$	0.2700	(0.3482)	0	3.3071	0.3093	(0.3894)	0	4.9333	***	0.3035	(0.3839)	0	4.9333
avgAuthors$_t$	9.7116	(106.8249)	1	3037.8410	10.0056	(84.8425)	1	3174.5000	***	9.9625	(88.4012)	1.0000	3174.5000
ImpactFact5$_t$	1.0873	(0.5430)	0.0720	6.9260	1.0989	(0.5326)	0.0160	7.7020		1.0972	(0.5341)	0.0160	7.7020
avgPubFundO3$_t$	$116,861.1	($262,970.7)	$0	$5,333,932	$145,446.0	($292,136.8)	$0	$11,700,000	***	$141,256.4	($288,212.5)	$0	$11,700,000
avgPrivFundO3$_t$	$10,941.0	($37,737.2)	$0	$413,252.3	$21,770.9	($82,750.6)	$0	$3,049,197	***	$20,183.6	($77,892.3)	$0	$3,049,197
avgPhilFundO3$_t$	$6,490.9	($33,675.7)	$0	$1,106,413	$9,033.5	($35,685.7)	$0	$8,50,003.9		$8,660.8	($35,408.1)	$0	$1,106,413
dFemale										0.1466	(0.3537)	0	1

Notes: [a]level of significance of the bi-lateral mean-comparison test (t-test) between men and women (***, **, * show significance at the 1%, 5% and 10% levels respectively). [b]Fractionnal counts are given for comparative purposes only.

Table A2 – Descriptive statistics – 2006-2012

Variables	mean	std. dev.	min	max	mean	std. dev.	min	max	Sig[a]	mean	std. dev.	min	max
HEALTH	**Women (N = 1679, n = 442)**				**Men (N = 4378, n = 1020)**					**Total (N =6075, n = 1462)**			
normCit10	1.2752	(2.2378)	0	53.1700	1.3742	(1.9855)	0	56.1210	***	1.3465	(2.0594)	0	56.1210
nbCit10	32.4791	(55.3476)	0	708	54.9354	(128.2758)	0	3035	***	48.6624	(113.2004)	0	3035
fracCit10	1.7098	(3.1696)	0	60.6667	1.9458	(3.3224)	0	86	***	1.8799	(3.2818)	0	86
nbArticles$_t$	3.3930	(2.9056)	1	22	4.2325	(4.1831)	1	67	***	3.9980	(3.8870)	1	67
nbArtFirst$_t$	0.4031	(0.7907)	0	8	0.3709	(0.8012)	0	9	***	0.3799	(0.7983)	0	9
nbArtLast$_t$	1.1715	(1.6403)	0	13	1.5619	(2.1844)	0	44	***	1.4528	(2.0544)	0	44
nbArtMiddle$_t$	1.8185	(1.9879)	0	15	2.2997	(2.6949)	0	35	***	2.1653	(2.5265)	0	35
fracArticles$_t^b$	0.6978	(0.6391)	0.0027	5.0512	0.8629	(0.8522)	0.0088	8.1290	***	0.8168	(0.8018)	0.0027	8.1290
fracArtFirst$_t$	0.1231	(0.2864)	0	3.3333	0.1471	(0.3981)	0	6	**	0.1404	(0.3704)	0	6
fracArtLast$_t$	0.2895	(0.4318)	0	4.1500	0.3780	(0.5185)	0	4.6354	***	0.3533	(0.4973)	0	4.6354
fracArtMiddle$_t$	0.2853	(0.3241)	0	2.7595	0.3378	(0.3742)	0	3.2799	***	0.3231	(0.3617)	0	3.2799
avgAuthors$_t$	7.0812	(11.4891)	1	374.0000	7.4622	(11.4857)	1	507.4000	***	7.3557	(11.4870)	1	507.4000
ImpactFact5$_t$	1.1359	(0.6866)	0.0410	11.4170	1.2218	(0.7262)	0.0280	11.4170	***	1.1978	(0.7163)	0.0280	11.4170
avgPubFundO3$_t$	$103,470.8	($192,184.2)	$0	$3,299,431.0	$150,146.7	($352,237.2)	$0	$10,100,000	***	$137,108.2	($316,478.8)	$0	$10,100,000
avgPrivFundO3$_t$	$9,354.2	($41,269.3)	$0	$797,889.8	$28,581.4	($163,877.6)	$0	$6,865,308	***	$23,210.4	($141,076.8)	$0	$6,865,308
avgPhilFundO3$_t$	$16,310.1	($41,454.7)	$0	$382,631.3	$41,347.5	($218,643.7)	$0	$6,604,800	***	$34,353.5	($187,229.8)	$0	$6,604,800
shareWomen	0.6162	(0.2039)	0	1	0.3244	(0.1929)	0	1	***	0.4059	(0.2357)	0	1
dFemale										0.2793	(0.4487)	0	1
NSE	**Women (N = 894, n = 235)**				**Men (N = 4904, n = 1289)**					**Total (N =5798, n = 1524)**			
normCit10	0.9962	(1.1126)	0	8.6740	1.1040	(1.6160)	0	43.2915	**	1.0874	(1.5495)	0	43.2915
nbCit10	20.3087	(49.7601)	0	988	29.5850	(75.8688)	0	1288	***	28.1547	(72.5330)	0	1288
fracCit10	1.4947	(2.2443)	0	22.2115	1.5689	(2.6024)	0	44.3500		1.5575	(2.5504)	0	44.3500
nbArticles$_t$	3.2081	(3.1262)	1	48	4.1752	(5.0034)	1	85	***	4.0260	(4.7750)	1	85
nbArtFirst$_t$	0.3501	(0.7266)	0	5	0.3921	(0.8782)	0	17		0.3857	(0.8566)	0	17
nbArtLast$_t$	1.3367	(1.4889)	0	13	1.9017	(2.3599)	0	25	***	1.8146	(2.2568)	0	25
nbArtMiddle$_t$	1.5213	(2.8219)	0	48	1.8813	(4.3562)	0	84	***	1.8258	(4.1585)	0	84
fracArticles$_t$	0.8919	(0.6989)	0.0055	6.5000	1.1736	(1.0645)	0.0006	12.2000	***	1.1301	(1.0217)	0.0006	12.2000
fracArtFirst$_t$	0.1463	(0.3376)	0	2	0.1896	(0.4694)	0	7.5000		0.1829	(0.4519)	0	7.5000
fracArtLast$_t$	0.4476	(0.5524)	0	6.5000	0.6345	(0.8061)	0	8.2444	***	0.6057	(0.7754)	0	8.2444
fracArtMiddle$_t$	0.2980	(0.3620)	0	2.1045	0.3495	(0.4212)	0	4.9333	***	0.3416	(0.4130)	0	4.9333
avgAuthors$_t$	12.0417	(137.0484)	1	3037.8410	11.6968	(107.7940)	1	3174.5000	**	11.7500	(112.7871)	1	3174.5000
ImpactFact5$_t$	1.0941	(0.5015)	0.0960	5.8530	1.1044	(0.5085)	0.0160	7.7020		1.1028	(0.5074)	0.0160	7.7020
avgPubFundO3$_t$	$121,383.6	($329,168.4)	$0	$5,333,932.0	$134,576.8	($205,930.0)	$0	$4,573,357	***	$132,542.6	($229,305.5)	$0	$5,333,932
avgPrivFundO3$_t$	$9,697.9	($35,595.2)	$0	$379,741.3	$16,636.8	($54,296.7)	$0	$1,080,265	***	$15,566.9	($51,912.7)	$0	$1,080,265
avgPhilFundO3$_t$	$5,854.8	($20,045.7)	$0	$368,804.2	$9,082.0	($33,807.6)	$0	$850,003.9		$8,584.4	($32,092.8)	$0	$850,003.9
shareWomen	0.5715	(0.2268)	0	1	0.1820	(0.2044)	0	1	***	0.2420	(0.2511)	0	1
dFemale										0.1542	(0.3612)	0	1

Notes: [a]level of significance of the bi-lateral mean-comparison test (t-test) between men and women (***, **, * show significance at the 1%, 5% and 10% levels respectively). [b]Fractionnal counts are given for comparative purposes only.

Table A3 – Correlation table – 2000-2012

Health \ NSE		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
ln(normCit10$_t$)	**1**	**1**	0.431 *	0.634 *	0.214 *	0.147 *	0.099 *	0.028 *	0.051 *	0.182 *	0.050 *	0.094 *	0.147 *
nbCit10$_t$	**2**	0.448 *	**1**	0.398 *	0.585 *	0.266 *	0.111 *	0.068 *	0.054 *	0.351 *	0.068 *	0.150 *	0.257 *
ln(fracCit10$_t$)	**3**	0.519 *	0.476 *	**1**	0.051 *	0.214 *	0.106 *	0.068 *	0.052 *	0.060 *	0.151 *	0.084 *	-0.044 *
nbArticles$_t$	**4**	0.179 *	0.496 *	0.084 *	**1**	0.573 *	0.168 *	0.084 *	0.069 *	0.729 *	0.109 *	0.400 *	0.511 *
ln(fracArticles$_t$)	**5**	0.106 *	0.419 *	0.229 *	0.832 *	**1**	0.199 *	0.095 *	0.083 *	0.750 *	0.366 *	0.647 *	0.307 *
ln(avgPubFundO3$_t$)	**6**	0.114 *	0.268 *	0.144 *	0.320 *	0.296 *	**1**	0.454 *	0.210 *	0.275 *	-0.023 *	0.244 *	0.184 *
ln(avgPrivFundO3$_t$)	**7**	0.079 *	0.191 *	0.096 *	0.208 *	0.146 *	0.294 *	**1**	0.133 *	0.139 *	-0.023 *	0.116 *	0.097 *
ln(avgPhilFundO3$_t$)	**8**	0.100 *	0.250 *	0.103 *	0.250 *	0.197 *	0.370 *	0.211 *	**1**	0.114 *	-0.041 *	0.116 *	0.085 *
ln(nbArticles$_t$)	**9**	0.169 *	0.449 *	0.074 *	0.912 *	0.805 *	0.354 *	0.201 *	0.244 *	**1**	0.261 *	0.644 *	0.646 *
ln(nbArtFirst$_t$)	**10**	0.054 *	0.168 *	0.172 *	0.302 *	0.467 *	0.052 *	0.066 *	0.047 *	0.333 *	**1**	-0.076 *	-0.079 *
ln(nbArtLast$_t$)	**11**	0.052 *	0.281 *	0.106 *	0.635 *	0.664 *	0.331 *	0.129 *	0.189 *	0.672 *	0.067 *	**1**	0.062 *
ln(nbArtMiddle$_t$)	**12**	0.186 *	0.370 *	-0.051 *	0.713 *	0.455 *	0.234 *	0.172 *	0.188 *	0.774 *	0.040 *	0.169 *	**1**
propNbArtFirst$_t$	**13**	-0.001	0.020 *	0.176 *	-0.008	0.192 *	-0.051 *	-0.012	-0.018	-0.010	0.782 *	-0.165 *	-0.208 *
propNbArtMiddle$_t$	**14**	-0.069 *	-0.001	0.056 *	0.058 *	0.124 *	0.131 *	-0.004	0.049 *	0.073 *	-0.220 *	0.657 *	-0.331 *
ln(fracArticles$_t$)	**15**	0.085 *	0.374 *	0.216 *	0.786 *	0.941 *	0.311 *	0.134 *	0.187 *	0.869 *	0.479 *	0.692 *	0.507 *
ln(fracArtFirst$_t$)	**16**	0.041 *	0.157 *	0.231 *	0.260 *	0.553 *	0.062 *	0.052 *	0.056 *	0.290 *	0.876 *	0.080 *	0.011
ln(fracArtLast$_t$)	**17**	0.033 *	0.253 *	0.158 *	0.564 *	0.695 *	0.296 *	0.083 *	0.151 *	0.606 *	0.082 *	0.941 *	0.104 *
ln(fracArtMiddle$_t$)	**18**	0.113 *	0.325 *	-0.007	0.673 *	0.495 *	0.219 *	0.137 *	0.154 *	0.736 *	0.055 *	0.169 *	0.943 *
propFracArtFirst$_t$	**19**	-0.001	0.020 *	0.176 *	-0.008	0.192 *	-0.051 *	-0.012	-0.018	-0.010	0.782 *	-0.165 *	-0.208 *
propFracArtMiddle$_t$	**20**	-0.069 *	-0.001	0.056 *	0.058 *	0.124 *	0.131 *	-0.004	0.049 *	0.073 *	-0.220 *	0.657 *	-0.331 *
ln(ImpactFact5$_t$)	**21**	0.512 *	0.263 *	0.279 *	0.105 *	0.037 *	0.130 *	0.052 *	0.105 *	0.108 *	-0.019	0.050 *	0.127 *
ln(avgAuthors$_t$)	**22**	0.236 *	0.131 *	-0.145 *	0.061 *	-0.143 *	0.026 *	0.099 *	0.055 *	0.054 *	-0.096 *	-0.115 *	0.219 *
dFemale	**23**	-0.047 *	-0.097 *	-0.063 *	-0.101 *	-0.096 *	-0.070 *	-0.107 *	-0.061 *	-0.076 *	0.022 *	-0.081 *	-0.063 *

Health \ NSE		**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**
ln(normCit10$_t$)	**1**	-0.005	-0.035 *	0.115 *	0.029 *	0.076 *	0.106 *	-0.005	-0.035 *	0.390 *	0.183 *	-0.024 *
nbCit10$_t$	**2**	-0.011	-0.028 *	0.190 *	0.059 *	0.139 *	0.132 *	-0.011	-0.028 *	0.189 *	0.213 *	-0.037 *
ln(fracCit10$_t$)	**3**	0.138 *	0.016	0.190 *	0.172 *	0.118 *	0.010	0.138 *	0.016	0.263 *	-0.140 *	0.005
nbArticles$_t$	**4**	-0.054 *	0.015	0.482 *	0.101 *	0.373 *	0.346 *	-0.054 *	0.015	0.120 *	0.199 *	-0.057 *
ln(fracArticles$_t$)	**5**	0.114 *	0.123 *	0.910 *	0.423 *	0.681 *	0.383 *	0.114 *	0.123 *	0.072 *	-0.186 *	-0.081 *
ln(avgPubFundO3$_t$)	**6**	-0.098 *	0.077 *	0.224 *	-0.033 *	0.215 *	0.180 *	-0.098 *	0.077 *	0.085 *	0.045 *	-0.063 *
ln(avgPrivFundO3$_t$)	**7**	-0.053 *	0.031 *	0.103 *	-0.026 *	0.096 *	0.095 *	-0.053 *	0.031 *	0.011	0.028 *	-0.062 *
ln(avgPhilFundO3$_t$)	**8**	-0.067 *	0.041 *	0.087 *	-0.041 *	0.098 *	0.080 *	-0.067 *	0.041 *	0.042 *	0.023 *	-0.038 *
ln(nbArticles$_t$)	**9**	-0.052 *	0.054 *	0.834 *	0.242 *	0.606 *	0.560 *	-0.052 *	0.054 *	0.131 *	0.121 *	-0.072 *
ln(nbArtFirst$_t$)	**10**	0.800 *	-0.333 *	0.419 *	0.922 *	-0.053 *	-0.058 *	0.800 *	-0.333 *	-0.003	-0.115 *	0.002
ln(nbArtLast$_t$)	**11**	-0.283 *	0.660 *	0.685 *	-0.056 *	0.962 *	0.084 *	-0.283 *	0.660 *	0.078 *	-0.097 *	-0.078 *
ln(nbArtMiddle$_t$)	**12**	-0.247 *	-0.368 *	0.356 *	-0.092 *	0.029 *	0.902 *	-0.247 *	-0.368 *	0.104 *	0.296 *	-0.038 *
propNbArtFirst$_t$	**13**	**1**	-0.433 *	0.131 *	0.781 *	-0.265 *	-0.238 *	1.000 *	-0.433 *	-0.055 *	-0.144 *	0.033 *
propNbArtMiddle$_t$	**14**	-0.313 *	**1**	0.132 *	-0.316 *	0.660 *	-0.362 *	-0.433 *	1.000 *	0.018	-0.176 *	-0.022 *
ln(fracArticles$_t$)	**15**	0.196 *	0.156 *	**1**	0.475 *	0.719 *	0.438 *	0.131 *	0.132 *	0.077 *	-0.200 *	-0.075 *
ln(fracArtFirst$_t$)	**16**	0.776 *	-0.209 *	0.552 *	**1**	-0.030 *	-0.072 *	0.781 *	-0.316 *	-0.001	-0.167 *	-0.011
ln(fracArtLast$_t$)	**17**	-0.147 *	0.668 *	0.725 *	0.104 *	**1**	0.065 *	-0.265 *	0.660 *	0.072 *	-0.181 *	-0.074 *
ln(fracArtMiddle$_t$)	**18**	-0.201 *	-0.333 *	0.554 *	0.025 *	0.120 *	**1**	-0.238 *	-0.362 *	0.061 *	0.083 *	-0.044 *
propFracArtFirst$_t$	**19**	1.000 *	-0.313 *	0.196 *	0.776 *	-0.147 *	-0.201 *	**1**	-0.433 *	-0.055 *	-0.144 *	0.033 *
propFracArtMiddle$_t$	**20**	-0.313 *	1.000 *	0.156 *	-0.209 *	0.668 *	-0.333 *	-0.313 *	**1**	0.018	-0.176 *	-0.022 *
ln(ImpactFact5$_t$)	**21**	-0.049 *	-0.023 *	0.030 *	-0.024 *	0.024 *	0.071 *	-0.049 *	-0.023 *	**1**	0.106 *	-0.022 *
ln(avgAuthors$_t$)	**22**	-0.135 *	-0.183 *	-0.185 *	-0.153 *	-0.189 *	0.060 *	-0.135 *	-0.183 *	0.219 *	**1**	-0.010
dFemale	**23**	0.036 *	-0.034 *	-0.076 *	-0.006	-0.080 *	-0.055 *	0.036 *	-0.034 *	-0.040 *	-0.009	**1**

Note: * show significance at the 5% level. Natural logarithm of the z-scores used for all exogenous variables.

Table A4 – Correlation table – 2006-2012

Health \ NSE		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
ln(normCit10$_t$)	**1**	**1**	0.388 *	0.613 *	0.226 *	0.161 *	0.089 *	0.022	0.080 *	0.196 *	0.067 *	0.098 *	0.155 *
nbCit10$_t$	**2**	0.426 *	**1**	0.393 *	0.662 *	0.248 *	0.088 *	0.044 *	0.097 *	0.384 *	0.056 *	0.140 *	0.281 *
ln(fracCit10$_t$)	**3**	0.510 *	0.483 *	**1**	0.060 *	0.234 *	0.081 *	0.046 *	0.056 *	0.073 *	0.145 *	0.104 *	-0.036 *
nbArticles$_t$	**4**	0.200 *	0.509 *	0.133 *	**1**	0.531 *	0.169 *	0.089 *	0.069 *	0.723 *	0.092 *	0.368 *	0.517 *
ln(fracArticles$_t$)	**5**	0.128 *	0.428 *	0.262 *	0.837 *	**1**	0.212 *	0.109 *	0.078 *	0.740 *	0.341 *	0.663 *	0.316 *
ln(avgPubFundO3$_t$)	**6**	0.120 *	0.282 *	0.155 *	0.329 *	0.326 *	**1**	0.449 *	0.216 *	0.294 *	-0.007	0.251 *	0.201 *
ln(avgPrivFundO3$_t$)	**7**	0.107 *	0.199 *	0.091 *	0.265 *	0.206 *	0.321 *	**1**	0.108 *	0.150 *	-0.009	0.120 *	0.099 *
ln(avgPhilFundO3$_t$)	**8**	0.108 *	0.279 *	0.112 *	0.264 *	0.217 *	0.379 *	0.232 *	**1**	0.114 *	-0.048 *	0.116 *	0.101 *
ln(nbArticles$_t$)	**9**	0.190 *	0.452 *	0.125 *	0.905 *	0.813 *	0.376 *	0.251 *	0.263 *	**1**	0.246 *	0.639 *	0.667 *
ln(nbArtFirst$_t$)	**10**	0.068 *	0.186 *	0.167 *	0.315 *	0.466 *	0.071 *	0.101 *	0.057 *	0.350 *	**1**	-0.057 *	-0.056 *
ln(nbArtLast$_t$)	**11**	0.062 *	0.293 *	0.143 *	0.637 *	0.678 *	0.355 *	0.146 *	0.204 *	0.674 *	0.085 *	**1**	0.053 *
ln(nbArtMiddle$_t$)	**12**	0.205 *	0.379 *	0.005	0.721 *	0.487 *	0.248 *	0.224 *	0.208 *	0.791 *	0.086 *	0.187 *	**1**
propNbArtFirst$_t$	**13**	0.016	0.029 *	0.157 *	0.009	0.197 *	-0.036 *	0.002	-0.017	0.013	0.777 *	-0.143 *	-0.169 *
propNbArtMiddle$_t$	**14**	-0.078 *	0.006	0.069 *	0.054 *	0.135 *	0.146 *	-0.028 *	0.052 *	0.065 *	-0.201 *	0.649 *	-0.340 *
ln(fracArticles$_t$)	**15**	0.104 *	0.375 *	0.240 *	0.787 *	0.942 *	0.343 *	0.181 *	0.206 *	0.877 *	0.479 *	0.702 *	0.540 *
ln(fracArtFirst$_t$)	**16**	0.062 *	0.180 *	0.229 *	0.271 *	0.542 *	0.090 *	0.090 *	0.067 *	0.305 *	0.875 *	0.091 *	0.054 *
ln(fracArtLast$_t$)	**17**	0.043 *	0.261 *	0.187 *	0.565 *	0.709 *	0.321 *	0.098 *	0.167 *	0.607 *	0.099 *	0.940 *	0.122 *
ln(fracArtMiddle$_t$)	**18**	0.121 *	0.316 *	0.031 *	0.676 *	0.526 *	0.229 *	0.186 *	0.170 *	0.750 *	0.097 *	0.188 *	0.940 *
propFracArtFirst$_t$	**19**	0.016	0.029 *	0.157 *	0.009	0.197 *	-0.036 *	0.002	-0.017	0.013	0.777 *	-0.143 *	-0.169 *
propFracArtMiddle$_t$	**20**	-0.078 *	0.006	0.069 *	0.054 *	0.135 *	0.146 *	-0.028 *	0.052 *	0.065 *	-0.201 *	0.649 *	-0.340 *
ln(ImpactFact5$_t$)	**21**	0.495 *	0.265 *	0.271 *	0.126 *	0.061 *	0.154 *	0.064 *	0.108 *	0.135 *	-0.010	0.064 *	0.150 *
ln(avgAuthors$_t$)	**22**	0.270 *	0.160 *	-0.109 *	0.083 *	-0.130 *	0.028 *	0.089 *	0.056 *	0.077 *	-0.075 *	-0.114 *	0.238 *
shareWomen$_t$	**23**	-0.085 *	-0.096 *	-0.096 *	-0.091 *	-0.104 *	-0.021	-0.073 *	-0.043 *	-0.089 *	-0.038 *	-0.052 *	-0.075 *
dFemale	**24**	-0.045 *	-0.089 *	-0.055 *	-0.097 *	-0.092 *	-0.060 *	-0.088 *	-0.069 *	-0.069 *	0.019	-0.073 *	-0.062 *

Health \ NSE		**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
ln(normCit10$_t$)	**1**	0.016	-0.047 *	0.128 *	0.047 *	0.073 *	0.113 *	0.016	-0.047 *	0.367 *	0.186 *	-0.008	-0.025
nbCit10$_t$	**2**	-0.007	-0.044 *	0.174 *	0.053 *	0.123 *	0.116 *	-0.007	-0.044 *	0.192 *	0.231 *	0.037 *	-0.046 *
ln(fracCit10$_t$)	**3**	0.132 *	0.027 *	0.202 *	0.169 *	0.135 *	0.023	0.132 *	0.027 *	0.259 *	-0.132 *	-0.055 *	-0.009
nbArticles$_t$	**4**	-0.044 *	-0.018	0.441 *	0.089 *	0.337 *	0.318 *	-0.044 *	-0.018	0.126 *	0.235 *	0.066 *	-0.073 *
ln(fracArticles$_t$)	**5**	0.118 *	0.115 *	0.910 *	0.400 *	0.700 *	0.399 *	0.118 *	0.115 *	0.071 *	-0.182 *	-0.060 *	-0.100 *
ln(avgPubFundO3$_t$)	**6**	-0.085 *	0.054 *	0.239 *	-0.016	0.207 *	0.200 *	-0.085 *	0.054 *	0.081 *	0.051 *	-0.027 *	-0.057 *
ln(avgPrivFundO3$_t$)	**7**	-0.041 *	0.015	0.118 *	-0.014	0.097 *	0.103 *	-0.041 *	0.015	0.004	0.028 *	-0.012	-0.054 *
ln(avgPhilFundO3$_t$)	**8**	-0.069 *	0.023	0.083 *	-0.048 *	0.086 *	0.097 *	-0.069 *	0.023	0.048 *	0.037 *	-0.033 *	-0.042 *
ln(nbArticles$_t$)	**9**	-0.045 *	0.015	0.822 *	0.228 *	0.594 *	0.567 *	-0.045 *	0.015	0.138 *	0.138 *	0.004	-0.099 *
ln(nbArtFirst$_t$)	**10**	0.800 *	-0.323 *	0.396 *	0.921 *	-0.034 *	-0.033 *	0.800 *	-0.323 *	-0.002	-0.097 *	-0.022	-0.007
ln(nbArtLast$_t$)	**11**	-0.258 *	0.638 *	0.703 *	-0.035 *	0.957 *	0.079 *	-0.258 *	0.638 *	0.076 *	-0.109 *	-0.044 *	-0.101 *
ln(nbArtMiddle$_t$)	**12**	-0.215 *	-0.419 *	0.369 *	-0.070 *	0.013	0.895 *	-0.215 *	-0.419 *	0.112 *	0.299 *	0.023	-0.050 *
propNbArtFirst$_t$	**13**	**1**	-0.415 *	0.131 *	0.795 *	-0.239 *	-0.203 *	1.000 *	-0.415 *	-0.047 *	-0.122 *	-0.015	0.027 *
propNbArtMiddle$_t$	**14**	-0.288 *	**1**	0.118 *	-0.308 *	0.644 *	-0.410 *	-0.415 *	1.000 *	0.008	-0.194 *	-0.018	-0.027 *
ln(fracArticles$_t$)	**15**	0.201 *	0.163 *	**1**	0.448 *	0.740 *	0.462 *	0.131 *	0.118 *	0.077 *	-0.204 *	-0.070 *	-0.098 *
ln(fracArtFirst$_t$)	**16**	0.784 *	-0.193 *	0.540 *	**1**	-0.008	-0.046 *	0.795 *	-0.308 *	0.001	-0.145 *	-0.039 *	-0.019
ln(fracArtLast$_t$)	**17**	-0.127 *	0.665 *	0.735 *	0.113 *	**1**	0.056 *	-0.239 *	0.644 *	0.065 *	-0.197 *	-0.055 *	-0.093 *
ln(fracArtMiddle$_t$)	**18**	-0.166 *	-0.340 *	0.587 *	0.064 *	0.141 *	**1**	-0.203 *	-0.410 *	0.067 *	0.063 *	-0.023	-0.055 *
propFracArtFirst$_t$	**19**	1.000 *	-0.288 *	0.201 *	0.784 *	-0.127 *	-0.166 *	**1**	-0.415 *	-0.047 *	-0.122 *	-0.015	0.027 *
propFracArtMiddle$_t$	**20**	-0.288 *	1.000 *	0.163 *	-0.193 *	0.665 *	-0.340 *	-0.288 *	**1**	0.008	-0.194 *	-0.018	-0.027 *
ln(ImpactFact5$_t$)	**21**	-0.041 *	-0.028 *	0.056 *	-0.004	0.034 *	0.086 *	-0.041 *	-0.028 *	**1**	0.120 *	-0.023	-0.024
ln(avgAuthors$_t$)	**22**	-0.124 *	-0.200 *	-0.169 *	-0.131 *	-0.196 *	0.068 *	-0.124 *	-0.200 *	0.203 *	**1**	0.061 *	-0.011
shareWomen$_t$	**23**	-0.022	0.013	-0.100 *	-0.073 *	-0.054 *	-0.058 *	-0.022	0.013	-0.080 *	-0.041 *	**1**	0.524 *
dFemale	**24**	0.035 *	-0.024	-0.070 *	-0.011	-0.071 *	-0.053 *	0.035 *	-0.024	-0.039 *	-0.018	0.503 *	**1**

Note: * show significance at the 5% level. Natural logarithm of the z-scores used for all exogenous variables.

Appendix B – Marginal effects calculations

Because we normalized the explanatory variables using a group-mean centered transformation (so that they follow a normal distribution), the variables that enter the regressions are transformed using the natural logarithm of the z-score of each explanatory variable plus 3 (5 in the case of **nbAuthor**). The marginal effects will therefore be calculated by comparing the predicted value of the discipline-normalised citations by adding one unit to the mean of each variable. For each variable, the unit is defined in Table B1.

Table B1 – Increments used for the calculation of the marginal effects

Variables		HEALTH – Total (N =10258, n = 1578)				NSE – Total (N =10630, n = 1734)			
2000-2012	**Increment**	**Mean**	**Std. dev.**	**Min**	**Max**	**Mean**	**Std. dev.**	**Min**	**Max**
nbArticles$_t$	1	3.6645	(3.5445)	1	67	3.5487	(4.2205)	1	85
nbArtFirst$_t$	1	0.3952	(0.7975)	0	9	0.4250	(0.8993)	0	17
nbArtLast$_t$	1	1.3458	(1.8852)	0	44	1.5469	(2.0874)	0	25
nbArtMiddle$_t$	1	1.9235	(2.3327)	0	35	1.5768	(3.5678)	0	84
avgAuthors$_t$	1	7.0596	(12.3976)	1	668.5000	9.9625	(88.4012)	1	3174.5000
ImpactFact5$_t$	0.50	1.2012	(0.7359)	0.0220	12.4760	1.0972	(0.5341)	0.0160	7.7020
avgPubFundO3$_t$	$100,000	$142,936	($333,046)	$0	$10,100,000	$141,256	($288,213)	$0	$11,700,000
avgPrivFundO3$_t$	$30,000	$30,862	($136,529)	$0	$6,865,308	$20,184	($77,892)	$0	$3,049,197
avgPhilFundO3$_t$	$30,000	$33,411	($159,974)	$0	$6,604,800	$8,661	($35,408)	$0	$1,106,413
shareWomen	0.25	0.4059421	0.2357169	0	1	0.2420	0.2511	0	1
Variables		**HEALTH – Total (N =6075, n = 1462)**				**NSE – Total (N =5798, n = 1524)**			
2006-2012		**Mean**	**Std. dev.**	**Min**	**Max**	**Mean**	**Std. dev.**	**Min**	**Max**
nbArticles$_t$	1	3.9980	(3.8870)	1	67	4.0260	(4.7750)	1	85
nbArtFirst$_t$	1	0.3799	(0.7983)	0	9	0.3857	(0.8566)	0	17
nbArtLast$_t$	1	1.4528	(2.0544)	0	44	1.8146	(2.2568)	0	25
nbArtMiddle$_t$	1	2.1653	(2.5265)	0	35	1.8258	(4.1585)	0	84
avgAuthors$_t$	1	7.3557	(11.4870)	1	507.4000	11.7500	(112.7871)	1	3174.5000
ImpactFact5$_t$	0.50	1.1978	(0.7163)	0.0280	11.4170	1.1028	(0.5074)	0.0160	7.7020
avgPubFundO3$_t$	$100,000	$137,108	($316,479)	$0	$10,100,000	$132,543	($229,306)	$0	$5,333,932
avgPrivFundO3$_t$	$30,000	$23,210	($141,077)	$0	$6,865,308	$15,567	($51,913)	$0	$1,080,265
avgPhilFundO3$_t$	$30,000	$34,354	($187,230)	$0	$6,604,800	$8,584	($32,093)	$0	$850,004
shareWomen	0.25	0.4059	(0.2357)	0	1	0.2420	(0.2511)	0	1

Suppose that

$$x_1' = x_1 + a$$

$$x_{1z} = \frac{x_1 - \bar{x}_1}{\sigma_{x_1}}$$

$$x_{1z}' = \frac{x_1 + a - \bar{x}_1}{\sigma_{x_1}} = x_{1z} + \frac{a}{\sigma_{x_1}} = x_{1z} + a_z$$

where a is the increment considered for the calculation of the marginal effect. The core of the analysis relies upon the results of the interactive variables with **dFemale** (dummy variable), which will be represented by x_2 in the equations below, and **shareWomen** (continuous variable). The equations above are also valid for x_2.

The predicted value of ln(y), where all the x's are calculated at the mean is given by:

$$\ln(\hat{y}) = \alpha + \beta_1 \ln(\bar{x}_{1z}) + \beta_2 \ln(\bar{x}_{2z}) + \beta_3 \ln(\bar{x}_{1z}) \times \ln(\bar{x}_{2z})$$

adding a constant a to x_1 yields

$$\ln\left(\hat{y}_{\bar{x}_1}\right) = \alpha + \beta_1 \ln\left(\bar{x}'_{1z}\right) + \beta_2 \ln\left(\bar{x}_{2z}\right) + \beta_3 \ln\left(\bar{x}'_{1z}\right) \times \ln\left(\bar{x}_{2z}\right)$$
$$= \alpha + \beta_1 \ln\left(\bar{x}_{1z} + a_z\right) + \beta_2 \ln\left(\bar{x}_{2z}\right) + \beta_3 \ln\left(\bar{x}_{1z} + a_z\right) \times \ln\left(\bar{x}_{2z}\right)$$

In our analysis, we x_2 may be a dummy variable (**dFemale**) or a continuous variable (**ShareWomen**). In the former case, x_2 takes two values, 0 and 1, and as a consequence, the predicted value of ln(y) will be calculated for x_2 = 0 (men) and for x_2 = 1(women) for each of the continuous variables x_1.

$$\ln\left(\hat{y}_{x_1}\right) = \begin{cases} \alpha + \beta_1 \ln\left(\bar{x}_{1z}\right) & \text{if } x_2 = 0 \\ \alpha + \beta_1 \ln\left(\bar{x}_{1z}\right) + \beta_2 + \beta_3 \ln\left(\bar{x}_{1z}\right) & \text{if } x_2 = 1 \end{cases}$$

adding a constant a to x_1 yields

$$\ln\left(\hat{y}_{x'_1}\right) = \begin{cases} \alpha + \beta_1 \ln\left(\bar{x}'_{1z}\right) & \text{if } x_2 = 0 \\ \alpha + \beta_1 \ln\left(\bar{x}'_{1z}\right) + \beta_2 + \beta_3 \ln\left(\bar{x}'_{1z}\right) & \text{if } x_2 = 1 \end{cases}$$

The impact of adding a to x_1 on y is given by the following equation:

$$\hat{y}_{\bar{x}'_1} - \hat{y}_{\bar{x}_1} = e^{\ln \hat{y}_{x'_1}} - e^{\ln \hat{y}_{x_1}}$$

Three types of marginal effects reported in each table for each of the continuous variables. The first, identified in the rows titled Men, compare the variation of x_1 when **dFemale** = 0, keeping all other variables at the mean.

$$\hat{y}_{\bar{x}'_1, x_2=0} - \hat{y}_{\bar{x}_1, x_2=0} = e^{\ln \hat{y}_{x'_1, x_2=0}} - e^{\ln \hat{y}_{x_1, x_2=0}}$$

The second, identified in the rows titled Women, compare the variation of x_1 when **dFemale** = 1, keeping all other variables at the mean.

$$\hat{y}_{\bar{x}'_1, x_2=1} - \hat{y}_{\bar{x}_1, x_2=1} = e^{\ln \hat{y}_{x'_1, x_2=1}} - e^{\ln \hat{y}_{x_1, x_2=1}}$$

The third correspond to the case where **dFemale** takes the value 0 and 1 while keeping all other variables at the mean. Next to the interactive variables with **dFemale**, the marginal effects (row titled Women vs Men) represent the difference of being male and female for the predicted value at the mean of all other variables.

$$\ln\left(\hat{y}_{x_2}\right) = \begin{cases} \alpha + \beta_1 \ln\left(\bar{x}_{1z}\right) & \text{if } x_2 = 0 \\ \alpha + \beta_1 \ln\left(\bar{x}_{1z}\right) + \beta_2 + \beta_3 \ln\left(\bar{x}_{1z}\right) & \text{if } x_2 = 1 \end{cases}$$

$$\hat{y}_{x_2=1} - \hat{y}_{x_2=0} = e^{\ln \hat{y}_{x_2=1}} - e^{\ln \hat{y}_{x_2=0}}$$

In the case where the interactive variable is a continuous variable (**shareWomen**), only the first two types of marginal effects will be presented. In the particular case where the marginal effect of **shareWomen** is examined, the predicted value corresponding to each interactive variable will be identical, hence only two values will be presented, one for dFemale = 1 and another for dFemale = 0.

Appendix C – Alternative measures of the dependent variable

The results presented in this appendix are available from the authors in an unpublished appendix.

Articles

For comparative purposes, we estimated the same models using the natural logarithm of the fractional number of articles (fracArticle) as a dependent variable (available from the authors in an unpublished appendix). In this case, private funding in the health fields is now positive and significant (the coefficients of the other two sources of funding remain positive and significant). The influence of the Impact Factor disappears and not surprisingly, the number of co-authors is now negative and significant. This last result was to be expected as the number of co-authors appears in the denominator of the dependent variable (the fractional number of articles). Regarding the interactive variables with gender, the impact of public funding for women in the health fields disappears and the negative impact of philanthropic funding is now significant. All other effects are similar in the health fields. In the NSE fields, women who are equally funded, regardless of the funding source, are equally prolific as their male colleagues. And contrarily to their female health colleagues, natural scientists and engineers seem to benefit even less than men from a greater number of co-authors. If the average number of authors is removed from the regressions, the results remain sensibly the same.

Citations

For comparative purposes, we have estimated the models using the number of citations (**nbCit10**), the natural logarithm of the fractional number of citations (**fracCit10**) as well as of the relative number of citations (**normCit10**). Only the latter measure normalizes the number of citations with respect to the discipline. The first two measures are plagued with disciplinary biases, i.e. some disciplines publish more, and thus cite more, other disciplines have a greater number of authors and hence publications are more likely to be cited[33]. Furthermore, there is a clear incidence of the number of articles (and the fractional number of articles) written as first, last or middle author on the number of citations obtained by an individual for the articles published in a given year. What we actually measure is the size effect related to the fact that specific disciplines publish more and hence receive more citation than others. When one examines the impact on the fractional citation count, only articles written as first or last author matter. Accounting for disciplinary discrepancies, only the number of first- and middle-author articles influences the relative citation rate (in fractional counts, the significance of middle-author articles disappears[34]). If we do not distinguish the articles by author

[33] As examples, electrical and computer engineering, mechanical and industrial engineering, as well as civil engineering have less than half of the publication record of physics and astronomy. In the health fields, public health is the most prolific, with most of medicine reaching about 75% the number of articles published in public health.

[34] An important point regarding the use of the fractional number of articles as an independent variable needs to be made at this point. We postulated at the start of this article that the increased visibility granted by a greater yearly publication record and by publishing with a greater number of authors should yield a larger relative number of citations, *ceteris paribus* (in general this is what we find). Although the use of the fractional number of articles is probably a better measure of the true workload/productivity of an individual (i.e. accounting for the collective nature of co-writing), its use in our regression models yields mitigated results. On the one hand, a greater fractional number of articles should have a positive effect on the citation rate in general. On the other hand, a greater number authors has a positive impact on the citation rate (this result is robust through all our models), but indirectly contributes to reducing the fractional number of articles (for each article, the number of authors is the denominator of the fraction used to calculate the fractional number of articles), which should thus reduce the citation rate. Both opposite effects are embedded within the fractional number of articles, hence the poor significance obtained for this family of variables (**fracArtFirst**, **fracArtLast**, **fracArtMiddle**, **fracArticles**). For this reason, we prefer the use of the raw number of articles as an independent variable.

rank, and only consider the number (and fractional number) of articles, their impact is similar regardless of the form of the dependent variable. Similar results are obtained for both the health and NSE fields.

In contrast, when considering the raw number of citations in the health fields, the amount of public funding has a negative effect and private funding has a positive effect. When using the fractional citation count as a dependent variable, public funding has a negative effect but philanthropic funding has a positive effect. In the NSE fields, private funding has a negative and significant impact on the fractional citation count, but positive effect on the number of citations. Hence depending on the indicator used as a dependent variable the results vary greatly. We estimated both models using random effects[35] in order to be able to add dummy variables for each discipline[36] and not surprisingly, this addition is enough to change the sign and significance of the funding variables according to discipline. In an ideal world, we would have enough critical mass to be able to estimate the models for each discipline, but our database does not allow such level of disaggregation while maintaining an appropriate number of observations and individual scientists. We must therefore resort to using dependent variables that are free of disciplinary bias. In our case, the relative number of citations is one such variable.

[35] Throughout our analysis, we have always checked that the fixed effect models are preferred to the random effect models. To be able to add discipline fixed effects, we had to estimate the models using random effects, which are technically rejected using Hausman specification tests.

[36] The disciplines added for the health fields are: dentistry, general medicine, kinesiology, laboratory medicine, medical specialities, nursing, other health sciences, public health, rehabilitation and surgical specialities. For the NSE fields, they are: agriculture, biology and botany, chemical engineering, chemistry, civil engineering, computer and information science, earth and ocean science, electrical and computer engineering, mathematics, mechanical and industrial engineering, other engineering, physics and astronomy, resource management and forestry.

Appendix D – Regression results

Table D1 – Regression results – Number of articles – Health fields 2000-2012

Variables	Art-H-1	Art-H-2	Art-H-3	Art-H-4	Art-H-5	Art-H-6
ln(avgPubFundO3$_t$)	0.1361 ***	0.0848 *	0.1359 ***	0.1386 ***	0.0847 *	0.0888 **
	(0.0402)	(0.0448)	(0.0403)	(0.0403)	(0.0448)	(0.0451)
ln(avgPrivFundO3$_t$)	0.0570	0.0592	0.0570	0.0547	0.0591	0.0625
	(0.0378)	(0.0403)	(0.0378)	(0.0380)	(0.0403)	(0.0406)
ln(avgPhilFundO3$_t$)	0.1942 ***	0.2224 ***	0.1943 ***	0.1919 ***	0.2223 ***	0.2243 ***
	(0.0339)	(0.0367)	(0.0339)	(0.0340)	(0.0367)	(0.0370)
ln(ImpactFact5$_t$)	0.0657 **	0.0674 **	0.0683 *	0.0611 **	0.0691 *	0.0820 **
	(0.0311)	(0.0311)	(0.0356)	(0.0312)	(0.0356)	(0.0359)
ln(avgAuthors$_t$)	0.0647	0.0640	0.0652	-0.0470	0.0643	-0.0595
	(0.0793)	(0.0794)	(0.0793)	(0.0909)	(0.0794)	(0.0915)
dFemale x		0.2492 **			0.2493 **	0.2295 **
ln(avgPubFundO3$_t$)		(0.1047)			(0.1047)	(0.1043)
dFemale x		-0.0693			-0.0688	-0.1077
ln(avgPrivFundO3$_t$)		(0.1134)			(0.1135)	(0.1135)
dFemale x		-0.1346			-0.1338	-0.1670 *
ln(avgPhilFundO3$_t$)		(0.0959)			(0.0962)	(0.0974)
dFemale x			-0.0103		-0.0069	-0.0767
ln(ImpactFact5$_t$)			(0.0682)		(0.0686)	(0.0720)
dFemale x				0.4239 ***		0.4885 ***
ln(avgAuthors$_t$)				(0.1518)		(0.1601)
Year dummies	yes ***	yes ***	yes ***	yes ***	yes ***	yes ***
Constant	3.1644 ***	3.1856 ***	3.1643 ***	3.2348 ***	3.1851 ***	3.2594 ***
	(0.2003)	(0.2023)	(0.2004)	(0.2047)	(0.2025)	(0.2058)
Nb observations	10154	10154	10154	10154	10154	10154
Nb academics	1474	1474	1474	1474	1474	1474
Avg nb of years	6.89	6.89	6.89	6.89	6.89	6.89
Log likelihood	-15146.7	-15143.1	-15146.7	-15142.6	-15143.1	-15138.2
χ^2	195.61 ***	203.55 ***	195.66 ***	203.35 ***	203.57 ***	212.91 ***

Note: ***, **, * show significance at the 1%, 5% and 10% level respectively, standard errors in parentheses. Natural logarithm of the z-scores used for all continuous variables.

Table D2 – Regression results – Number of articles – NSE fields 2000-2012

Variables	Art-NSE-1	Art-NSE-2	Art-NSE-3	Art-NSE-4	Art-NSE-5	Art-NSE-6
ln(avgPubFundO3$_t$)	0.0923 **	0.0924 **	0.0921 **	0.0918 **	0.0928 **	0.0938 **
	(0.0444)	(0.0466)	(0.0444)	(0.0444)	(0.0466)	(0.0467)
ln(avgPrivFundO3$_t$)	0.0525	0.0336	0.0523	0.0551	0.0341	0.0351
	(0.0450)	(0.0466)	(0.0450)	(0.0451)	(0.0466)	(0.0467)
ln(avgPhilFundO3$_t$)	0.1103 ***	0.0978 **	0.1106 ***	0.1102 ***	0.0984 **	0.0996 **
	(0.0384)	(0.0405)	(0.0384)	(0.0385)	(0.0405)	(0.0406)
ln(ImpactFact5$_t$)	0.0534 **	0.0533 **	0.0411	0.0540 **	0.0460	0.0470
	(0.0265)	(0.0265)	(0.0285)	(0.0265)	(0.0287)	(0.0287)
ln(avgAuthors$_t$)	0.3297 ***	0.3322 ***	0.3287 ***	0.3050 ***	0.3321 ***	0.3211 ***
	(0.0534)	(0.0537)	(0.0535)	(0.0570)	(0.0537)	(0.0573)
dFemale x		-0.0541			-0.0571	-0.0688
ln(avgPubFundO3$_t$)		(0.1509)			(0.1509)	(0.1521)
dFemale x		0.3046 *			0.2932 *	0.2956 *
ln(avgPrivFundO3$_t$)		(0.1712)			(0.1720)	(0.1722)
dFemale x		0.1387			0.1334	0.1199
ln(avgPhilFundO3$_t$)		(0.1144)			(0.1149)	(0.1177)
dFemale x			0.0826		0.0475	0.0426
ln(ImpactFact5$_t$)			(0.0702)		(0.0721)	(0.0726)
dFemale x				0.1399		0.0697
ln(avgAuthors$_t$)				(0.1070)		(0.1246)
Year dummies	yes ***	yes ***	yes ***	yes ***	yes ***	yes ***
Constant	2.3306 ***	2.3036 ***	2.3307 ***	2.3246 ***	2.3074 ***	2.3103 ***
	(0.1397)	(0.1395)	(0.1396)	(0.1393)	(0.1396)	(0.1396)
Nb observations	10492	10492	10492	10492	10492	10492
Nb academics	1596	1596	1596	1596	1596	1596
Avg nb of years	6.57	6.57	6.57	6.57	6.57	6.57
Log likelihood	-15407.0	-15403.6	-15406.3	-15406.2	-15403.4	-15403.2
χ^2	133.49 ***	139.38 ***	134.67 ***	134.63 ***	139.66 ***	139.32 ***

Note: ***, **, * show significance at the 1%, 5% and 10% level respectively, standard errors in parentheses. Natural logarithm of the z-scores used for all continuous variables.

Table D3 – Regression results – Normalised citations – Health fields 2000-2012

Variables	H-1	H-2	H-3	H-4	H-5	H-6	H-7	H-7alt	H-7sq
ln(avgPubFundO3$_t$)	-0.0554 *	-0.0615 *	-0.0548 *	-0.0606 *	-0.0571 *	-0.0556 *	-0.0614 *	-0.0516	0.2723 *
	(0.0320)	(0.0366)	(0.0320)	(0.0320)	(0.0320)	(0.0319)	(0.0365)	(0.0367)	(0.1432)
ln(avgPubFundO3$_t$)^2									-0.1036 **
									(0.0429)
ln(avgPrivFundO3$_t$)	0.0025	0.0501	0.0053	0.0006	0.0039	0.0083	0.0537 *	0.0933 ***	-0.0354
	(0.0298)	(0.0323)	(0.0298)	(0.0298)	(0.0298)	(0.0298)	(0.0322)	(0.0321)	(0.1576)
ln(avgPrivFundO3$_t$)^2									0.0331
									(0.0481)
ln(avgPhilFundO3$_t$)	0.0220	0.0275	0.0204	0.0170	0.0222	0.0217	0.0267	0.0305	-0.1368
	(0.0283)	(0.0320)	(0.0283)	(0.0284)	(0.0283)	(0.0283)	(0.0320)	(0.0322)	(0.1261)
ln(avgPhilFundO3$_t$)^2									0.0511
									(0.0368)
ln(nbArtFirst$_t$)	0.0654 ***	0.0650 ***	0.0516 ***		0.0650 ***	0.0664 ***	0.0550 ***	0.0629 ***	0.0556 ***
	(0.0148)	(0.0148)	(0.0172)		(0.0148)	(0.0148)	(0.0172)	(0.0172)	(0.0172)
ln(nbArtLast$_t$)	0.0004	-0.0003	-0.0050		0.0005	0.0023	0.0005	-0.0054	-0.0002
	(0.0160)	(0.0160)	(0.0183)		(0.0160)	(0.0160)	(0.0184)	(0.0184)	(0.0184)
ln(nbArtMiddle$_t$)	0.0958 ***	0.0955 ***	0.1213 ***		0.0949 ***	0.0958 ***	0.1120 ***	0.1092 ***	0.1098 ***
	(0.0169)	(0.0169)	(0.0195)		(0.0169)	(0.0169)	(0.0195)	(0.0196)	(0.0195)
ln(nbArticles$_t$)				0.1330 ***					
				(0.0242)					
ln(ImpactFact5$_t$)	0.8219 ***	0.8228 ***	0.8219 ***	0.8241 ***	0.8554 ***	0.8275 ***	0.8431 ***	0.8590 ***	0.8417 ***
	(0.0187)	(0.0187)	(0.0187)	(0.0187)	(0.0218)	(0.0187)	(0.0219)	(0.0219)	(0.0219)
ln(avgAuthors$_t$)	0.6230 ***	0.6308 ***	0.6243 ***	0.6416 ***	0.6334 ***	0.8242 ***	0.8064 ***	0.6080 ***	0.8104 ***
	(0.0504)	(0.0504)	(0.0504)	(0.0496)	(0.0505)	(0.0599)	(0.0604)	(0.0506)	(0.0604)
dFemale x		0.0639					0.0573	0.0488	-0.5594 *
ln(avgPubFundO3$_t$)		(0.0766)					(0.0766)	(0.0771)	(0.3228)
dFemale x									0.1971 *
ln(avgPubFundO3$_t$)^2									(0.1015)
dFemale x		-0.3149 ***					-0.2895 ***	-0.3267 ***	0.0707
ln(avgPrivFundO3$_t$)		(0.0822)					(0.0822)	(0.0825)	(0.4566)
dFemale x									-0.1309
ln(avgPrivFundO3$_t$)^2									(0.1475)
dFemale x		-0.0375					-0.0405	-0.0351	0.5009
ln(avgPhilFundO3$_t$)		(0.0692)					(0.0691)	(0.0696)	(0.4127)
dFemale x									-0.1818
ln(avgPhilFundO3$_t$)^2									(0.1412)
dFemale x			0.0519				0.0406	0.0529	0.0396
ln(nbArtFirst$_t$)			(0.0333)				(0.0333)	(0.0334)	(0.0333)
dFemale x			0.0204				0.0027	0.0060	0.0047
ln(nbArtLast$_t$)			(0.0370)				(0.0371)	(0.0372)	(0.0372)
dFemale x			-0.0977 ***				-0.0653 *	-0.1004 ***	-0.0637 *
ln(nbArtMiddle$_t$)			(0.0377)				(0.0381)	(0.0380)	(0.0381)
dFemale x				-0.0259					
ln(nbArticles$_t$)				(0.0486)					
dFemale x					-0.1209 ***		-0.0560	-0.1056 ***	-0.0547

Variables	H-1	H-2	H-3	H-4	H-5	H-6	H-7	H-7alt	H-7sq
ln(ImpactFact5$_t$)					(0.0406)		(0.0420)	(0.0409)	(0.0420)
dFemale x						-0.6407 ***	-0.5438 ***		-0.5352 ***
ln(avgAuthors$_t$)						(0.1036)	(0.1093)		(0.1102)
Year dummies	yes ***	yes ***	yes ***	yes ***	yes ***	yes ***	yes ***	yes ***	yes ***
Constant	-1.4685 ***	-1.4494 ***	-1.4727 ***	-1.4567 ***	-1.4852 ***	-1.5322 ***	-1.5143 ***	-1.3834 ***	-1.6330 ***
	(0.0991)	(0.0994)	(0.0991)	(0.0953)	(0.0992)	(0.0994)	(0.0998)	(0.0996)	(0.1856)
Nb observations	10258	10258	10258	10258	10258	10258	10258	10258	10258
Nb academics	1578	1578	1578	1578	1578	1578	1578	1578	1578
Log likelihood	-3118.29	-3109.30	-3112.30	-3126.12	-3113.04	-3095.68	-3084.21	-3152.23	-3078.78
F	157.39 ***	134.73 ***	134.39 ***	166.16 ***	149.27 ***	151.41 ***	109.94 ***	118.15 ***	88.99 ***
Avg nb of years	6.50	6.50	6.50	6.50	6.50	6.50	6.50	6.50	6.50
R^2 within	0.2360	0.2373	0.2369	0.2348	0.2368	0.2393	0.2410	0.2309	0.2418
R^2 overall	0.2956	0.2322	0.2974	0.2941	0.2852	0.0788	0.0615	0.1694	0.0765
R^2 between	0.4278	0.2619	0.4278	0.4237	0.3868	0.0622	0.0485	0.1681	0.0609

Note: ***, **, * show significance at the 1%, 5% and 10% level respectively, standard errors in parentheses. Natural logarithm of the z-scores used for all continuous variables.

Table D4 – Regression results – Normalised citations – NSE fields 2000-2012

Variables	NSE-1	NSE-2	NSE-3	NSE-4	NSE-5	NSE-6	NSE-7	NSE-7sq
ln(avgPubFundO3$_t$)	-0.0132	-0.0024	-0.0134	-0.0182	-0.0132	-0.0130	-0.0020	0.0496
	(0.0296)	(0.0313)	(0.0296)	(0.0296)	(0.0296)	(0.0296)	(0.0313)	(0.1316)
ln(avgPubFundO3$_t$)^2								-0.0157
								(0.0400)
ln(avgPrivFundO3$_t$)	-0.0294	-0.0371	-0.0293	-0.0318	-0.0294	-0.0298	-0.0370	-0.1665
	(0.0307)	(0.0316)	(0.0307)	(0.0307)	(0.0307)	(0.0307)	(0.0316)	(0.1531)
ln(avgPrivFundO3$_t$)^2								0.0405
								(0.0468)
ln(avgPhilFundO3$_t$)	0.0368	0.0152	0.0365	0.0336	0.0368	0.0370	0.0159	0.0869
	(0.0262)	(0.0276)	(0.0262)	(0.0262)	(0.0262)	(0.0262)	(0.0276)	(0.1542)
ln(avgPhilFundO3$_t$)^2								-0.0227
								(0.0482)
ln(nbArtFirst$_t$)	0.0902 ***	0.0905 ***	0.0927 ***		0.0902 ***	0.0904 ***	0.0933 ***	0.0932 ***
	(0.0165)	(0.0165)	(0.0178)		(0.0165)	(0.0165)	(0.0179)	(0.0179)
ln(nbArtLast$_t$)	0.0428 ***	0.0424 ***	0.0340 **		0.0428 ***	0.0424 ***	0.0346 **	0.0342 **
	(0.0162)	(0.0161)	(0.0172)		(0.0162)	(0.0162)	(0.0172)	(0.0172)
ln(nbArtMiddle$_t$)	0.0650 ***	0.0649 ***	0.0636 ***		0.0650 ***	0.0652 ***	0.0628 ***	0.0622 ***
	(0.0160)	(0.0160)	(0.0171)		(0.0160)	(0.0160)	(0.0172)	(0.0172)
ln(nbArticles$_t$)				0.1245 ***				
				(0.0218)				
ln(ImpactFact5$_t$)	0.4611 ***	0.4609 ***	0.4608 ***	0.4596 ***	0.4591 ***	0.4614 ***	0.4591 ***	0.4591 ***
	(0.0152)	(0.0152)	(0.0152)	(0.0152)	(0.0166)	(0.0152)	(0.0166)	(0.0166)

Variables	NSE-1	NSE-2	NSE-3	NSE-4	NSE-5	NSE-6	NSE-7	NSE-7sq
ln(avgAuthors$_t$)	0.5717 ***	0.5720 ***	0.5733 ***	0.5697 ***	0.5715 ***	0.5849 ***	0.5846 ***	0.5843 ***
	(0.0372)	(0.0372)	(0.0372)	(0.0363)	(0.0372)	(0.0396)	(0.0397)	(0.0397)
dFemale x ln(avgPubFundO3$_t$)		-0.1173					-0.1196	-0.4843
		(0.0984)					(0.0985)	(0.4183)
dFemale x ln(avgPubFundO3$_t$)^2								0.1202
								(0.1368)
dFemale x ln(avgPrivFundO3$_t$)		0.1539					0.1456	0.5388
		(0.1299)					(0.1307)	(0.8264)
dFemale x ln(avgPrivFundO3$_t$)^2								-0.1518
								(0.3125)
dFemale x ln(avgPhilFundO3$_t$)		0.2120 **					0.2033 **	0.2849
		(0.0860)					(0.0862)	(0.3692)
dFemale x ln(avgPhilFundO3$_t$)^2								-0.0263
								(0.1175)
dFemale x ln(nbArtFirst$_t$)			-0.0136				-0.0148	-0.0160
			(0.0464)				(0.0465)	(0.0465)
dFemale x ln(nbArtLast$_t$)			0.0768				0.0663	0.0672
			(0.0498)				(0.0503)	(0.0503)
dFemale x ln(nbArtMiddle$_t$)			0.0155				0.0223	0.0235
			(0.0464)				(0.0478)	(0.0479)
dFemale x ln(nbArticles$_t$)				0.0595				
				(0.0636)				
dFemale x ln(ImpactFact5$_t$)					0.0120		0.0115	0.0123
					(0.0409)		(0.0412)	(0.0412)
dFemale x ln(avgAuthors$_t$)						-0.1061	-0.0918	-0.0946
						(0.1078)	(0.1133)	(0.1136)
Year dummies	yes ***	yes ***	yes ***	yes ***	yes ***	yes ***	yes ***	yes ***
Constant	-1.1799 ***	-1.1981 ***	-1.1861 ***	-1.1035 ***	-1.1794 ***	-1.1761 ***	-1.1986 ***	-1.1993 ***
	(0.0843)	(0.0849)	(0.0844)	(0.0793)	(0.0843)	(0.0844)	(0.0852)	(0.1737)
Nb observations	10630	10630	10630	10630	10630	10630	10630	10630
Nb academics	1734	1734	1734	1734	1734	1734	1734	1734
Log likelihood	-3118.59	-3114.12	-3116.98	-3121.26	-3118.54	-3118.01	-3112.29	-3111.19
F	91.26 ***	77.98 ***	77.70 ***	96.65 ***	86.19 ***	86.24 ***	62.50 ***	50.43 ***
Avg nb of years	6.13	6.13	6.13	6.13	6.13	6.13	6.13	6.13
R^2 within	0.1487	0.1495	0.1490	0.1483	0.1487	0.1488	0.1497	0.1499
R^2 overall	0.1920	0.1456	0.1832	0.1865	0.1916	0.1751	0.1623	0.1508
R^2 between	0.2523	0.1560	0.2342	0.2416	0.2516	0.2076	0.1875	0.1650

Note: ***, **, * show significance at the 1%, 5% and 10% level respectively, standard errors in parentheses. Natural logarithm of the z-scores used for all continuous variables.

Table D5 – Regression results – Normalised citations (adding shareWomen) – Health and NSE fields 2006-2012

Variables	H-1b	H-7b	H-10b	NSE-1b	NSE-7b	NSE-10b
ln(avgPubFundO3$_t$)	-0.1412 ***	-0.1337 **	-0.1324 **	-0.0626	-0.0739	-0.0744
	(0.0535)	(0.0610)	(0.0609)	(0.0533)	(0.0585)	(0.0584)
ln(avgPrivFundO3$_t$)	0.0541	0.1394 **	0.1388 **	-0.0067	-0.0169	-0.0183
	(0.0605)	(0.0707)	(0.0707)	(0.0541)	(0.0558)	(0.0557)
ln(avgPhilFundO3$_t$)	0.0368	0.0775	0.0732	-0.0212	-0.0566	-0.0586
	(0.0426)	(0.0476)	(0.0476)	(0.0452)	(0.0479)	(0.0479)
ln(nbArtFirst$_t$)	0.0811 ***	0.0534 **		0.1111 ***	0.1081 ***	
	(0.0211)	(0.0247)		(0.0253)	(0.0274)	
ln(nbArtLast$_t$)	0.0141	0.0028		0.0612 **	0.0542 **	
	(0.0231)	(0.0267)		(0.0240)	(0.0254)	
ln(nbArtMiddle$_t$)	0.0722 ***	0.1102 ***		0.0839 ***	0.0858 ***	
	(0.0244)	(0.0286)		(0.0231)	(0.0249)	
ln(nbArticles$_t$)			0.1259 ***			0.1664 ***
			(0.0356)			(0.0328)
propNbArtFirst$_t$			0.0611 *			0.0864 ***
			(0.0325)			(0.0323)
propNbArtMiddle$_t$			0.0908 ***			0.0508 **
			(0.0315)			(0.0257)
ln(ImpactFact5$_t$)	0.8157 ***	0.8157 ***	0.8143 ***	0.4253 ***	0.4229 ***	0.4221 ***
	(0.0280)	(0.0334)	(0.0334)	(0.0241)	(0.0264)	(0.0264)
ln(avgAuthors$_t$)	1.0144 ***	1.0202 ***	1.0089 ***	0.5312 ***	0.5282 ***	0.5156 ***
	(0.0804)	(0.0934)	(0.0935)	(0.0506)	(0.0539)	(0.0539)
shareWomen$_t$	-0.0227 **	-0.0236 **	-0.0234 **	0.0035	0.0034	0.0026
	(0.0102)	(0.0102)	(0.0102)	(0.0097)	(0.0097)	(0.0097)
dFemale x	0.0389 **	0.0330 *	0.0331 *	-0.0096	-0.0125	-0.0141
shareWomen$_t$	(0.0191)	(0.0192)	(0.0192)	(0.0218)	(0.0226)	(0.0225)
dFemale x		-0.0286	-0.0285		0.0408	0.0438
ln(avgPubFundO3$_t$)		(0.1303)	(0.1302)		(0.1590)	(0.1586)
dFemale x		-0.2828 **	-0.2893 **		0.1307	0.1360
ln(avgPrivFundO3$_t$)		(0.1379)	(0.1379)		(0.2321)	(0.2314)
dFemale x		-0.2605 **	-0.2555 **		0.3307 **	0.3402 **
ln(avgPhilFundO3$_t$)		(0.1085)	(0.1084)		(0.1463)	(0.1460)
dFemale x		0.1069 **			0.0145	
ln(nbArtFirst$_t$)		(0.0471)			(0.0705)	
dFemale x		0.0417			0.0603	
ln(nbArtLast$_t$)		(0.0526)			(0.0762)	
dFemale x		-0.1410 ***			-0.0084	
ln(nbArtMiddle$_t$)		(0.0545)			(0.0674)	
dFemale x			-0.0294			0.0558
ln(nbArticles$_t$)			(0.0708)			(0.0952)
dFemale x			0.0967 *			0.0120
propNbArtFirst$_t$			(0.0573)			(0.0772)

Variables	H-1b	H-7b	H-10b	NSE-1b	NSE-7b	NSE-10b
dFemale x propNbArtMiddle$_t$			-0.1344 **			-0.1293 **
			(0.0588)			(0.0649)
dFemale x ln(ImpactFact5$_t$)		0.0004	-0.0059		0.0111	0.0185
		(0.0609)	(0.0609)		(0.0646)	(0.0646)
dFemale x ln(avgAuthors$_t$)		-0.0234	-0.0047		-0.0035	0.0666
		(0.1821)	(0.1828)		(0.1557)	(0.1583)
Year dummies	yes ***	yes ***	yes ***	yes ***	yes ***	yes ***
Constant	-2.0748 ***	-2.0476 ***	-2.1504 ***	-1.0289 ***	-1.0449 ***	-1.0966 ***
	(0.1597)	(0.1604)	(0.1604)	(0.1281)	(0.1292)	(0.1303)
Nb observations	6075	6075	6075	5798	5798	5798
Nb academics	1462	1462	1462	1524	1524	1524
Log likelihood	-1699.38	-1682.67	-1678.50	-1424.24	-1419.31	-1410.82
F	81.22 ***	56.45 ***	56.78 ***	339.11 ***	230.46 ***	230.10 ***
Avg nb of years	4.16	4.16	4.16	3.80	3.80	3.80
R^2 within	0.2310	0.2352	0.2363	0.1297	0.1312	0.1338
R^2 overall	0.2926	0.1362	0.1216	0.1777	0.0569	0.0618
R^2 between	0.3729	0.1188	0.1027	0.2055	0.0475	0.0526

Note: ***, **, * show significance at the 1%, 5% and 10% level respectively, standard errors in parentheses. Natural logarithm of the z-scores used for all continuous variables. In bold are highlighted the differences with Table 3 and Table 4.

Table D6 – Regression results – Normalised citations (adding shareWomen and its interactions) – Health fields 2006-2012

Variables	H-2c	H-3c	H-4c	H-5c	H-6c	H-7c	H-7c alt	H-7sqc
ln(avgPubFundO3$_t$)	-0.1478 ***	-0.1352 **	-0.1447 ***	-0.1446 ***	-0.1448 ***	-0.1484 ***	-0.1465***	-0.1051
	(0.0538)	(0.0535)	(0.0536)	(0.0535)	(0.0535)	(0.0538)	(0.0538)	(0.2109)
ln(avgPubFundO3$_t$)^2								-0.0136
								(0.0639)
ln(avgPrivFundO3$_t$)	0.0422	0.0503	0.0544	0.0503	0.0554	0.0402	0.0388	0.2104
	(0.0608)	(0.0605)	(0.0606)	(0.0605)	(0.0605)	(0.0608)	(0.0608)	(0.2794)
ln(avgPrivFundO3$_t$)^2								-0.0542
								(0.0918)
ln(avgPhilFundO3$_t$)	0.0192	0.0335	0.0288	0.0391	0.0359	0.0193	0.0199	0.0561
	(0.0431)	(0.0426)	(0.0426)	(0.0426)	(0.0426)	(0.0431)	(0.0431)	(0.1753)
ln(avgPhilFundO3$_t$)^2								-0.0181
								(0.0526)
ln(nbArtFirst$_t$)	0.0817 ***	0.0839 ***		0.0818 ***	0.0814 ***	0.0846 ***	0.0847***	0.0851 ***
	(0.0211)	(0.0211)		(0.0211)	(0.0211)	(0.0211)	(0.0211)	(0.0211)
ln(nbArtLast$_t$)	0.0133	0.0143		0.0145	0.0127	0.0137	0.0143	0.0118
	(0.0230)	(0.0230)		(0.0230)	(0.0231)	(0.0230)	(0.0230)	(0.0231)
ln(nbArtMiddle$_t$)	0.0735 ***	0.0706 ***		0.0721 ***	0.0737 ***	0.0722 ***	0.0714***	0.0733 ***
	(0.0244)	(0.0243)		(0.0244)	(0.0244)	(0.0244)	(0.0244)	(0.0244)
ln(nbArticles$_t$)			0.1245 ***					
			(0.0307)					

Variables	H-2c	H-3c	H-4c	H-5c	H-6c	H-7c	H-7c alt	H-7sqc
ln(ImpactFact5$_t$)	0.8164 ***	0.8127 ***	0.8170 ***	0.8140 ***	0.8156 ***	0.8127 ***	0.8124***	0.8125 ***
	(0.0280)	(0.0279)	(0.0279)	(0.0280)	(0.0280)	(0.0280)	(0.0280)	(0.0280)
ln(avgAuthors$_t$)	1.0127 ***	1.0171 ***	1.0167 ***	1.0143 ***	0.9826 ***	1.0008 ***	1.0148***	0.9991 ***
	(0.0804)	(0.0804)	(0.0791)	(0.0804)	(0.0817)	(0.0818)	(0.0803)	(0.0819)
shareWomen$_t$	0.1282 *	0.0154	0.0251	0.0423	0.2677 **	0.2829 *	0.1637**	-0.1534
	(0.0696)	(0.0480)	(0.0374)	(0.0303)	(0.1332)	(0.1552)	(0.0793)	(0.2844)
dFemale x shareWomen$_t$	0.0356 *	0.0362 *	0.0405 **	0.0378 **	0.0369 *	0.0324 *	0.0329*	0.0334 *
	(0.0191)	(0.0191)	(0.0191)	(0.0191)	(0.0191)	(0.0191)	(0.0191)	(0.0191)
shareWomen x ln(avgPubFundO3$_t$)	0.0398					0.0444	0.0457	0.0223
	(0.0402)					(0.0413)	(0.0413)	(0.1875)
shareWomen x ln(avgPubFundO3$_t$)^2								0.0094
								(0.0620)
shareWomen x ln(avgPrivFundO3$_t$)	-0.0746					-0.0533	-0.0567	0.2942
	(0.0549)					(0.0554)	(0.0553)	(0.3381)
shareWomen x ln(avgPrivFundO3$_t$)^2								-0.1184
								(0.1186)
shareWomen x ln(avgPhilFundO3$_t$)	-0.1056 ***					-0.0994 **	-0.0983**	0.2428
	(0.0398)					(0.0401)	(0.0401)	(0.2109)
shareWomen x ln(avgPhilFundO3$_t$)^2								-0.1155
								(0.0706)
shareWomen x ln(nbArtFirst$_t$)		0.0285				0.0291	0.0310	0.0282
		(0.0201)				(0.0202)	(0.0201)	(0.0202)
shareWomen x ln(nbArtLast$_t$)		0.0211				0.0219	0.0241	0.0199
		(0.0225)				(0.0233)	(0.0231)	(0.0234)
shareWomen x ln(nbArtMiddle$_t$)		-0.0806 ***				-0.0640 **	-0.0687***	-0.0659 ***
		(0.0240)				(0.0251)	(0.0245)	(0.0251)
shareWomen x ln(nbArticles$_t$)			-0.0399					
			(0.0290)					
shareWomen x ln(ImpactFact5$_t$)				-0.0623 **		-0.0454	-0.0501*	-0.0456
				(0.0273)		(0.0283)	(0.0278)	(0.0284)
shareWomen x ln(avgAuthors$_t$)					-0.1829 **	-0.0791		-0.0863
					(0.0837)	(0.0886)		(0.0887)
Year dummies	yes ***	yes ***	yes ***	yes ***	yes ***	yes ***	yes***	yes ***
Constant	-2.0344 ***	-2.0751 ***	-2.0311 ***	-2.0693 ***	-2.0196 ***	-2.0096 ***	-2.0334***	-2.1737 ***
	(0.1602)	(0.1596)	(0.1542)	(0.1596)	(0.1616)	(0.1623)	(0.1601)	(0.2758)
Nb observations	6075	6075	6075	6075	6075	6075	6075	6075
Nb academics	1462	1462	1462	1462	1462	1462	1462	1462
Log likelihood	-1693.24	-1690.12	-1702.03	-1695.94	-1696.22	-1682.26	-1682.78	-1679.05
F	69.59 ***	69.90 ***	85.99 ***	77.07 ***	76.43 ***	56.48 ***	58.80***	45.69 ***
Avg nb of years	4.16	4.16	4.16	4.16	4.16	4.16	4.16	4.16
R^2 within	0.2326	0.2334	0.2303	0.2319	0.2318	0.2353	0.2352	0.2361
R^2 overall	0.2927	0.2951	0.2924	0.2930	0.2945	0.2954	0.2950	0.2960
R^2 between	0.3723	0.3779	0.3722	0.3716	0.3795	0.3779	0.3756	0.3782

Note: ***, **, * show significance at the 1%, 5% and 10% level respectively, standard errors in parentheses. Natural logarithm of the z-scores used for all continuous variables. In bold are highlighted the differences with Table 3 and Table 4.

Table D7 – Regression results – Normalised citations (adding shareWomen and its interactions) – NSE fields 2006-2012

Variables	NSE-2c	NSE-3c	NSE-4c	NSE-5c	NSE-6c	NSE-7c	NSE-7sqc
ln(avgPubFundO3$_t$)	-0.0595	-0.0642	-0.0689	-0.0612	-0.0627	-0.0593	0.1774
	(0.0534)	(0.0534)	(0.0533)	(0.0533)	(0.0533)	(0.0535)	(0.2385)
ln(avgPubFundO3$_t$)^2							-0.0766
							(0.0754)
ln(avgPrivFundO3$_t$)	-0.0112	-0.0056	-0.0071	-0.0057	-0.0067	-0.0098	0.0355
	(0.0544)	(0.0541)	(0.0541)	(0.0541)	(0.0541)	(0.0545)	(0.2865)
ln(avgPrivFundO3$_t$)^2							-0.0141
							(0.0912)
ln(avgPhilFundO3$_t$)	-0.0139	-0.0215	-0.0247	-0.0216	-0.0211	-0.0142	0.4411 *
	(0.0455)	(0.0452)	(0.0452)	(0.0452)	(0.0452)	(0.0455)	(0.2383)
ln(avgPhilFundO3$_t$)^2							-0.1465 *
							(0.0749)
ln(nbArtFirst$_t$)	0.1110 ***	0.1088 ***		0.1114 ***	0.1111 ***	0.1090 ***	0.1089 ***
	(0.0253)	(0.0254)		(0.0253)	(0.0253)	(0.0254)	(0.0254)
ln(nbArtLast$_t$)	0.0604 **	0.0604 **		0.0619 ***	0.0612 **	0.0601 **	0.0588 **
	(0.0240)	(0.0240)		(0.0240)	(0.0240)	(0.0240)	(0.0240)
ln(nbArtMiddle$_t$)	0.0838 ***	0.0832 ***		0.0837 ***	0.0839 ***	0.0828 ***	0.0816 ***
	(0.0231)	(0.0231)		(0.0231)	(0.0231)	(0.0231)	(0.0232)
ln(nbArticles$_t$)			0.1725 ***				
			(0.0307)				
ln(ImpactFact5$_t$)	0.4255 ***	0.4255 ***	0.4237 ***	0.4250 ***	0.4253 ***	0.4255 ***	0.4243 ***
	(0.0241)	(0.0241)	(0.0241)	(0.0241)	(0.0241)	(0.0241)	(0.0241)
ln(avgAuthors$_t$)	0.5294 ***	0.5330 ***	0.5292 ***	0.5326 ***	0.5307 ***	0.5309 ***	0.5299 ***
	(0.0506)	(0.0506)	(0.0497)	(0.0506)	(0.0512)	(0.0513)	(0.0513)
shareWomen$_t$	-0.0454	0.0572	0.0425	0.0497 **	-0.0020	0.0236	0.1494
	(0.0669)	(0.0520)	(0.0364)	(0.0251)	(0.0823)	(0.1129)	(0.2755)
dFemale x	-0.0075	-0.0109	-0.0114	-0.0128	-0.0094	-0.0115	-0.0114
shareWomen$_t$	(0.0219)	(0.0221)	(0.0219)	(0.0219)	(0.0219)	(0.0222)	(0.0222)
shareWomen x	0.0267					0.0369	-0.1011
ln(avgPubFundO3$_t$)	(0.0453)					(0.0457)	(0.2083)
shareWomen x							0.0499
ln(avgPubFundO3$_t$)^2							(0.0708)
shareWomen x	-0.0351					-0.0350	-0.1265
ln(avgPrivFundO3$_t$)	(0.0512)					(0.0514)	(0.3068)
shareWomen x							0.0302
ln(avgPrivFundO3$_t$)^2							(0.1051)
shareWomen x	0.0521					0.0540	0.0984
ln(avgPhilFundO3$_t$)	(0.0391)					(0.0394)	(0.2473)
shareWomen x							-0.0143
ln(avgPhilFundO3$_t$)^2							(0.0815)

Variables	NSE-2c	NSE-3c	NSE-4c	NSE-5c	NSE-6c	NSE-7c	NSE-7sqc
shareWomen x		-0.0278				-0.0261	-0.0269
ln(nbArtFirst$_t$)		(0.0233)				(0.0234)	(0.0234)
shareWomen x		-0.0035				-0.0052	-0.0027
ln(nbArtLast$_t$)		(0.0229)				(0.0235)	(0.0235)
shareWomen x		-0.0149				-0.0152	-0.0131
ln(nbArtMiddle$_t$)		(0.0224)				(0.0239)	(0.0239)
shareWomen x			-0.0310				
ln(nbArticles$_t$)			(0.0270)				
shareWomen x				-0.0441 **		-0.0444 **	-0.0450 **
ln(ImpactFact5$_t$)				(0.0221)		(0.0224)	(0.0224)
shareWomen x					0.0035	0.0113	0.0064
ln(avgAuthors$_t$)					(0.0516)	(0.0549)	(0.0551)
Year dummies	yes ***	yes ***	yes ***	yes ***	yes ***	yes ***	yes ***
Constant	-1.0312 ***	-1.0267 ***	-0.9340 ***	-1.0335 ***	-1.0280 ***	-1.0308 ***	-1.5398 ***
	(0.1283)	(0.1282)	(0.1215)	(0.1281)	(0.1288)	(0.1290)	(0.3022)
Nb observations	5798	5798	5798	5798	5798	5798	5798
Nb academics	1524	1524	1524	1524	1524	1524	1524
Log likelihood	-1422.57	-1423.06	-1424.46	-1421.54	-1424.24	-1418.57	-1414.57
F	288.45 ***	288.65 ***	356.50 ***	321.30 ***	320.21 ***	231.75 ***	186.84 ***
Avg nb of years	3.80	3.80	3.80	3.80	3.80	3.80	3.80
R^2 within	0.1302	0.1301	0.1297	0.1305	0.1297	0.1314	0.1326
R^2 overall	0.1768	0.1776	0.1764	0.1785	0.1777	0.1776	0.1802
R^2 between	0.2026	0.2062	0.2041	0.2039	0.2056	0.2016	0.2051

Note: ***, **, * show significance at the 1%, 5% and 10% level respectively, standard errors in parentheses. Natural logarithm of the z-scores used for all continuous variables. In bold are highlighted the differences with Table 3 and Table 4.