# The Effect of Gender and Funding on Research Performance

**Mohammad Soleymanifar**

**A Thesis**
**In the Department of**
**Concordia Institute of Information Systems Engineering**

**Presented in Partial Fulfillment of the Requirements**
**For the Degree of**
**Master of Applied Science (Quality Systems Engineering)**

**At Concordia University**
**Montreal, Quebec, Canada**

**May 2024**

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: Mohammad Soleymanifar

Entitled: The Effect of Gender and Funding on Research Performance

and submitted in partial fulfillment of the requirements for the degree of

Master of Applied Science (Quality Systems Engineering)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

______________________________ Chair
Dr. Farnoosh Naderkhani

______________________________ Examiner
Dr. Chun Wang

______________________________ Examiner

______________________________ Thesis Supervisor(s)
Dr. Andrea Schiffauerova

______________________________ Thesis Supervisor(s)
Dr. Ashkan Ebadi

Approved by ______________________________
Dr. J. Yan (GPD) Chair of Department or Graduate Program Director

______________________________
Dr. Mourad Debbabi Dean of ENG & Computer Science

**Abstract**

The Effect of Gender and Funding on Research Performance

Mohammad Soleymanifar

In spite of various improvements and increasing involvement of female researchers in scientific activities in recent years, the gender gap still persists and women remain greatly underrepresented in technology, engineering, and computer science fields. This thesis attempts to shed some light on the effect of gender and funding on research output of Canadian researchers in natural sciences and engineering. In this research, using NSERC and Scopus data from 1982 to 2018, we apply descriptive statistical analysis and regression analysis to study the influence of funding and gender on the quantity of published journal papers and their scientific impact. The study concludes that funding has a positive impact on both the number of papers published and the number of citations received by their respective author. However, we also observe that as career age of authors increases, researchers become less productive, they publish less papers and their citation counts slightly diminish with time as well, even though their funding amounts typically increase. In terms of gender, even though we find that female researchers are indeed greatly underrepresented and receive lower amounts of funding than their male counterparts, they produce on average similar number of articles with similar scientific impact. This means that female researchers can generate comparable research output with lower research costs compared to male researchers and are thus more efficient in their research production. These findings suggest that governmental funding agencies should introduce more effective gender-related funding strategies and greater support for early-career researchers.

**Keywords:** Funding, Gender, Regression Analysis, Impact, Quantity, Productivity, Male, Female

**Acknowledgments**

It is with immense pleasure and gratitude that I take this opportunity to give my deepest appreciation for my family who have always been by my side through the most intense phases of my life and continue to do so. I would not be where I am today if it were not for their support.

Next, I would like to thank my supervisor Professor, Andrea Schiffauerova, for her invaluable support and especially for putting her trust in me and giving me this opportunity to do my studies under her supervision. She has always supported me and my work, encouraging me to keep going even when facing challenges in certain phases of my research.

Lastly, I would also like to thank Dr. Ashkan Ebadi for his input and recommendations on certain aspects of my research and for providing the gender information dataset for NSERC awardees.

## Table of Contents

## List of Tables

## List of Figures

**Table of Abbreviations**

| Abbreviation | Definition |
| --- | --- |
| NSERC | The Natural Sciences and Engineering Research Council of Canada |
| SPSS | Statistical Product and Service Solutions |
| CIHR | Canadian Institutes of Health Research |
| NRC | National Research Council Canada |
| MRC | Medical Research Council Canada |
| NIH | National Institutes of Health |

# 1. Introduction

Although women have greatly contributed to science and technology, produced remarkable discoveries, and are currently leading ground-breaking research across the world, female researchers remain underrepresented in technology, engineering, and computer science (Hill et al., 2010). Despite covering half of the global workforce, women are also underrepresented in high academic positions, have on average access to fewer research resources, and receive lower salaries (O'Dorchai et al., 2009). In spite of improvements in recent years and increasing growth in female researchers' involvement in scientific activities, these problems persist (Hajibabaei et al., 2022 a, b). Consequently, there is a growing interest of the researchers in the study of various aspects of gender differences in scientific performance. Gender disparities in science – measured by various factors such as the presence of each gender among research teams, number of publications, citation counts, and funding amounts – have been investigated across different countries and disciplines (Shen 2013, Lariviere et al., 2013, Ebadi and Schiffauerova 2016, Witteman et al., 2019, Hajibabaei et al., 2022). Understanding the role of gender in scientific productivity, research impact and collaboration thus lies in the heart of this thesis.

An important aspect of scientific performance evaluation is research productivity (proxied by scientific publications) and research impact (proxied by citations), where we are interested in the difference in research performance based on gender. In general, male researchers have been mostly found more productive (in terms of scientific publications) than female researchers (Sax et al., 2002, Stack 2004, Puuska 2010, Hunter et al., 2010, Lariviere et al., 2011), even though some studies found no difference between the genders (Lewison 2001, Gallivan et al., 2006, Tower et al., 2011, and Mauleon et al., 2008). Similarly, some researchers found a lower number of citations of articles authored by female authors (Lariviere et al., 2011, Hunter et al., 2010), while others found no clear relationship between gender and the number of citations (Mauleon et al., 2008, Cole et al., 1984, Ledin et al., 2007, Copenheaver et al., 2010). Hence, the findings of various studies are not consistent.

Another major theme this thesis is going to tackle is research funding. Every year governments all over the world spend considerable amounts of money on funding various research projects through different funding programs with an aim of improving the scientific potential of the country. Various aspects of the research funding have been analyzed at different levels, including the gender aspects. Several studies showed that female researchers receive lower amounts in funding on average compared to male researchers (e.g. Hajibabaei et al., 2022, Beaudry et al., 2014, Eloy et al., 2013, Lariviere et al., 2011, Stack 2004, Feldt 1986), while others did not observe such differences between the genders (e.g. Waisbren et al., 2008, Zuckerman 1987). Since funding is not an unlimited resource, the effectiveness of such investment and its impact on society should be considered and thoroughly analyzed. This should help funding agencies with the decision-making in terms of the fund allocation to the most productive researchers and to the most important and impactful research programs. There has been a growing amount of research focusing on the relation between the amounts of funding received and the scientific production of funded researchers, where most of the studies conclude that funding has a great impact on the research productivity (e.g. Ebadi and Schiffauerova 2016, Beaudry et al., 2014).

Even though there are research studies analyzing the impact of funding on the output of funded research programs, to our knowledge, the difference in effectiveness of the funds awarded to male or female researchers has not been studied yet. There is so far no research study which would include the gender factor in the equation, and which would estimate the different and interacting effects of funding and gender on scientific performance. And this is the main purpose of this thesis where we aim to shed light on various factors influencing research performance, while specifically focusing on the factors of research funding and gender. We use the funding and publications data related to Canadian researchers in the science, technology, and engineering fields and with the help of descriptive statistics and regression models we identify the gender gap, characterize various aspects of the relationships between funding, gender, and performance, and evaluate the impact of funding and gender on the number of publications and citations.

The rest of the thesis is organized as follows: The next chapter introduces relevant literature, then we describe the data and methodology, carry out the descriptive analysis while presenting its results, then we introduce the regression models and present the results, and, finally, the conclusions are presented, and limitations and suggestions for future research are discussed.

# 2. Literature Review

In this section the relevant literature is discussed, and different papers are analyzed and compared based on their area of focus. First, the concept of funding is introduced and the literature addressing the impact of funding on research performance is examined. Afterwards the role of collaboration is examined and relevant research papers reviewed. Finally, the gender disparity issue is explained, the relevant literature reviewed and its importance discussed. The literature review of each section is summarized in a corresponding table. Finally, the results are discussed and objectives for determining the scope of this analysis presented.

## 2-1. Funding

It has been shown by many studies that funding plays a crucial role in scientific development (e.g. Martin 2003). It is viewed as a key factor that affects research output since it provides better access to available research resources (Lee and Bozeman 2005). However, evaluation of the relationship between funding and the quantity (number of publications) and the scientific impact (number of citations) of the research output has proven to be a challenge where several methodologies (e.g. bibliometrics, statistical analysis, interviews, surveys, data and text mining) have been used for this purpose (Ebadi and Schiffauerova 2016). The following section reviews the literature which studies the impact of funding received by researchers for their scientific programs on the number of publications which resulted from this research and on the number of citations which these publications received.

## 2-2. Number of Publications and Citations

Several studies analyzed different aspects of the connection between funding and scientific performance at various levels using different methods.

In many instances, funding proved to increase the number of publications. Tahmooresnejad et al. (2014) in a study focusing on the role of public funding in the US and Canada for researchers active in the nanotechnology field found that funding has a positive effect on researchers' productivity in the United States. In other words, funded researchers produce more publications on average every year compared to non-funded researchers. A similar effect was observed for researchers in Canada. Even though in their data the number of researchers in the US (33,655) was ten times higher than the number of researchers in Canada (3,684), both countries showed similar patterns. Ebadi and Schiffauerova (2015 and 2016) concluded that funding promotes productivity among Canadian researchers. Their data covered 47,789 and 36,124 researchers from 1996-2010 respectively. Heyard and Hottenrot (2021), in another study in Switzerland covering publications from 2005 to 2019, analyzed the impact of funding on scientific output and concluded that funding boosts research productivity. A recent paper by Sattari et al. (2022) covering publications for the period of 1985-2020, studied the effect of funding on the output of research in the US and found that funding heavily affects productivity.

In several cases, funding did not appear to affect the number of publications. A study of Jacob and Lefgren (2011), focusing on publications from 1980-2000 in the US, revealed that the number of papers was not affected by funding. A study by Agarwal and Tu (2021) covering applicants—both successful and unsuccessful in attaining grants from NIH (National Institutes of Health) in the US in the year 2000—uncovered that applicants who received funding from NIH were not significantly more productive compared to applicants who did not receive any funding from NIH.

A study by Roshani et al. (2021) focusing on publications from 2015, found varying results across different fields. In computer science there are more funded publications than unfunded publications. An equal number of funded and unfunded publications in medicine was observed. Lastly, economics journals published fewer funded papers than unfunded papers. It can be concluded that most studies found that funding has a positive impact on the number of publications. The results of these studies are presented in Table 1.

| Authors | Publishing Year | Data | Area of Focus | Region | Results |
|---|---|---|---|---|---|
| B. Jacob, L. Lefgren | 2011 | 1980-2000 | Impact of research grant funding on scientific productivity | US | Funding has no impact on the number of papers |
| L. Tahmooresnejad, C. Beaudry, A. Schiffauerova | 2014 | 33,655 researchers (US) 3,684 researchers (Canada) 1996-2005 | Role of Public Funding in Nanotechnology Scientific Production | US and Canada | Funding has a positive impact on the number of papers in the US and Canada |
| A. Ebadi, A. Schiffauerova | 2015 | 47,789 researchers 1996-2010 | Analysis of the Impact of Funding on Scientific Productivity | Canada | Funding has a positive impact on the number of papers |
| A. Ebadi, A. Schiffauerova | 2016 | 36,124 researchers 1996-2010 | Statistical Analysis of Funding and Other Factors | Canada | Funding has a positive impact on the number of papers |
| R. Heyard, H. Hottenrot | 2021 | 2005-2019 | Impact of funding on research outcomes | Switzerland | Funding has a positive impact on the number of papers |
| S. Roshani, M. Bagherylooieh, M. Mosleh, M. Coccia | 2021 | 2015 | Relationship between research funding and productivity | Global | Funding has a positive impact on the number of papers in 2 out of 3 fields and no impact in the other field |
| R. Agarwal, W. Tu | 2021 | 55,000 publications 2000 | Funding and research productivity | US | Funding has no impact on the number of papers |
| R. Sattari, J. Bae, E. Berkes, B. Weinberg | 2022 | 1985-2020 | Effect of funding on research output | US | Funding has a positive impact on the number of papers |

**Table 1 - Publications Focusing on the Number of Publications**

A few other studies analyzed the impact of funding on the scientific impact of publications measured by citation counts.

In many instances, funding proved to increase the number of citations. Gok et al. (2014) investigated the effect of research funding on scientific output. Their data included 242,406 funded researchers from Belgium, Denmark, Netherlands, Norway, Switzerland, and Sweden for the period of 2009-2011. According to their results, funding is highly linked to average citation counts. Another paper by Wang and Shapira (2015), covering 89,605 publications from 2008-2009 from researchers around the world, investigated the impact of research sponsorship on research output in the nanotechnology field. Their results also concluded that funding leads to greater citation numbers. A study by Roshani et al. (2011) focusing on publications published in the year 2015 in the scientometrics field concluded that in all three fields of computer science, medicine, and economics, funded publications receive more citations on average compared to unfunded publications. A recent study by Gonzalez and Gonzalez (2023) analyzed different citation patterns among funded researchers on a global scale. This study solely focused on publications from 2016. According to their findings, researchers who are funded get more citations.

In other studies, funding did not appear to affect the number of citations. A study by Agarwani and Tu (2021) investigated the impact of funding on research output for researchers who were granted funding from NIH in the year 2000 in the US. Based on their findings, funding does not necessarily lead to an increase in citations.

The study by Tahmooresnejad et al. (2014), found varying results in two different countries. While funding has a positive effect on the number of citations for researchers in the US, there was no observed impact from funding on citation counts in Canada.

To summarize, most papers found that funding positively impacts citation counts with a few exceptions where no impact was observed. More research is needed to understand the effects funding has on the quality of the resulting scientific outcome. The results from these studies are presented in Table 2.

| Authors | Publishing Year | Data | Area of Focus | Region | Results |
|---|---|---|---|---|---|
| A. Gok, J. Rigby, P. Shapira | 2014 | 242,406 researchers<br><br>2009-2011 | Impact of Research Funding on Scientific Outputs | Belgium, Denmark, Netherlands, Norway, Switzerland, and Sweden | Funding has a positive impact on the number of citations |
| L. Tahmooresnejad, C. Beaudry, A. Schiffauerova | 2014 | 33,655 researchers (US)<br>3,684 researchers (Canada)<br><br>1996-2005 | Role of Public Funding in Nanotechnology Scientific Production | US and Canada | Funding has a positive impact on the number of citations in the US but no impact is shown in Canada |
| J. Wang, P. Shapira | 2015 | 89,605 publications<br><br>2008-2009 | Relationship Between Sponsorship and Publication Impact on Nanotechnology | Global | Funding has a positive impact on the number of citations |
| A. Ebadi, A. Schiffauerova | 2016 | 36,124 researchers<br><br>1996-2010 | Statistical Analysis of Funding and Other Factors | Canada | Funding has a positive impact on the number of citations |
| S. Roshani, M. Bagherylooieh, M. Mosleh, M. Coccia | 2021 | 2015 | Relationship between research funding and citation-based performance | Global | Funding has a positive impact on the number of citations |
| P. Gonzalez, M. Gonzalez | 2023 | 2016 | Citation differences across research funding | Global | Funding has a positive impact on the number of citations |

**Table 2 – Publications Focusing on the Number of Citations**

**2-3. Collaboration**

In today's world, geographical limitations are not as impactful as they once were in limiting researchers' ability to collaborate from different regions as they are constantly working together on a global scale to gain knowledge. The result of scientific knowledge progression is technological developments (Subramanyam 1983). With the growth of complexity in scientific projects, researchers need to collaborate (Katz and Martin 1997, Wood and Gray 1991) and form diverse teams to better address challenges (Bennet and Gadlin 2012). Collaboration brings several advantages to scientific activities (Ubfal and Maffioli 2011), and can be used to attain available skills that further develop new expertise (Lee and Bozeman 2005). The importance of scientific collaboration is thus acknowledged in scientific communities (Wray 2006).

Several studies have been conducted in the literature analyzing the role and importance of collaboration in scientific activities. In some studies, funding led to a higher collaboration rate amongst researchers. This relationship was also true in the opposite way, as collaboration led to securing more funding.  For example, a study by Ebadi and Schiffauerova (2013) showed that funded researchers around the world tend to collaborate more often than non-funded researchers. Shin et al. (2022) found that funding helps researchers living in different regions of Europe collaborate more often, while Bornstein and Bordons (2020) found varying results across different fields in Spain. While in three out of seven disciplines funding positively impacted collaboration, no impact was observed in the other four disciplines.

Ebadi and Schiffauerova (2015) revealed that collaboration helps researchers receive more funding in Canada, while Davis et al. (2022) reached similar results for New Zealand.

Based on the results, it can be concluded that there is a relationship between funding and collaboration. The results of these studies are presented in Table 3.

| Authors | Publishing Year | Data | Area of Focus | Region | Results |
|---|---|---|---|---|---|
| A. Ebadi, A. Schiffauerova | 2013 | N/A | Impact of Funding on Scientific Output and Collaboration | Global | Funding has a positive impact on collaboration |
| A. Ebadi, A. Schiffauerova | 2015 | 228,417 researchers<br><br>1996-2010 | Funding and Scientific Collaboration Networks | Canada | Funding has a positive impact on collaboration |
| A. Ebadi, A. Schiffauerova | 2015 | 174,773 researchers<br><br>1996-2010 | The Role of Collaboration and Networking in Funding | Canada | Collaboration has a positive impact on receiving funding |
| B. Bornstein, M. Bordons | 2020 | 12,461 publications<br><br>2010-2014 | Role of Funding and Collaboration in Higher Research in Several Disciplines | Spain | Funding has a positive impact on collaboration in 3 out of 7 fields and no impact in the other fields |
| B. Davis, J. Gush, S. Hendy, A. Jaffe | 2022 | 7,854,938 publications<br><br>1996-2018 | Research funding and collaboration | New Zealand | Collaboration has a positive impact on receiving funding |
| H. Shin, K. Kim, D.Kogler | 2022 | 3,077,225 publications<br><br>2008-2017 | Scientific collaboration and research funding | Europe | Funding has a positive impact on collaboration |

**Table 3 - Publications Focusing on Collaboration**

## 2-4. Gender

Despite observed improvements in gender equity and the continued growth of women's presence in leading positions across well-developed countries (Unesco 2018), gender disparity still exists (Nelson and Rogers 2003, Shaw and Stanton 2012, West et al., 2013). In recent years, studying gender disparity in scientific performance—measured by various factors such as the presence of each gender in research teams, number of publications, citation counts, and funding amounts—has gained popularity amongst researchers and has been studied across different countries and different disciplines (e.g. Shen 2013, Lariviere et al., 2013, Ebadi and Schiffauerova 2016, Witteman et al., 2019, Hajibabaei et al., 2022).

Several studies have shown that there is persistent gender bias in many aspects such as getting hired and promoted (Moss-Racusin et al., 2012, Nelson and Rogers 2003, M.W. Nielsen 2016), salaries (Shen 2013), receiving funding for research (Witteman et al., 2019), scientific impact (Lariviere et al., 2013), collaboration (Uhly et al., 2015, Zeng et al., 2016), and peer reviews (Murray et al., 2019).

There is also evidence that shows male and female collaboration patterns are different in scientific activities (Lariviere et al., 2013, Jadidi et al., 2018, Abramo et al., 2013, Sonnert and Holton 1995).

Hence, it is critical to study and understand the role of gender in scientific productivity, impact, and collaboration.

Some studies showed that men were producing more papers than women. For example, a study by Holman et al. (2018) found that female researchers worldwide in the STEM (Science, Technology, Engineering, and Mathematics) fields published fewer papers than male researchers. Additionally, two consecutive papers by Hajibabaei et al. (2022) found that female Canadian researchers in the artificial intelligence field were less productive than males despite improvements in the gender gap.

Other studies showed that men were both producing more papers and being cited more than women. For example, a study by Lariviere et al. (2011) found that women in Quebec secured less funding, produced fewer papers, and received fewer citations than men. Another study by Cowley et al. (2020), focusing on elite researchers in the US, Canada, and South Africa, revealed that men are more productive and get cited more often compared to women.

Additional studies showed that while men were more productive than women, both men and women were equally cited. For example, a study by Ebadi and Schiffauerova (2016), focusing on Canadian researchers, found male researchers are more productive than female researchers. However, both genders received an equal number of citations.

Even further studies showed that men collaborate more than women. A paper by Abramo et al. (2013), focusing on researchers in Italy, concluded that women struggle to collaborate on an international level compared to men. Another study, conducted by Hajibabaei et al. (2022), found that female Canadian researchers collaborate less than their male peers.

To summarize, some studies found that female researchers are less productive, get fewer citations, and are less collaborative while others found no difference between males and females in these aspects. The results of these studies are presented in Table 4.

| Authors | Publishing Year | Data | Area of Focus | Region | Results |
|---|---|---|---|---|---|
| V. Lariviere, E.V-Gagne, C. Villeneuve, P. Gelinas, Y. Gingras | 2011 | 13,636 researchers<br><br>2000- 2008 | Gender differences in funding, productivity, and impact | Quebec | Women receive less funding, are less productive, and receive fewer citations |
| G. Abramo, C.A. D'angelo, G. Murgia | 2013 | 43,379 researchers<br><br>2006-2010 | Gender differences in research collaboration | Italy | Women collaborate less than men internationally |
| A. Schiffauerova, A. Ebadi | 2016 | 173,773 researchers<br><br>1996-2010 | Gender differences in research output, funding, and collaboration | Canada | Men are more productive, but citation counts are the same for both genders |
| L. Holman, D. Stuart-Fox, C. Hauser | 2018 | 36,000,000 researchers<br><br>2002-2017 | Gender gap in science | Global | Women are less productive in STEM fields |
| C. Sa, S. Cowley, M. Martinez, N. Kachynska, E. Sabzalieva | 2020 | 943 elite researchers<br><br>2000-2020 | Gender Gaps in research productivity | US, Canada, and South Africa | Men are more productive and receive more citations |
| A. Hajibabaei, A. Schiffauerova, A. Ebadi | 2022 | 39,679 publications 2000-2019 | Women and key positions in collaboration networks | Canada | Women are less productive in the AI field |
| A. Hajibabaei, A. Schiffauerova, A. Ebadi | 2022 | 39,679 publications 2000-2019 | Gender-specific patterns in the AI scientific ecosystem | Canada | The gender gap between men and women is getting smaller in AI research |

**Table 4 - Publications Focusing on Gender**

## 2-5. Literature Review Conclusion and Objectives

To conclude the literature review section, we first recognize that there is a research interest in examining the effects of funding on the scientific outcome of the funded research projects, and that there is also a great interest in studying various aspects of gender disparity in scientific activities. This is evidenced by the number of papers presented in the preceding section. Reviewed studies confirm the existence of the gender gap and provide evidence of differences between male and female researchers, their different performance, different collaboration patterns and different funding success. Also, we note that, in general, the previous research investigating the influence of funding and collaboration on the scientific production finds positive effects. Nevertheless, to our knowledge, there is no research study which would combine these factors, and which would attempt to understand the difference in effectiveness of the funding awarded to male and to female researchers, which is a research gap which this thesis seeks to address. Hence, in this study, we are going to shed some light on the interaction between funding and gender, and the combined effects they have on the quantity of the scientific production (publications) and on the scientific impact of that production (citations). Consequently, we defined the following objectives:

- Objective #1: To investigate the impact of gender and funding on the quantity of scientific production
- Objective #2: To investigate the impact of gender and funding on the value of that production (scientific impact)

In the next section, the data used for the analysis is introduced and the methodology of the analysis is discussed in detail followed by the descriptive analysis results.

# 3. Data

## 3-1. Data preparation

In this section we discuss the databases and the procedure for preparing the data for analysis.

### 3-1-1. NSERC

For this research, we have decided to focus on the data related to Canadian researchers funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) between 1982 and 2018, which covers 337,330 grant installments (including separate grant amounts that correspond to the same grant application and the same award). The reason for choosing this period lies in the fact that in the years preceding this period the NSERC data is not as comprehensive. NSERC was chosen since federal funding is the main source of grants for university research in Canada. Another reason was that the NSERC data is available to the public. The Natural Sciences and Engineering Research Council of Canada (NSERC) is the major federal funding agency that is responsible for funding research in natural sciences and engineering in Canada. NSERC directly funds university professors and students as well as Canadian organizations to perform research and training. With funding from the Canadian Government, NSERC supports the research of over 41,000 students, trainees, and professors at universities and colleges in Canada with an annual budget of CA$1.43 billion in 2023 (NSERC). We collected the data related to all the grants in all the fields awarded to researchers affiliated with Canadian universities from 1982 to 2018. The final dataset which is our first major dataset contains information such as the year in which the grant installment was received, the affiliation of the researchers, the award amount, and the award duration.

### 3-1-2. Scopus

The publications' information and authorship data for the above-mentioned NSERC awardees were collected from Elsevier's Scopus using their first name, last name, and affiliation. Elsevier's Scopus was chosen since it provides accurate and comprehensive information regarding author affiliations. Scopus covers a wide variety of fields and was deemed more appropriate to our needs compared to others such as Microsoft Academic Research, Google Scholar, etc. This database was not as complete in its early years, especially before 1996, but has improved significantly (Ebadi and Schiffauerova, 2016).

The Scopus database covers 37,000 journal titles from approximately 11,678 publishers, of which 34,346 are peer-reviewed journals in top-level subject fields such as life sciences, social sciences, physical sciences, and health sciences. It covers several types of sources which are: book series, journals, trade journals, and conference proceedings. All journals covered in the Scopus database are evaluated for sufficiently high quality every year according to four types of numerical impact measures for each title (Scopus).

The Scopus database contains data for researchers' publications such as affiliation, publishing year, funding info, language, references, citations, etc.

After collecting the publication data for each author in the NSERC dataset from the Scopus website using their first name, last name, and affiliation, we have merged the entire data into one complementary dataset to form our second major dataset which contains information on the publications such as the publication year, the team members, the institution, citation counts and more. Some authors from the NSERC dataset could not be found on the Scopus website.

The dataset contains information on publications for the period of 1834-2023 and a total number of 2,434,453 publications. Some funding details are also available in the Scopus dataset. However, since many researchers do not mention the source of funding in their articles, the funding details for only 809,972 of the mentioned publications were found in Scopus. The funding information contained in Scopus was hence not considered reliable and useful for our analysis. Consequently, we only use the funding details from the NSERC dataset in this thesis.

### 3-1-3. Gender

As one of the main objectives of our analysis, we are going to study the differences in funding patterns and research output for each gender. A machine learning program developed by Ashkan Ebadi is used to determine the gender of each researcher in the NSERC dataset based on their first name, last name, and region of origin. In the next steps, we focus on the number of publications based on year and province for each gender. We also focus on funding details and collaboration for each gender.

### 3-1-4. Joining the NSERC and Scopus Datasets

To get the publications' information for the NSERC awardees from the Scopus dataset we matched the two datasets together. To accomplish this, we first had to find the author IDs in the Scopus dataset. Fuzzy matching was used for this matching process between the NSERC and Scopus datasets using the author's first name, last name, and affiliation. To deal with mismatches and two matches for the same entry (same name with different affiliations, as an example) we decided to delete the duplicate matches. This resulted in losing a small number of entries in the final dataset. Next, we added the Scopus author IDs to the NSERC dataset. After that, we used the Scopus author IDs as a common link between the two datasets to get the publication information from the Scopus dataset for the researchers funded by NSERC for our analysis.

### 3-2. Descriptive Analysis of the Data

In this section, the results of the descriptive analysis are presented, compared, and discussed based on certain categories. First, we are going to do a descriptive analysis of the NSERC data. This includes 451 individual research subjects in 90 areas of research. In this dataset, we have 879 institutions, 4372 departments, and 189 funding programs across Canada. For the scope of this project, we have decided to exclude scholarships and fellowships and only focus on academic institutions such as universities. The dataset contains the information for 29,131 researchers. In this dataset, we have 10 provinces in Canada. Yukon, Nunavut, and Northwest Territories are not included since there is not much data available for these regions. After that, we are going to do a descriptive analysis of the Scopus data, as well, which contains publications' information for over 2 million articles such as citation data, team members, and more.

### 3-2-1. Descriptive Analysis of Funding

In this part of the analysis, we are going to look at the funding-related data and discuss the results. Most of the NSERC grants are awarded for a period longer than one year, but they are disbursed to the researchers in annual installments. In this study, we considered all the funding data in the form of these annual installments received by each researcher as a part of each award. Hence, for example, for a 5-year grant we considered five installments received in each respective year. Considering the accurate timing of the fund disbursement enabled us to better assess their potential impact on the research performance in the following years. It is also important to note that in the NSERC dataset the total funding amount for team grants is allocated to the team leader, whereas in reality the funding amount is divided between team members who may be based in different provinces and affiliated with different institutions. To combat this issue, we assumed that each team member received an equal portion of the total amount and we assigned this portion to their respective province and/or institution. The average funding amount from NSERC per year per researcher is presented in Figure 1.

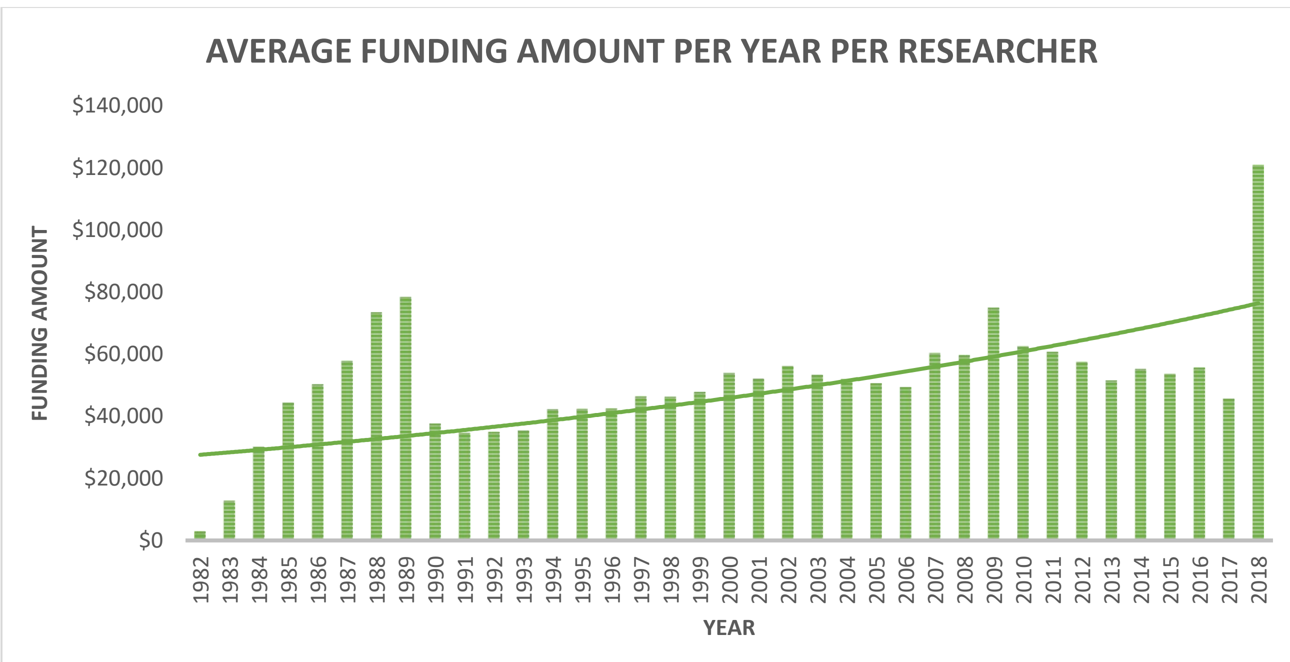


**Figure 1 - Average Funding Amount Per Year Per Researcher (NSERC)**

Figure 1 shows that during the 1980s, the average funding amount increased significantly. There was a drop in 1990, potentially because the number of researchers nearly doubled that year. After that, the average funding amount continued to steadily grow. In 2018 there was a large amount of funding but a relatively low number of researchers present in the dataset which may have caused the spike in the average funding amount. To get a better understanding of funding patterns, we will look at funding allocation per province in the next steps of our analysis.

The total number of grant installments for each province is presented in Figure 2.

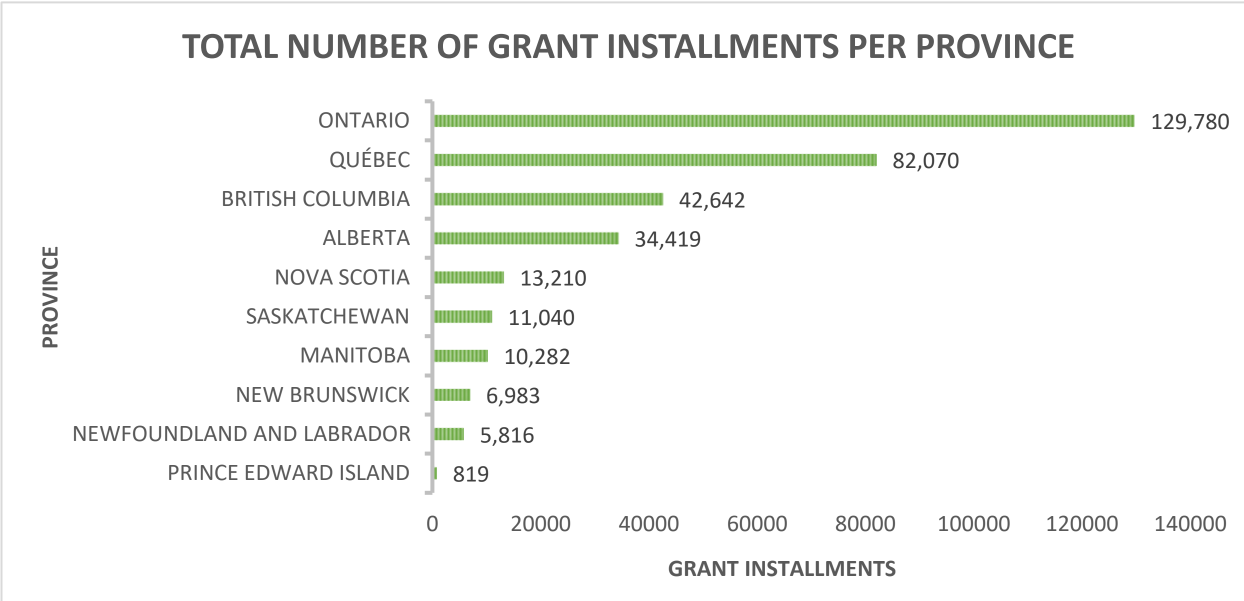


**Figure 2 - Total Number of Grant Installments Per Province**

Figure 2 shows that most of the grant installments are awarded to researchers in the province of Ontario which covers 38.49% of the data. This is to be expected since Ontario alone covers around a third of the total population of the country. This is followed by the province of Quebec which covers 24.34% of the grant installments in this dataset, then the province of British Columbia which covers 12.65% of the data, followed by the rest of the provinces. The top four provinces of Ontario, Quebec, British Columbia, and Alberta alone account for 80% of the total number of grant installments.

Since the total number of grant installments is directly linked to the total number of researchers, it is better to also compare the average annual funding amount per researcher for each province to better assess their performance.

The average annual funding amount per researcher by province is shown in Figure 3.

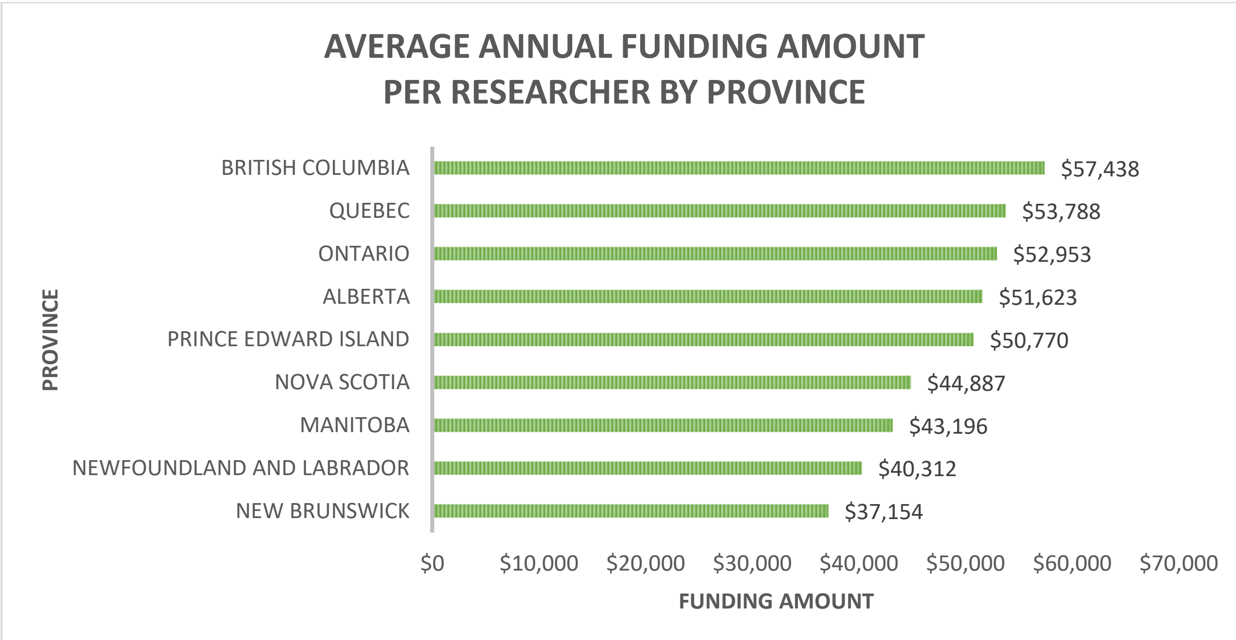


**Figure 3 - Average Annual Funding Amount Per Researcher by Province**

Figure 3 shows that researchers based in British Columbia, Quebec, Ontario, and Alberta are receiving the greatest amount of annual funding on average. This is in line with past findings in the literature (e.g. Ebadi and Schiffauerova 2013). If we compare Figure 2 and Figure 3, we can see that although researchers based in Ontario received the greatest number of grant installments, the average funding amount is lower than British Columbia and Quebec. This could be because Ontario has the highest number of researchers and universities but since British Columbia and Quebec have lower populations and fewer universities, the average award amount granted to their researchers is higher.

To better see how the funding is allocated within each province, it is important to go one step deeper and compare funding between institutions. It is worth mentioning that we are only looking at the top 10 institutions with the highest number of grant installments.

The total number of grant installments per institution is presented in Figure 4.

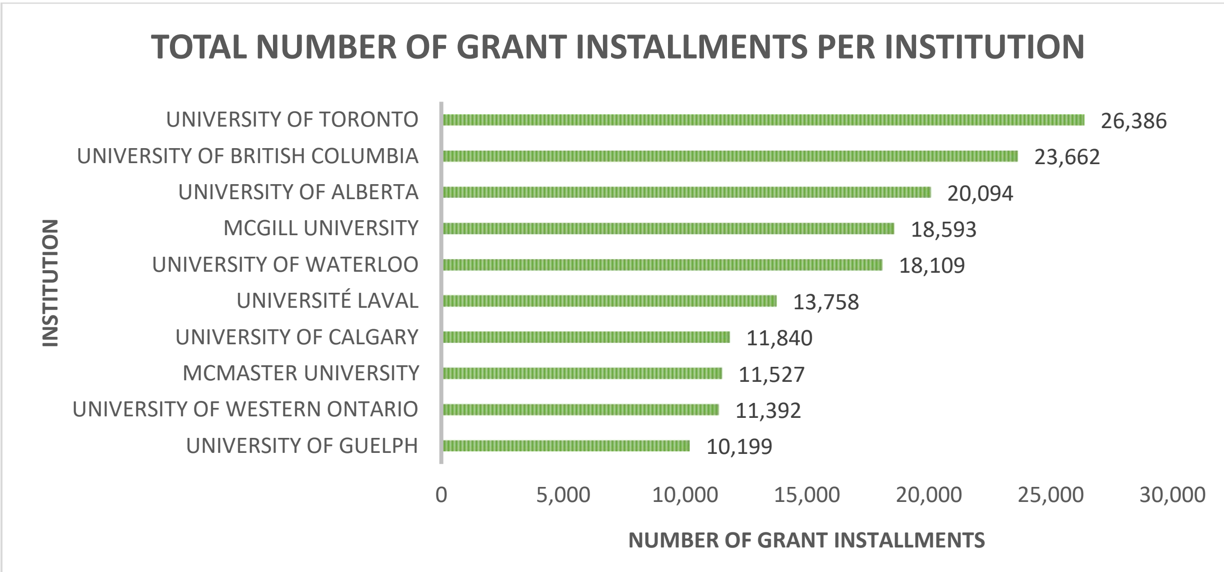


**Figure 4 – Total Number of Grant Installments Per Institution**

Figure 4 shows that researchers affiliated with universities in the provinces of Ontario, Quebec, British Columbia, and Alberta received the most grant installments. This corresponds to the results presented in Figure 2. Since the grant installments don't show the dollar amount of funding, we also looked at average annual funding amounts per researcher per institution.

The average annual funding amount per researcher per institution is presented in Figure 5.

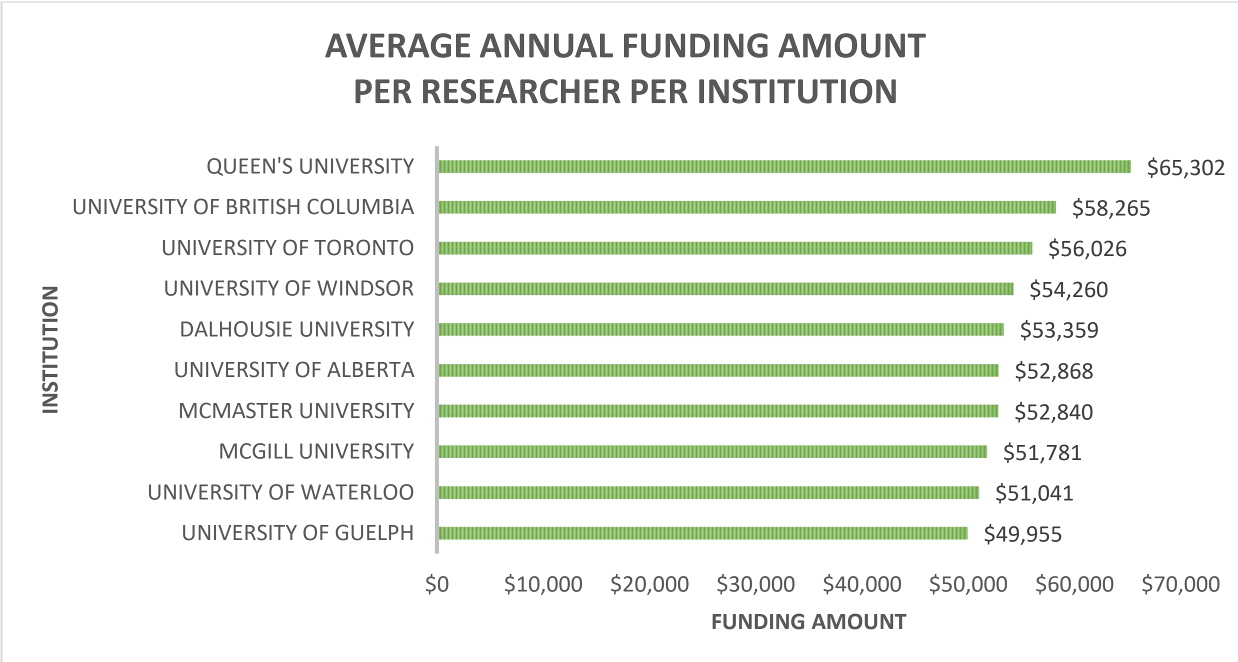


**Figure 5 – Average Annual Funding Amount Per Researcher Per Institution**

Figure 5 shows that researchers affiliated with Queen's University received the highest average annual funding amounts. This shows that although researchers from the University of Toronto received the highest number of grant installments (Figure 4), they did not receive the highest average dollar amount. This shows the importance of looking at both the number of grant installments and average dollar amounts to better observe the funding differences between institutions.

### 3-2-2. Descriptive Analysis of the Number of Publications and Citations

In this section, we focus on the productivity of researchers (measured by the number of publications), and scientific impact (commonly measured by number of citations in literature) (Lawani 1986; Moed 2006; Schiffauerova and Ebadi 2016).

Figure 6 shows the total number of publications by province. It is important to note that most publications in the dataset are collaborative. This poses a problem when counting number of publications and citations for provinces and institutions since there are papers with authors based in different provinces and affiliated with different institutions. To combat this issue, we counted the number of publications and citations for each different province and institution. i.e. A publication with authors at the University of Toronto in Ontario and authors from University of McGill in Quebec is counted for both provinces and institutions. The citations for that publication are also counted for both provinces and institutions.

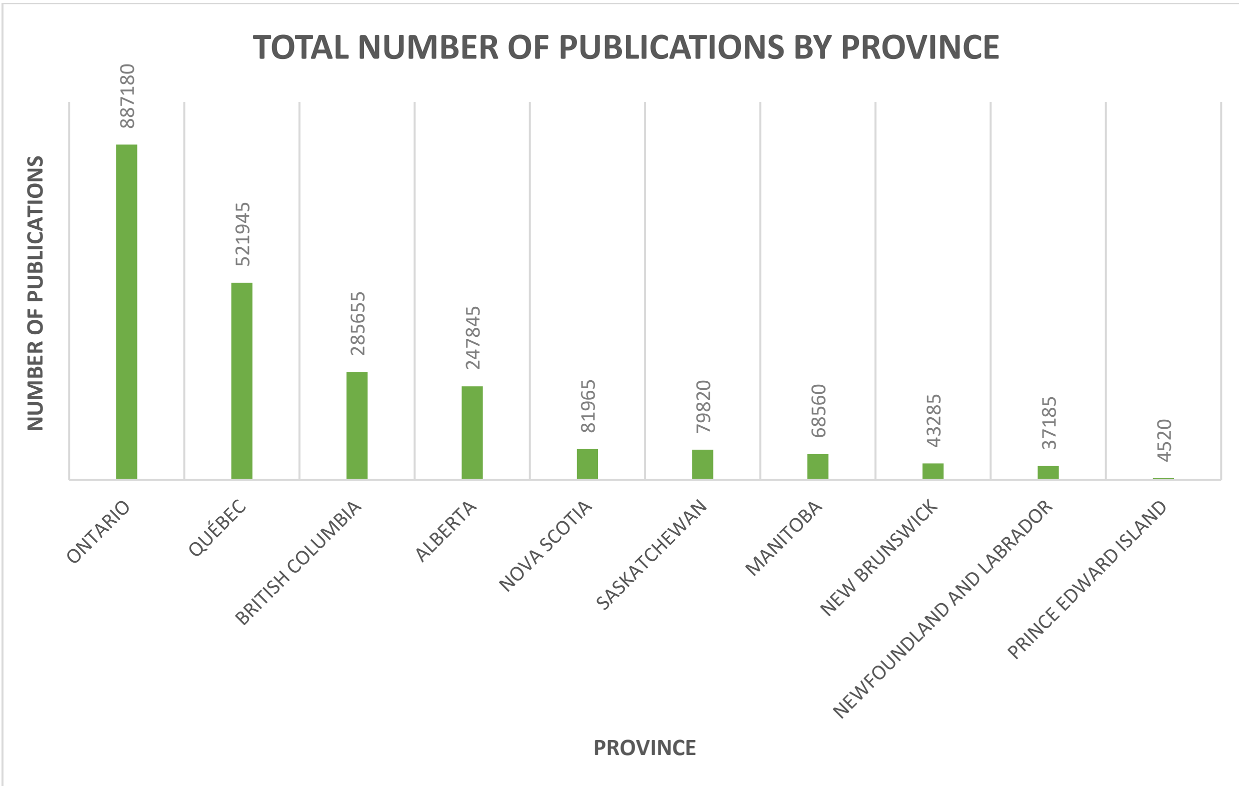


**Figure 6 – Total Number of Publications by Province**

Figure 6 shows that researchers based in Ontario, Quebec, British Columbia, and Alberta were the most productive. These researchers also received the highest funding amounts and the greatest number of grant installments. This shows that their productivity was positively affected by funding.

Next, we are going to compare total number of publications per institution to better assess performance.

Figure 7 shows the total number of publications per institution. It is worth mentioning that we are focusing on the top 10 institutions with the most publications.

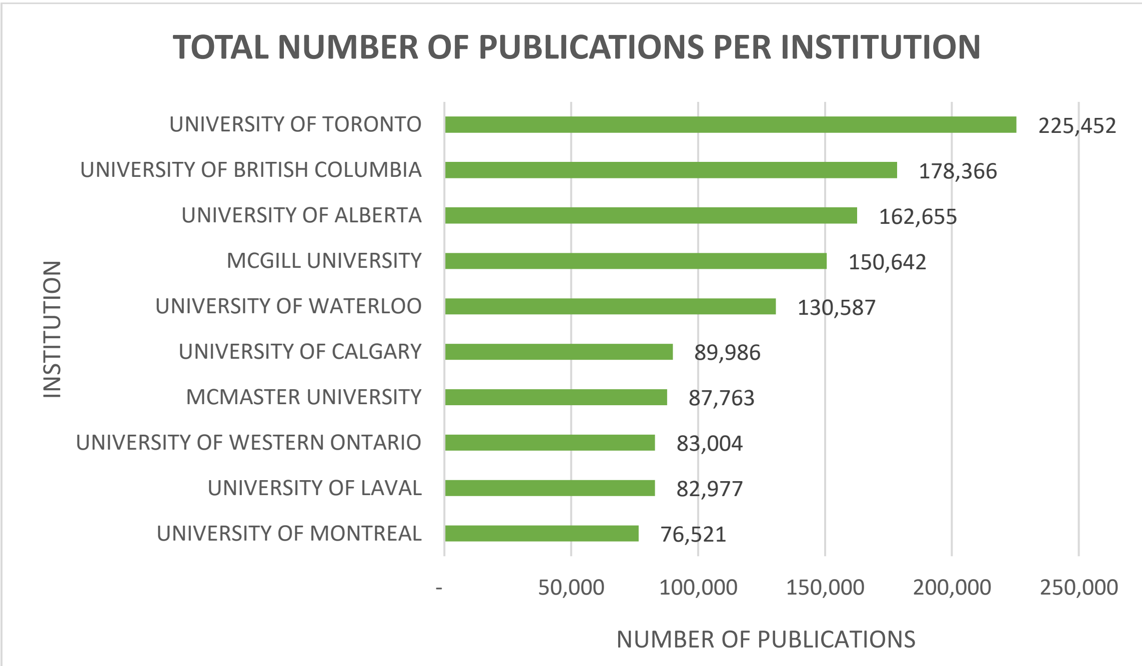


**Figure 7 – Total Number of Publications Per Institution**

There are similarities between the top institutions shown in Figure 7 and the top institutions for both grant installments (Figure 4) and funding amounts (Figure 5). This shows that funding positively affected the number of publications for these researchers. This aligns with the findings of Ebadi and Schiffauerova (2016).

In the next step we looked at the citation counts per province to assess scientific impact. Since the total number of citations directly correlates to the total number of publications, we instead examined the average number of citations per article to assess the performance of each province related to their scientific impact.

Average citations per article by province are shown in Figure 8.

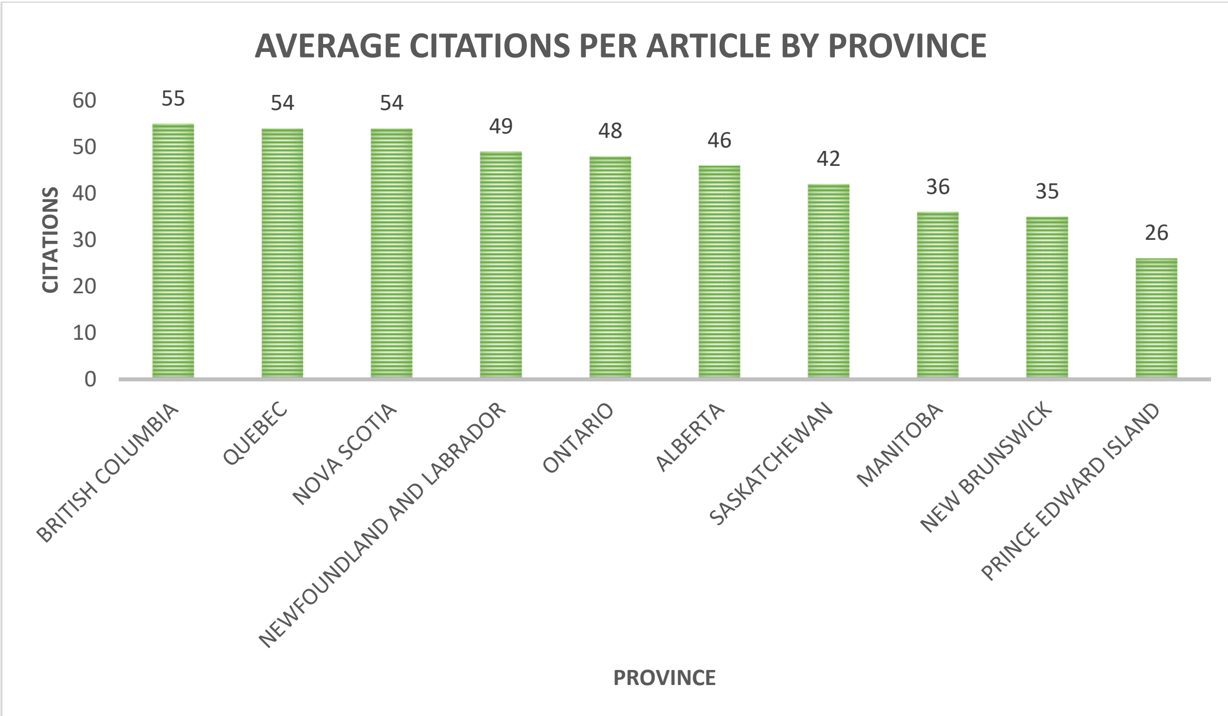


**Figure 8 - Average Citations Per Article by Province**

In Figure 8 we see that, throughout the entire dataset (1982-2018), publications that received the highest average amount of citations were from different provinces than the provinces with the highest number of publications (Figure 6). This is unexpected and requires further analysis. To accomplish this, we looked at the average citations per article by institution.

The information for the average citations per article for each institution is presented in Figure 9.

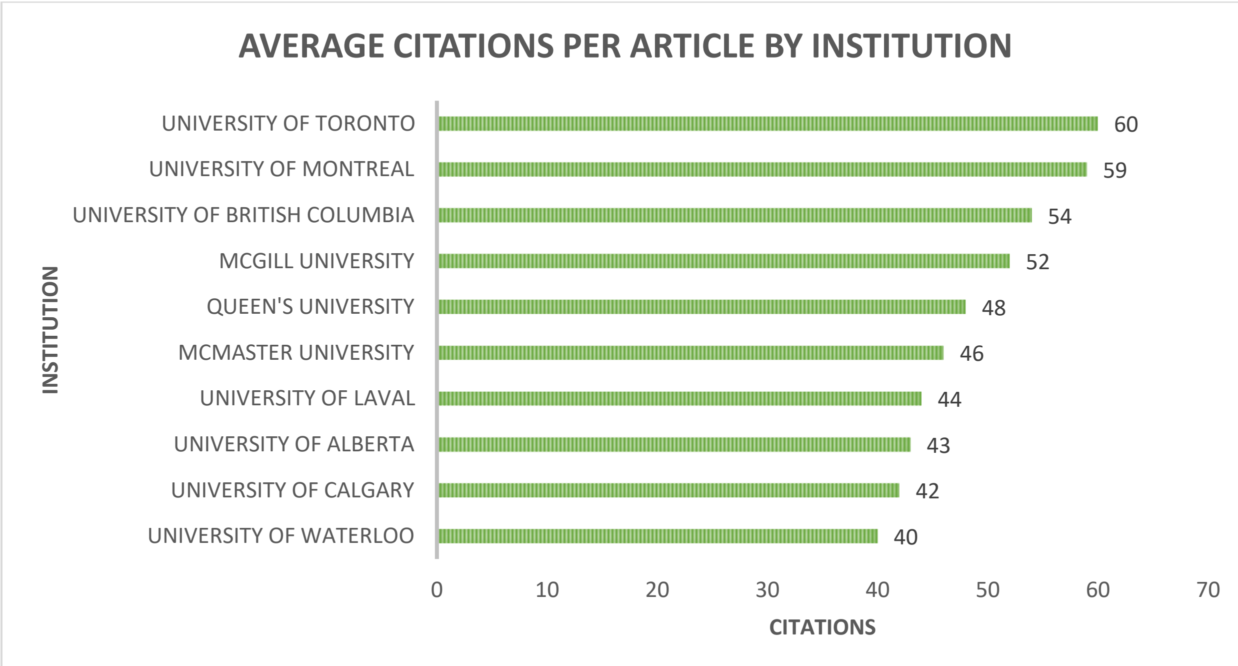


**Figure 9 - Average Citations Per Article by Institution**

In Figure 9, we observe that many of the most impactful institutions are not necessarily from the most impactful provinces. This could be because those provinces contain fewer universities compared to the rest of the country. For example, British Columbia is receiving the highest average citations despite having fewer universities than larger provinces. Thus, an institution like the University of British Columbia could be achieving higher citation numbers without as many lower-ranked universities bringing the average citation count down.

In contrast, some of the most impactful institutions are in Ontario – including University of Toronto, Queen's, McMaster, and Waterloo – despite Ontario falling lower on the previous chart. However, Ontario has many other universities that might not be receiving citations and leading to an overall lower average. These less impactful universities might be focused on undergraduate studies or programs that don't lead to citations.

It is worth mentioning that Concordia University has the 27$^{th}$ spot on the list with an overall average citation count of 29.

To conclude, the average citation numbers for a province can be affected by many factors including total number of universities and the programs offered at those universities.

### 3-2-3. Descriptive Analysis of Career Age

Another factor that we considered for this analysis is career age of researchers and its effect on their performance. In this thesis, the career age for each researcher is calculated by considering the time elapsed from their first publication present in the dataset. We assumed that this is their oldest publication and therefore subtracted the publishing year for that publication from the current year to get their career age.

It has been argued that researchers who have been productive in the past typically remain productive. Therefore, they continue to publish more articles and potentially gain access to higher amounts of funding as their career age grows (Merton, 1973; Kvik and Olsen, 2008). Nevertheless, it has been shown that although productivity increases as researchers' career age grows, scientific impact decreases (Ebadi and Schiffauerova, 2016). In the next steps, we are going to analyze the relationship between career age and distinct aspects of research.

First, we determined if there is a relationship between the researchers' career age and the amount of funding they received.

In Figure 10, we see the relationship between career age and the overall average amount of funding a researcher was granted in each specific year of their career. Since there was not much data available for the first 5 years, those have been removed from the figure.

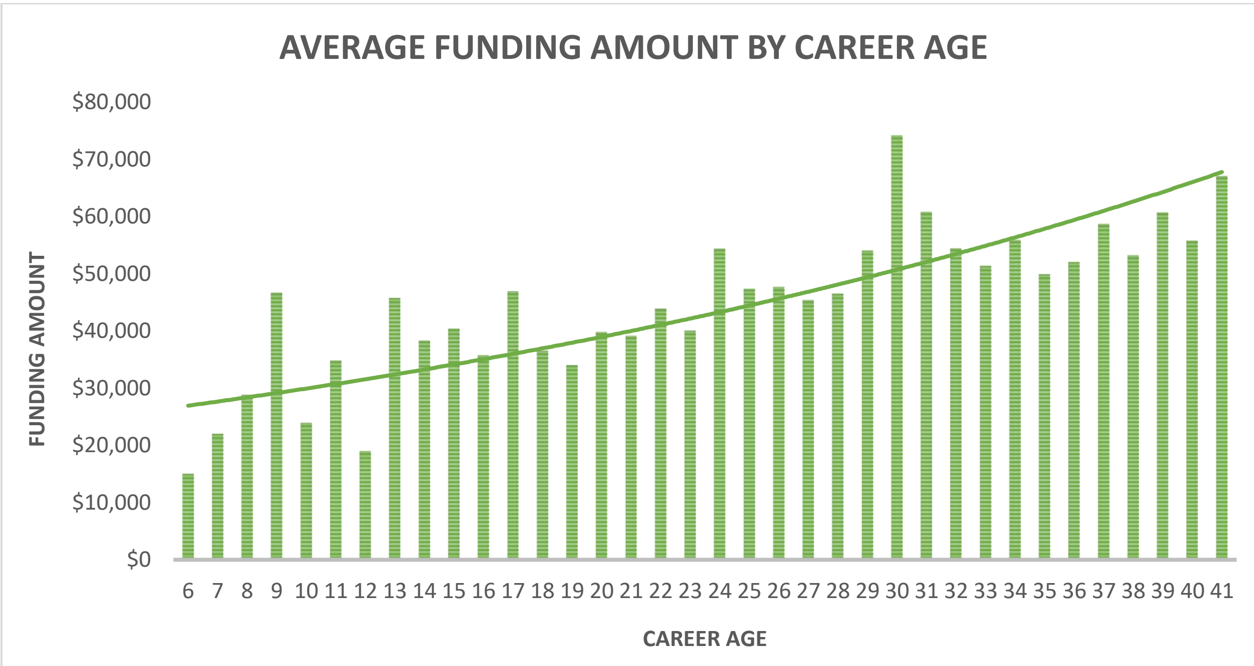


**Figure 10 - Average Funding Amount by Career Age**

In Figure 10, it is obvious that there is consistent growth in the average amount of funding granted to researchers as their career age goes up and this is clearly shown by the trend line in the chart. There are some spikes in the average funding amount for researchers with a certain career age which is not consistent with the overall trend. These are assumed to be outliers. Our findings are in line with the literature which states that researchers tend to have access to more funding as they gain more experience in their careers (Ebadi and Schiffauerova 2016).

After looking at average funding by career age, we looked at average number of publications by career age.

In Figure 11, we see the average number of publications by career age.

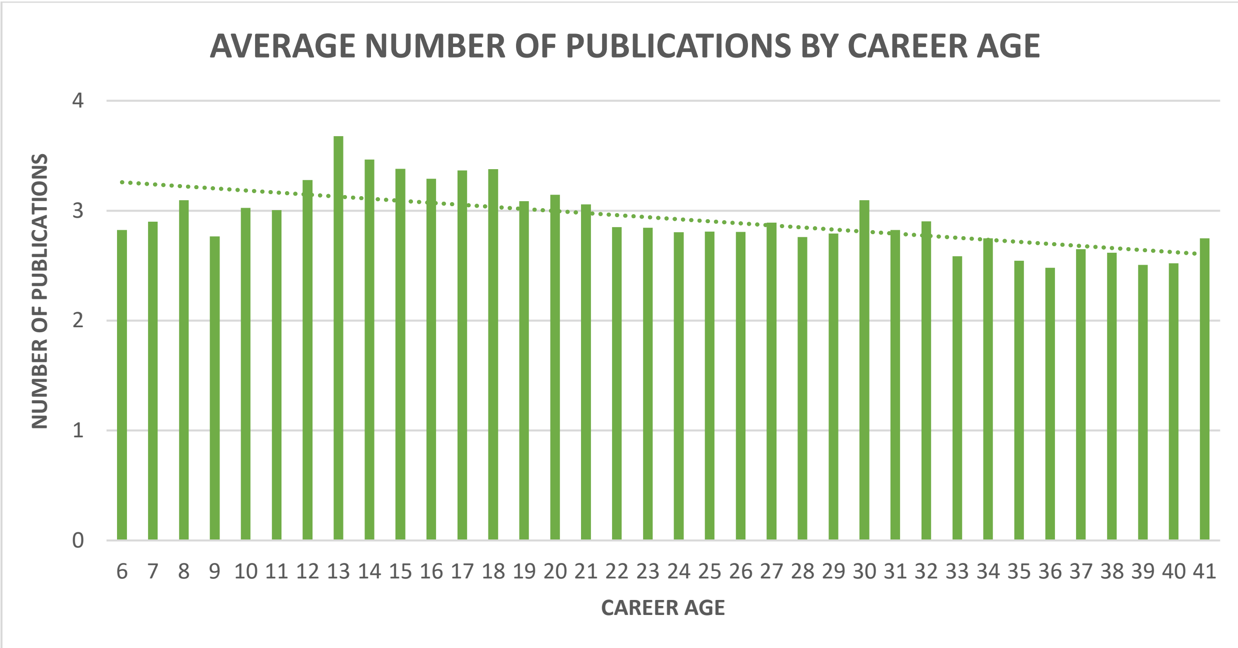


**Figure 11 - Average Number of Publications by Career Age**

In Figure 11 we see that after an initial rise in researchers' productivity, this is followed by a steady decline in productivity with a few small fluctuations. The overall trend is in line with past findings in the literature (e.g. Ebadi and Schiffauerova 2016) where they concluded that as researchers' career age increases, their productivity slowly decreases. This will be investigated further in the regression analysis section.

Next, we looked at average citations by career age.

Average citations by career age are shown in Figure 12. It is worth noting that due to the exceedingly small number of citations for career ages 1-5, we have removed them.

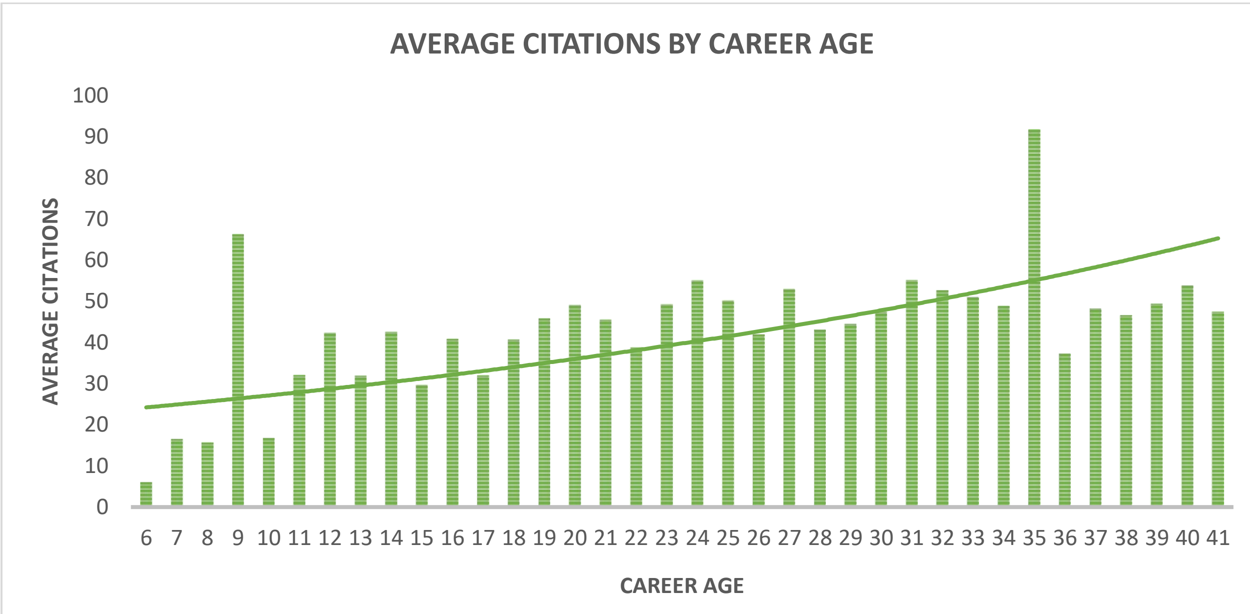


**Figure 12 - Average Citations by Career Age**

By looking at Figure 12, we see that after the first several years, the average citation of researchers remains relatively steady with a small amount of fluctuation. This is not in line with the findings of Ebadi and Schiffauerova (2016) which suggest that researchers tend to produce papers of lower scientific impact after a certain time in their careers. This will be investigated further in the regression analysis section.

### 3-2-4. Descriptive Analysis of Gender

In this section, the results of the descriptive analysis of gender are presented. The gender distribution info for the NSERC dataset is available in Table 5.

| Gender | Number of Researchers | Grant Installments |
|---|---|---|
| Male | 16,784 | 261,474 |
| Female | 3,861 | 54,348 |

**Table 5 - Gender Distribution of Researchers Funded by NSERC**

Table 5 shows that in the NSERC dataset we have a total of 16,784 male researchers and 3,861 female researchers which demonstrates the gender gap present in fields funded by NSERC. When it comes to the total number of grant installments, male researchers are receiving significantly more grant installments (261,474) compared to their female colleagues (54,348). This does not necessarily mean that female researchers are getting less funding amounts on average. This will be discussed in the next steps of the analysis.

In the next part of our gender analysis, we are going to compare the total number of grant installments awarded to researchers of each gender by NSERC on a year-by-year basis and discuss the trends. Since there were not many grant installments in the dataset before 1989, we chose to exclude those years from the following figures.

Figure 13 shows the total number of grant installments funded by NSERC annually by gender.

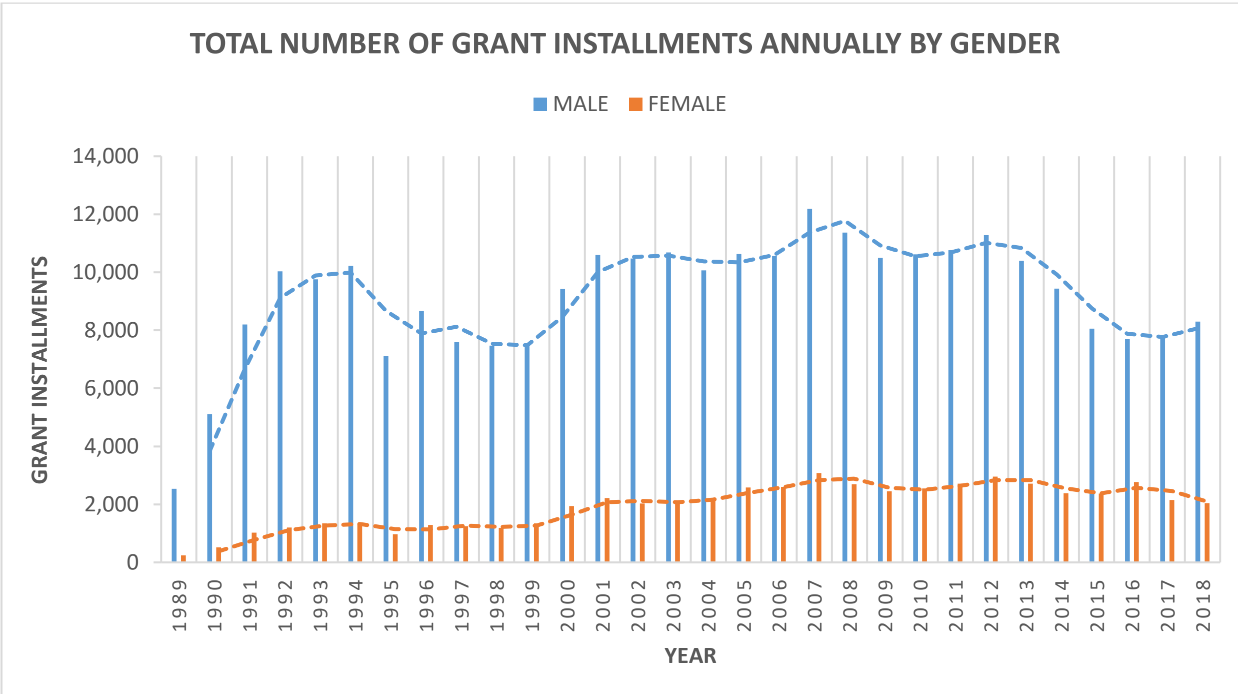


**Figure 13 - Total Number of Grant Installments Funded by NSERC Annually by Gender**

By observing Figure 13, it can be concluded that the trend for the number of grant installments received is similar for male and female researchers. Although both genders have been receiving a steady number of grant installments over the years, there is still a gender gap present. This shows that female researchers are underrepresented in fields funded by NSERC and aligns with findings in the literature (e.g. Holman et al., 2018, Ebadi and Shiffauerova 2016, Lariviere et al., 2011).

In the next phase of our gender analysis, we will compare the number of grant installments by gender in each province. It is worth mentioning that these figures don't reflect the dollar amount received by researchers in each province.

Figure 14 shows the total number of grant installments funded by NSERC by gender and province.

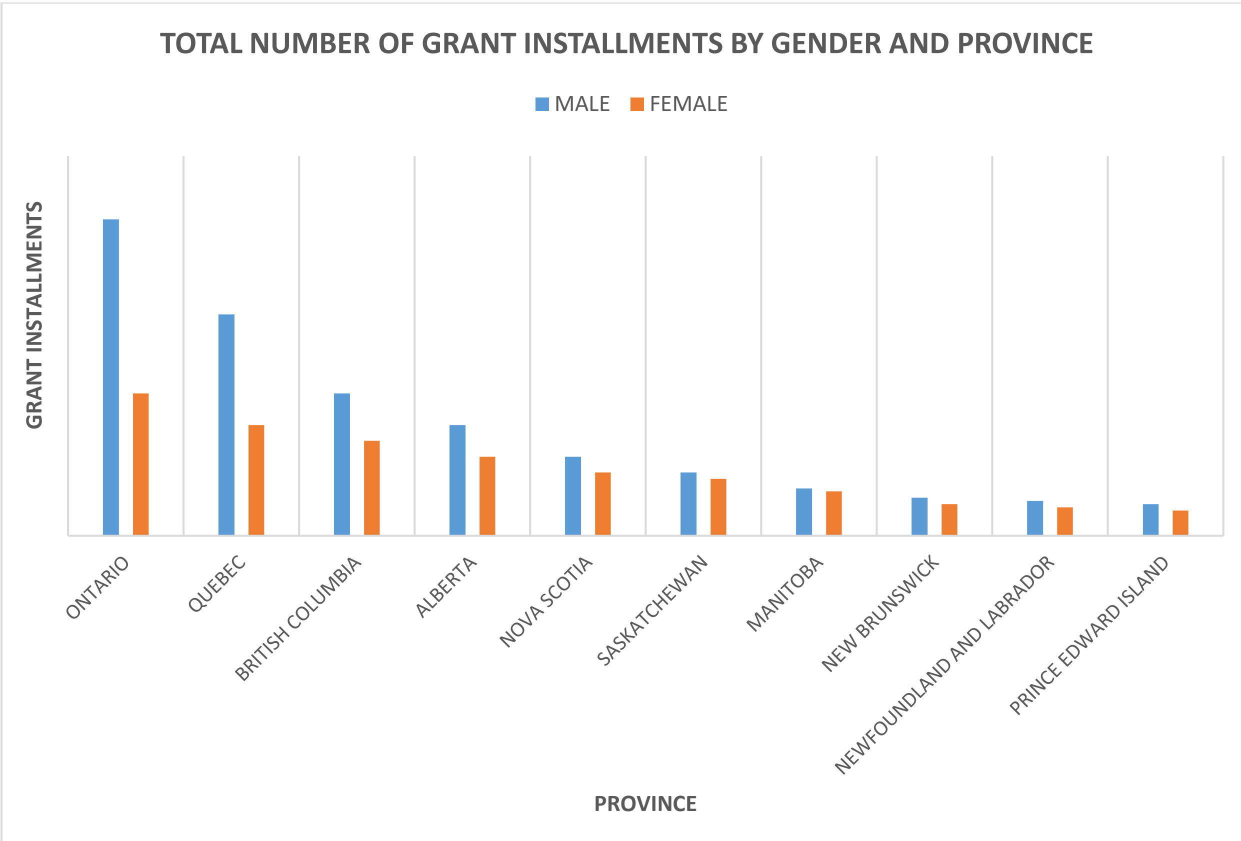


**Figure 14 - Total Number of Grant Installments Funded by NSERC by Gender and Province**

By observing Figure 14, we see that researchers of both genders located in the top five most populated provinces of Ontario, Quebec, British Columbia, Alberta, and Nova Scotia are receiving the highest number of grant installments. Although the gender gap is smaller in less populated provinces, it still exists. To conclude, province does not impact the gender gap as the issue is nationwide.

The next phase of our analysis includes the funding details for each gender.

The gender breakdown for the total funding amount, the maximum funding amount, the average funding amount (per individual), and the number of grant installments is presented in Figure 15.

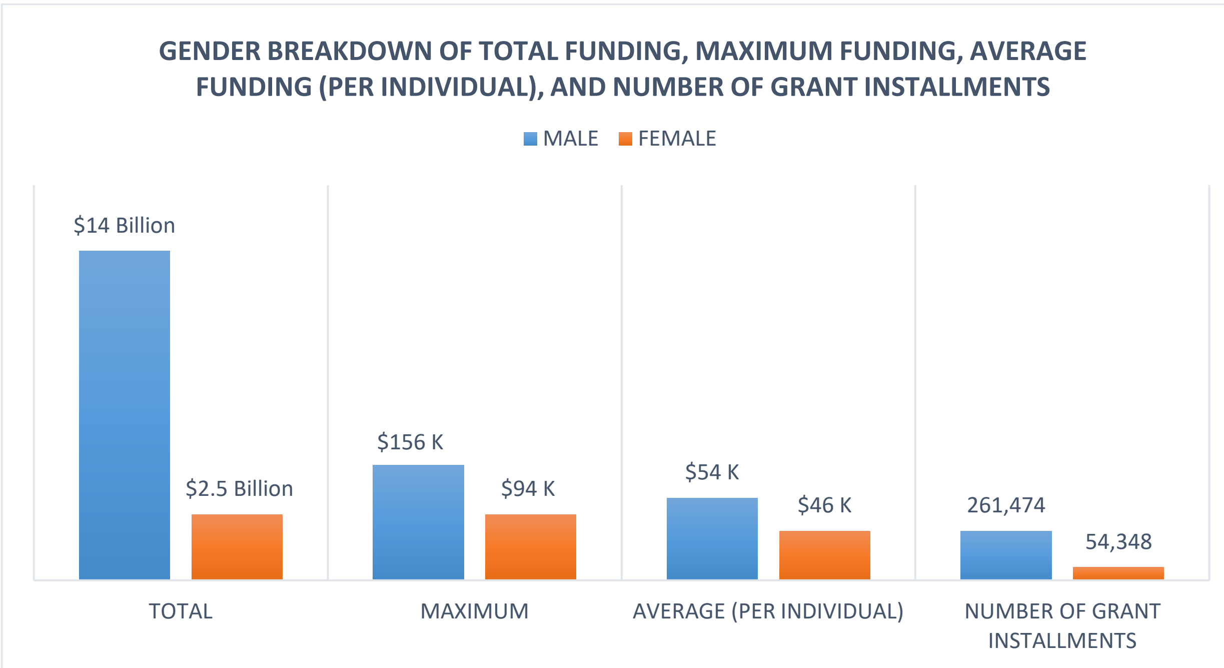


**Figure 15 – Gender Breakdown of Total Funding, Maximum Funding, Average Funding (Per Individual), and Number of Grant Installments**

In Figure 15, we see that the total amount of funds received by male researchers ($14 Billion) is approximately five times the amount received by female researchers ($2.5 Billion). It can also be observed that the number of grant installments received by male researchers (261,474) is about five times the amount received by female researchers (54,348). In addition, Figure 15 shows that the maximum funding granted to a male researcher ($156 K) is 60% higher than the maximum funding amount granted to a female researcher ($94 K). The observed gender gap is in line with the findings in past literature (e.g., Witteman et al., 2019).

Despite the large gender gap present in total funding, number of grant installments, and maximum funding, we see that male researchers received an average amount ($54 K) relatively close to the average amount received by female researchers ($46 K). This shows that while women are underrepresented in the fields funded by NSERC, they are receiving funding amounts close to their male colleagues.

For the next part of the gender analysis, we are going to analyze the differences between genders when it comes to receiving individual grant installments versus team grant installments from NSERC. It is worth noting that most grant installments in the NSERC dataset were individual grant installments received by one researcher who was the project lead with no co-researchers present.

In Figure 16, the number of individual grant installments versus team grant installments for each gender is presented.

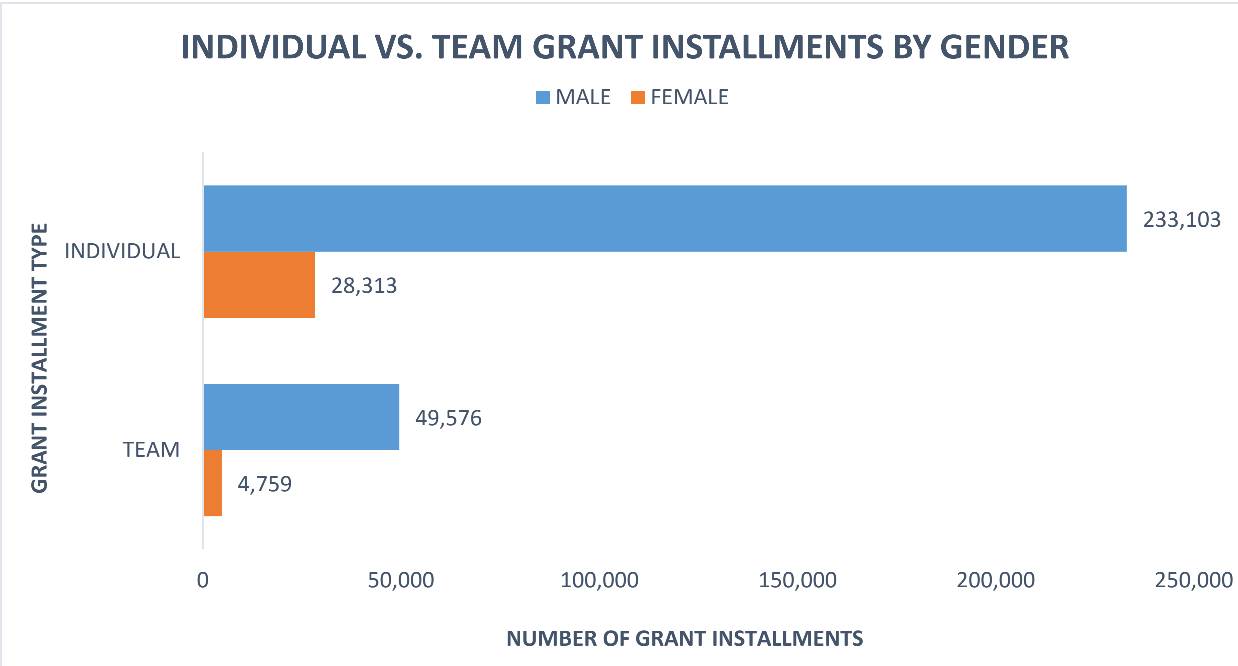


**Figure 16 - Individual Versus Team Grant Installments by Gender**

As observed in the previous phase, men are getting a higher number of grant installments overall. However, if we calculate the team grant installment percentage for both genders in Figure 16, we see that 10.83% of grant installments for men were team grant installments whereas 8.75% of grant installments for women were team grant installments. There is only a slight difference between genders. Therefore it can be concluded that, when it comes to receiving individual or team grant installments, both genders are performing similarly.

In the final step of our gender analysis, we look at each gender's performance by comparing the overall average amount of funding they receive, the average number of articles they publish per year, and the overall average number of citations they receive.

Figure 17 shows the overall average funding, average number of articles per year, and average number of citations per year by gender.

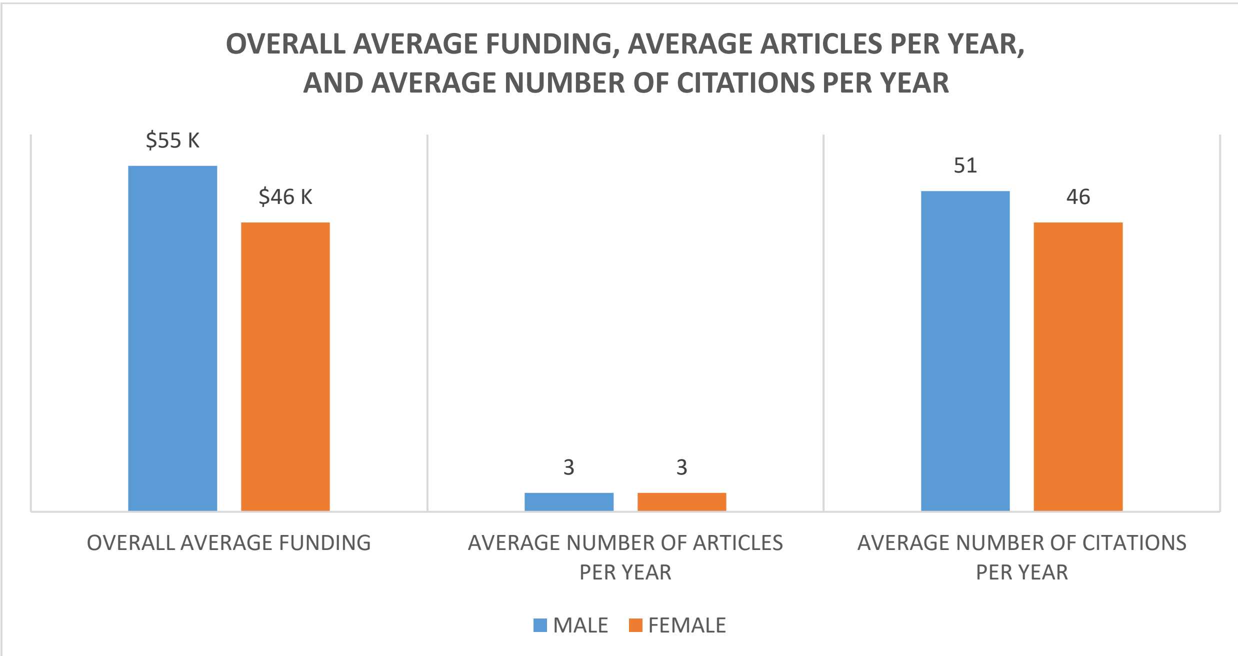


**Figure 17 - Overall Average Funding, Average Articles Per Year, and Average Number of Citations Per Year**

In Figure 17, we can observe that amongst Canadian researchers funded by NSERC, males are attaining slightly higher amounts of funding compared to females. This is in line with the findings in the literature (e.g. Lariviere et al., 2011). However, females are receiving a similar number of citations and are producing an equal number of articles – demonstrating that female researchers are performing similarly to their male colleagues. This contrasts with past findings of Lariviere et al. (2011). Further investigation using regression analysis is necessary and will be covered later.

### 3-2-5. Descriptive Analysis Conclusion

To conclude our descriptive analysis, we are going to assess the overall scientific performance of each gender. "Average cost of an article" refers to the average amount of annual funding for each gender divided by the average number of articles published per year for each gender. "Average cost of a citation" refers to the average amount of funding per year for each gender divided by the average number of citations per article for each gender.

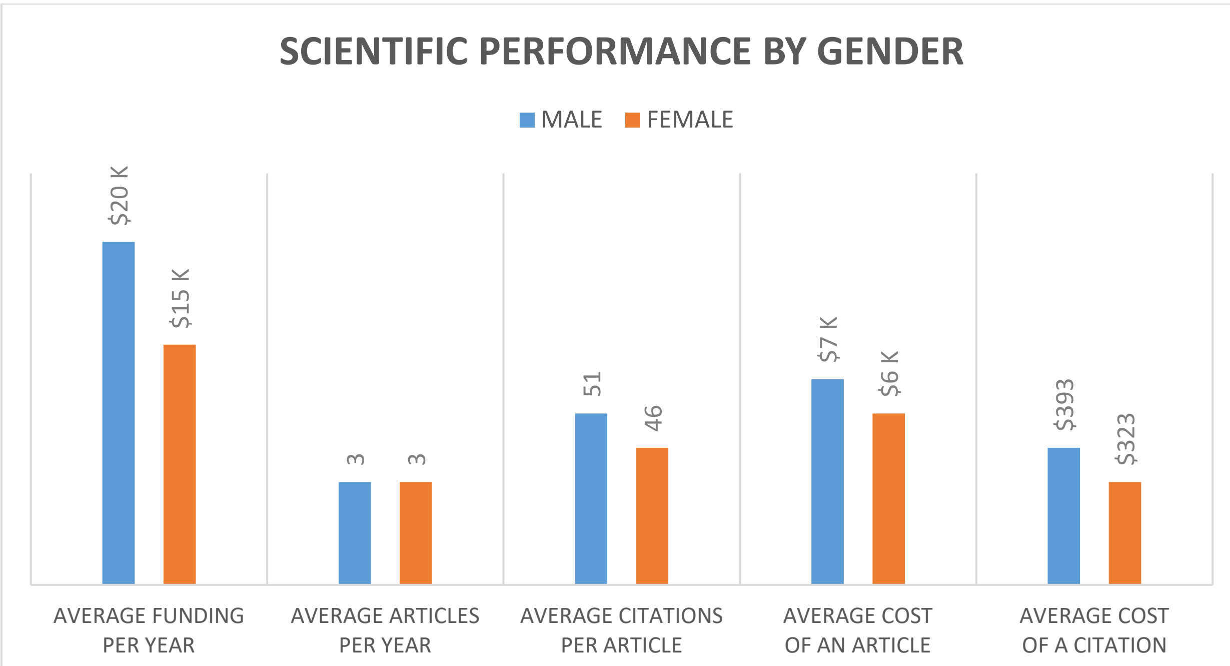


**Figure 18 - Scientific Performance by Gender**

Figure 18 shows that, on average, male researchers are receiving more funding than females. Their articles are also obtaining more citations than females. However, males and females are producing an equal average number of articles per year. However, when the average cost of articles and citations is considered, it demonstrates that despite male researchers' ability to secure higher amounts of funding, females are just as productive. We conclude that female researchers can generate similar or comparable research output with lower research costs than male researchers and are thus more efficient in their research.

# 4. Regression Analysis

## 4-1. Regression Methodology

In the final part of the analysis, we used regression analysis to see the impact of several factors on researcher productivity, where we are considering both publication quantity (number of publications) and publication impact (number of citations). It is worth mentioning that since the number of publications in the Scopus dataset is low in the early years of the analysis and only starts to increase by the late 90s, we have decided to only focus on the publications for the period of 1999-2018 for the scope of our regression analysis. This is also a relatively more recent period compared to past work in the literature. We are going to calculate the variable values on a year-by-year basis for the regression analysis, combine all the annual data to make one complementary dataset for our models, and then interpret the results. Two models have been proposed for this purpose which will be explained later in the analysis. For running the regressions, we use SPSS software by IBM.

Since the first model (quantity) deals with count measure data (publications each year), we first needed to check whether the Poisson or the Negative Binomial regression was better suited for our data. For the Poisson model to be a good fit, the variance must be roughly the same as the mean. If the variance is significantly higher than the mean, the Negative Binomial model would be better suited. For our dependent variable the mean was 2.85 and the variance was 8.06. Based on the results of the ANOVA test – proving that there is a significant variance between the values - and the deviance value we decided that the Poisson model would not be a good fit for our data since the result of this test was statistically significant, and we had a considerable amount of overdispersion. This was expected since when dealing with real data it is quite rare to have a Poisson distribution (Ebadi and Schiffauerova 2016, Coleman and Lazarsfeld 1981). Therefore, we have decided to use the Negative Binomial model for our analysis since it was better suited for our data. For the second model (impact = average citations per researcher per year) we decided to use multiple linear regression. Since linear regression is used in previous studies for this purpose with similar data (e.g., Beaudry et al 2014, Ebadi and Schiffauerova 2016) it has also been selected in this analysis. This approach has shortcomings such as not accounting for non-linear interactions between variables. To combat this, we also introduced an interaction effect between gender and funding variables which will be explained in the following step to account for non-linear interactions. This partially accounts for non-linearities in the values. The results of the analysis will be shown in the next step. It is worth noting that the data has been checked for normality as well. For the impact model, we normalized the data by transforming the values into logarithmic values using a Log10 function in SPSS. We chose this method of transformation because it is a suitable method for our data that is positively skewed. For the regression models it is worth noting that we will be using a confidence interval of 95%, therefore it means **alpha=0.05** for this analysis. Lastly, we looked at the goodness of fit for our models to assess their performance for our data. Based on the R-squared value (.41) the models are a good fit for the data although this value could be higher using different algorithms (e.g., machine learning).

### 4-1-1 Quantity of the Publications Model

For the quantity model, we looked at various factors affecting the number of articles produced by a researcher each year. This measure has been widely used in the industry as a proxy of scientific productivity (e.g., Centra 1983; Okubo 1997; Schiffauerova and Ebadi 2016). The following regression model has been proposed for the scope of this analysis:

**(1)** Article_number = f (average_fund3 + average_citation3 + average_article3 + career_age + average_teamsize + gender + genderXfund)

In the model, our dependent variable is the number of publications each year produced by a researcher called **Article_number**. Our first independent variable, **average_fund3,** is the average amount of funding the researcher received in the past three years. In literature, three years (e.g., Payne and Siow 2003) or five years (e.g., Jacob and Lefgren 2007) have been considered as the time window during which the funding received might have an impact on the production of publications in the subsequent year. Both time windows were considered and, after reviewing the results, we did not notice any major difference between the correlations or significance of the variable, so we decided to utilize the three-year window for our analysis to keep the results more relevant since it is also a common timeframe according to the literature (e.g., EbadI and Schiffauerova, 2016). The next independent variable in the model, **average_citation3,** is the average number of citations received by publications of the researcher in the past three years. This variable is used as a proxy to measure the quality/impact of publications and helps us in the model to determine if there is a connection between past citation numbers and the researcher's future productivity. A similar method has been used in literature. (e.g., Ebadi and Schiffauerova, 2016). The next variable is **average_article3,** which is the average number of publications by the researchers in the period of [***i***-1, ***i***-3] with ***I*** being the given year. This is to see if publications from the previous three years impact future productivity of researchers. According to the literature, past productivity has a positive impact on future productivity (e.g. Ebadi and Schiffauerova 2016), therefore we decided to investigate this for our analysis as well. We included another variable, **career_age,** in the model which, as discussed previously, is the time difference between the researcher's first publication and the current year. In literature (e.g., Beaudry and Schiffauerova, 2014) it has been shown that career age has an impact on the researcher's productivity, which is why we decided to investigate this further in our analysis. The next variable, **average_teamsize,** is the average number of co-authors co-publishing articles with a researcher each year. In the previous analysis, we found that having a collaborative research project leads to a higher number of published papers. We concluded this mostly based on our descriptive analysis results. Therefore, we wanted to see if having a bigger team has any impact on productivity. We included one dummy variable which is **gender**. This is to see the impact of each gender and to investigate the role of gender in research productivity which is one of the main objectives of this analysis. We also have an interaction variable, **genderXfund,** which shows the interaction effect between **average_fund3** and **gender** variables. Since we have seen in our descriptive analysis results that, on average, female researchers receive less funding, we decided to include this variable in the model to see the role of gender in securing funds and affecting productivity, which is yet another main objective of this analysis.

### 4-1-2 Scientific Impact of the Publications Model

To satisfy the second objective, we developed a model to estimate the effect of various variables on the quality (or impact) of the publications, measured by the average number of citations for all the publications of a researcher each year. It has been argued in the literature that this method comes with certain disadvantages (e.g., negative citations, self-citations, etc.), but it is still considered a standard measure to analyze the impact of publications (e.g., Lawani 1986; Moed 2006; Schiffauerova and Ebadi 2016), and since it provides critical information on the impact of the publications, it is widely used (Adler et al., 2009).

The following regression model has been proposed for this analysis:

**(2)** Average_citation = f (average_fund3 + average_citation3 + average_article3 + career_age + average_teamsize + gender + genderXfund)

In the model above, our dependent variable is going to be the average number of citations for all of a researcher's publications each year. The definition of all the independent and dummy variables is the same as in the first model.

#### 4-2. Regression Analysis Results

In the following sections, we present and interpret the results of our regression analysis. We look at the correlations table to check for highly correlated variables, and the coefficients table to see which variables are statistically significant in describing the changes in the dependent variable and see how much of an impact each variable has in the models.

#### 4-2-1. Quantity of the Publications Model

The result of the *ANOVA* test is presented in Table 6.

| Article_number | df | F | Sig. |
|---|---|---|---|
| Between Groups | 1 | 74.563 | <.001 |

**Table 6 - ANOVA Test Results**

According to Table 6, the test results are statistically significant, therefore we are going to use the Negative Binomial Regression model.

The correlations table for the first model is presented in Table 7. Pearson correlation values are used for this analysis.

| | (Intercept) | [gender=0] | [gender=1] | Career_Age | Average_citation3 | Average_fund3 | Average_article3 | Average_teamsize | genderXfund |
|---|---|---|---|---|---|---|---|---|---|
| (Intercept) | 1.000 | -.650 | .[a] | -.604 | -.082 | .026 | -.238 | -.101 | -.583 |
| [gender=0] | -.650 | 1.000 | .[a] | -.097 | .000 | -.131 | -.030 | .002 | .760 |
| [gender=1] | .[a] | .[a] | .[a] | .[a] | .[a] | .[a] | .[a] | .[a] | .[a] |
| Career_Age | -.604 | -.097 | .[a] | 1.000 | -.001 | -.037 | .024 | .026 | -.007 |
| Average_citation3 | -.082 | .000 | .[a] | -.001 | 1.000 | -.012 | .011 | -.038 | -.005 |
| Average_fund3 | .026 | -.131 | .[a] | -.037 | -.012 | 1.000 | -.012 | -.004 | -.165 |
| Average_article3 | .238 | -.030 | .[a] | .024 | .011 | -.012 | 1.000 | -.025 | -.002 |
| Average_teamsize | .101 | .002 | .[a] | .026 | -.038 | -.004 | -.025 | 1.000 | -.002 |
| genderXfund | -.583 | .760 | .[a] | -.007 | -.005 | -.165 | -.002 | -.002 | 1.000 |

**Table 7 - Correlations – Quantity Model**

The Pearson values in Table 7 show that no two variables in the model are highly correlated. If two variables were highly correlated, we would need to remove one of them to optimize the model. The number of articles in a year has a negative correlation with gender, which is to be expected, as female researchers tend to be less productive (Lariviere et al., 2013). The number of articles in a year has a negative correlation with career age. This finding is not in line with similar studies (e.g., Lee and Bozeman, 2005). The reason could be that the data for our analysis is more recent, and it focuses on a more diverse selection of disciplines. This can be studied further in future studies. There is a negative correlation between the number of articles per year and the average number of citations the researcher received in the past three years. This is not clear and needs to be investigated further. There is a positive correlation between the number of articles per year and the average amount of funding the researcher received in the past three years, the number of articles they published in the past three years, and the average size of their team. In other words, funding, past productivity, and bigger scientific teams boost productivity. There is a negative correlation between the number of articles published in a year and the interaction of funding and gender.

Table 8 shows each variable, its significance, and its impact on the dependent variable.

| Parameter | B | Std. Error | 95% Wald Confidence Interval | | df | Sig. |
|---|---|---|---|---|---|---|
| | | | Lower | Upper | | |
| (Intercept) | .820 | .0111 | .799 | .842 | 1 | <.001 |
| [gender=0] | .063 | .0088 | .045 | .080 | 1 | <.001 |
| [gender=1] | 0[a] | . | . | . | . | . |
| Career_Age | -.008 | .0003 | -.009 | -.008 | 1 | <.001 |
| Average_citation3 | -3.206E-5 | 2.0095E-5 | -7.144E-5 | 7.329E-6 | 1 | .001 |
| Average_fund3 | 7.099E-8 | 2.4908E-8 | 2.217E-8 | 1.198E-7 | 1 | .004 |
| Average_article3 | .110 | .0007 | .108 | .111 | 1 | <.001 |
| Average_teamsize | .000 | .0003 | -.001 | .000 | 1 | .008 |
| genderXfund | 2.500E-7 | 1.5007E-7 | -4.409E-8 | 5.442E-7 | 1 | .006 |

**Table 8 - Parameter Estimates – Quantity Model**

Table 8 shows that all the independent variables are statistically significant since all the values are lower than 0.05. To dive deeper, we will be looking at the values of B. For gender, we have index values of 0 for males and 1 for females. Gender has a slight impact, suggesting that male researchers are more productive. This is in line with past literature (e.g., Lariviere et al., 2013). Our findings suggest that although female researchers have access to lower funding amounts on average, they are more productive considering the number of publications they produce per year is quite close to their male colleagues. This suggests that female researchers can publish almost a comparable number of articles at a considerably lower cost (with much lower funding). Career age has a slight negative impact which means as researchers get older their productivity goes down. To elaborate, based on our findings, productivity of researchers increases with experience, but after reaching a certain number of years in academia the productivity of researchers starts slowly decreasing. This is in line with past findings in the literature (e.g., Ebadi and Schiffauerova, 2016). Average citations in the past three years have had a negative impact on productivity. In other words, it seems that researchers who get cited more often tend to publish fewer articles in the following years. This was not expected and could be investigated further in future studies. Funding in the past three years has had a significant positive impact on productivity meaning more funding leads to more papers. This is also in line with past findings (e.g., Godin, 2003). The average number of articles in the past three years had a positive impact, meaning that productive researchers will continue to be more productive in the future. This was expected since it aligns with past findings (e.g., Ebadi and Schiffauerova, 2016). Average team size has no impact on the average number of articles published per year. This was not expected and is different from some of the other studies we have reviewed (e.g., Ebadi and Schiffauerova, 2015, Plume and Van Wiejen, 2014). The reason behind this could be the difference in data and the variables used in other studies models. The interaction between gender and funding has a positive impact on the number of articles per year which implies that female researchers who get more funding turn out to be more productive compared to their male colleagues. In other words, female researchers' funding to productivity ratio is higher than that of male researchers. This means that if female researchers were to receive more funding than male researchers, they could publish even more papers compared to male researchers.

### 4-2-2. Scientific Impact of the Publications Model

For the impact model, we checked the values of our dependent variable, **Average_citations,** for normality. To do this we used a normality test, looked at the skewness, and the results proved the values were not normally distributed. The normal probability plot is presented in Figure 19.

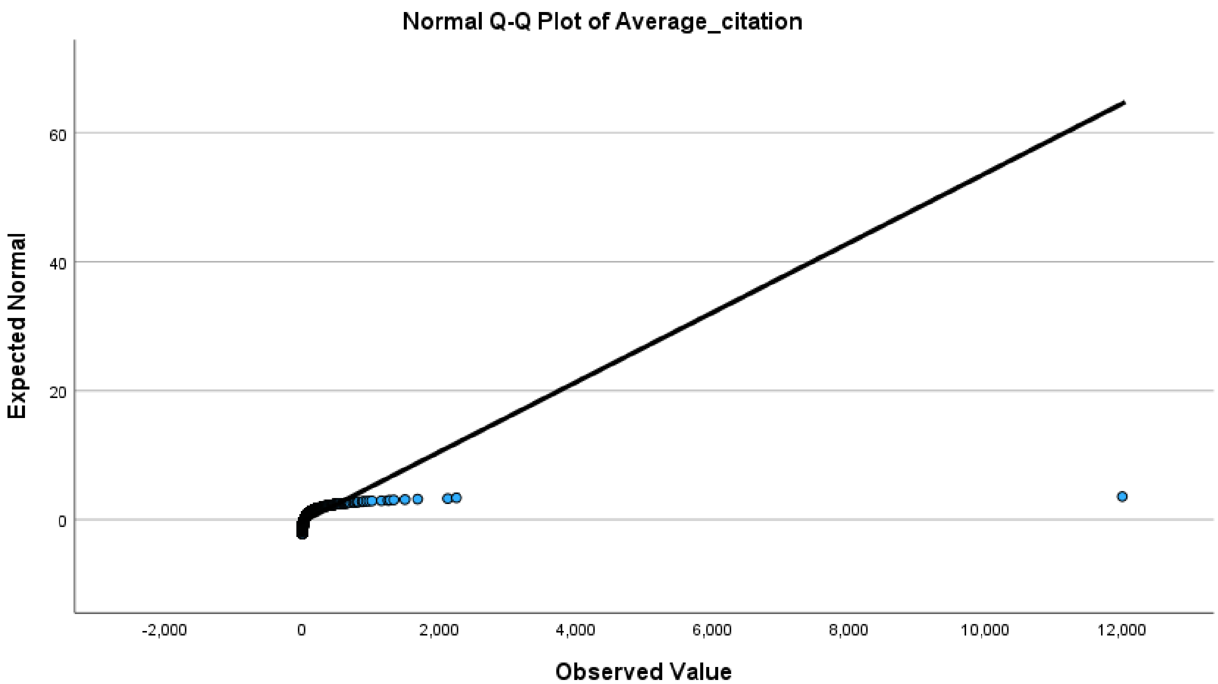


**Figure 19 - Normal Probability Plot – Original**

Figure 19 shows that the values start to derail from the normality line quite significantly, therefore proving that the values of our dependent variable are not normally distributed. It also shows that the distribution is positively skewed. To fix this, we have transformed the data using a log10 function and checked for normality again to see if the values are normally distributed after the transformation.

We conducted the tests once again and this time the result showed the data is now normally distributed.

The normal probability plot after the transformation is presented in Figure 20.

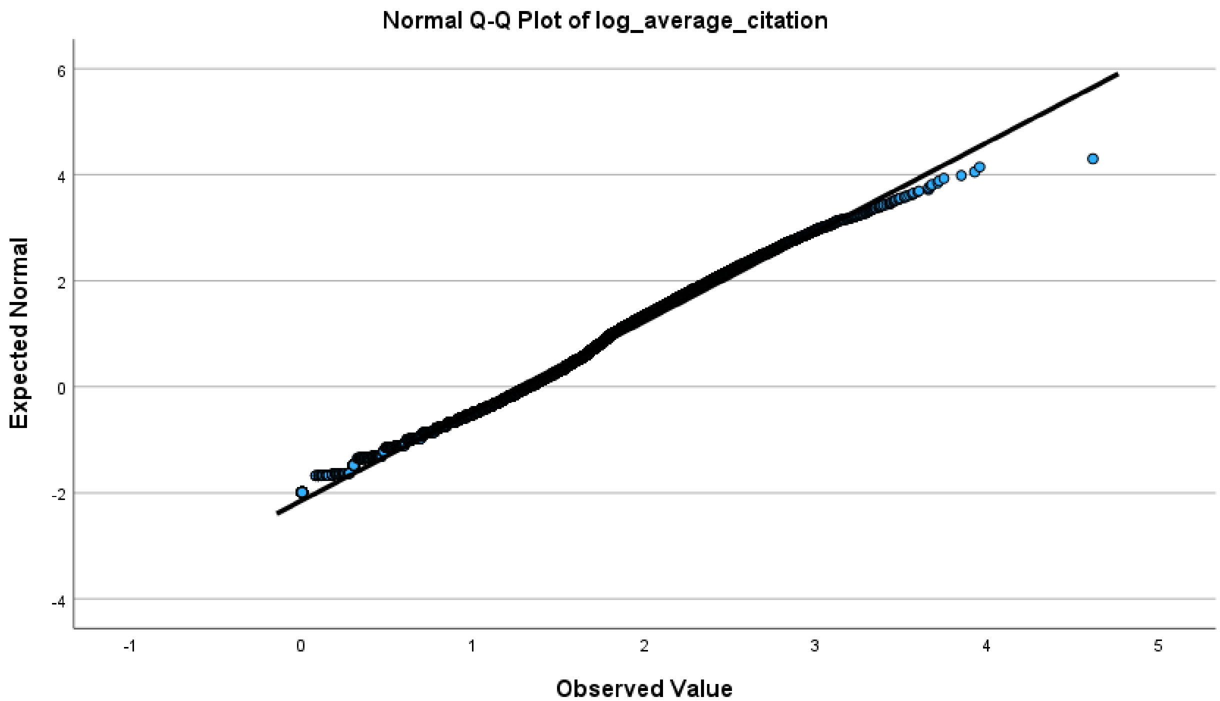


**Figure 20 - Normal Probability Plot – Transformed**

Figure 20 shows that the values follow the normality line closely, meaning that the data is now normally distributed. After making sure the data was normal, we continued with our regression analysis for the impact model. The results of the analysis are shown in the next steps.

The correlation values for the second model are presented in Table 9.

| | log_average_citation | Career_Age | gender | Average_citation3 | Average_fund3 | Average_article3 | Average_teamsize | genderXfund |
|---|---|---|---|---|---|---|---|---|
| log_average_citation | 1.000 | -.055 | .035 | .248 | .011 | .036 | .021 | .032 |
| Career_Age | -.055 | 1.000 | -.148 | -.002 | .037 | -.031 | -.025 | -.104 |
| gender | .035 | -.148 | 1.000 | .005 | -.014 | -.034 | .008 | .752 |
| Average_citation3 | .248 | -.002 | .005 | 1.000 | .012 | -.006 | .037 | .008 |
| Average_fund3 | .011 | .037 | -.014 | .012 | 1.000 | .007 | .004 | .103 |
| Average_article3 | .036 | -.031 | -.034 | -.006 | .007 | 1.000 | .024 | -.023 |
| Average_teamsize | .021 | -.025 | .008 | .037 | .004 | .024 | 1.000 | .008 |
| genderXfund | .032 | -.104 | .752 | .008 | .103 | -.023 | .008 | 1.000 |

**Table 9 - Correlations – Scientific Impact Model**

Table 9 shows that there is a negative correlation between the average number of citations a researcher receives and their career age. There is a positive correlation between the average number of citations a researcher gets and gender, the average number of citations they received in the past three years, the average amount of funding they received in the past three years, the average number of articles they published in the past three years, their scientific team size, and the interaction effect of funding and gender. There are no two highly correlated variables.

Next, we are going to look at the coefficients table to see the significance of each variable and its impact on our dependent variable in the second model. The details are shown in Table 10.

| | Unstandardized Coefficients | | Standardized Coefficients | | 95.0% Confidence Interval for B | |
|---|---|---|---|---|---|---|
| | B | Std. Error | Beta | Sig. | Lower Bound | Upper Bound |
| (Constant) | 1.274 | .007 | | <.001 | 1.261 | 1.287 |
| Career_Age | -.003 | .000 | -.050 | <.001 | -.004 | -.003 |
| gender | .031 | .007 | .020 | <.001 | .018 | .044 |
| Average_citation3 | .001 | .000 | .247 | <.001 | .001 | .001 |
| Average_fund3 | 6.035E-8 | .000 | .009 | .002 | .000 | .000 |
| Average_article3 | .008 | .001 | .036 | <.001 | .007 | .009 |
| Average_teamsize | .001 | .000 | .009 | .001 | .000 | .001 |
| genderXfund | 2.394E-7 | .000 | .009 | .032 | .000 | .000 |

**Table 10 - Coefficients Table – Scientific Impact Model**

By looking at Table 10, we see that all variables are statistically significant. Career age has a slight negative impact which means that as career age goes up the number of citations tends to decrease. This is similar to what we observed in the previous model where we noted the decline in productivity associated with higher career age. This is also in line with past findings in the literature (e.g., Ebadi and Schiffauerova, 2016). Gender has a slightly positive impact which translates to more citations, on average, for female researchers compared to male researchers. This is, however, different from past findings in the literature (e.g., Lariviere et al., 2013). The reason could be that in this analysis the focus is on Canadian researchers and the data is more recent. This could be further investigated in future studies.

The average citations in the past three years have had a slightly positive impact which means that researchers who get cited more in the past continue to get cited even more in the future. This is to be expected and aligns with the studies we reviewed (e.g., Ebadi and Schiffauerova 2016). The average funding in the past three years has had a significant positive impact on citations. This shows that researchers who receive higher amounts of funding tend to produce higher impact papers overall. This is interesting since there seems to be no relationship present between funding and the impact of papers in several previous studies (e.g., Godin 2003; Payne and Siow 2003; Tahmooresnejad et al., 2015).

The average number of articles in the past three years has had a small positive impact, meaning that past productivity leads to more citations in the future. In other words, researchers who were productive in the past tend to produce higher-impact papers in the following years. This aligns with the findings of Ebadi and Schiffauerova (2016). The average team size has a slightly positive impact, meaning that bigger teams get cited more often. This means that researchers who form bigger scientific teams tend to get access to better resources and therefore tend to produce higher-impact papers (Katz and Martin 1997; Melin 2000; Beaver 2001; Heinze and Kuhlmann 2008). The interaction between gender and funding has a positive effect on the impact of papers produced by researchers. This means that female researchers who get more funding get cited more often as opposed to their male colleagues. This proves that if female researchers were to receive more funding than their male colleagues, they would get cited even more in the future.

## 5. Summary of Results

In this section the results of our analysis are discussed. As we stated previously, we defined productivity as number of publications, and we defined scientific impact as number of citations. Our objectives looked at these two concepts and how they are affected by gender and funding.

The first variable we addressed was average citations in the past three years. There was a negative impact on productivity by past citations, meaning that researchers who received more citations in the past three years published fewer papers in the upcoming year. This could mean that there is a break period between publishing articles. However, citations from the previous three years boosted the number of citations in the upcoming year. This means that researchers who received citations in the past continue to get more cited. This could be due to the reputation gained in the field, which could be explained by the Matthew effect (Merton 1968).

The second variable we examined was the average number of publications from the previous three years by researchers. There was a positive impact from previous publications, meaning researchers who published in the past continued to publish in the future. Simply put, productive researchers continued to be productive. Nevertheless, publications from the previous three years had a very small positive effect on citations.

The third variable we studied was career age which we defined as the time difference between the researchers' first publication and the current year. We concluded that as career age increases researchers tend to become less productive. Additionally, career age has a very small negative impact on citations.

The fourth variable we analyzed was average team size which we defined as the average number of co-authors that are co-publishing. Publications with more than one author are more common. This means that most researchers prefer to collaborate with other peers for their research projects. Team size did not affect productivity but increased citations.

However, the main variables of interest were gender and funding from the previous three years. Men received higher funding amounts, on average, compared to their female colleagues though the difference in the average amount of funding was not substantial. Funding had a positive impact on both number of publications and citations. This is in line with past findings (e.g. Ebadi and Schiffauerova 2016). Men produced a significantly higher number of publications *in total* compared to their female counterparts. This was expected since there are more male researchers in these fields. However, both genders produced the same number of publications *on average* even though females received less funding. The results for citations were similar. Given the higher number of male researchers in these fields, male researchers received more citations *in total* compared to female researchers. but both genders received a similar number of citations *on average*. We speculate that if female researchers received an equal amount of funding, they would have the potential to be even *more* productive than their male counterparts.

## 6. Conclusion

Despite efforts to close the gender gap present in technology, engineering, and computer science fields, the issue persists and is reflected in lower salaries, access to fewer research resources and lower numbers of females in high academic positions. In this research, we used various techniques including descriptive statistical analysis and regression analysis to study the influence of funding and gender on research output of active researchers funded by NSERC from 1982 to 2018 and address our two main research objectives: (1) Assess the impact of gender and funding on the number of publications (productivity); (2) Explore the impact of gender and funding on the number of citations (scientific impact).

Regarding our first objective, we concluded that funding positively impacted productivity. However, we also observed that as career age increased, researchers became less productive even though funding typically increased. The results of the impact of gender on productivity showed that while female researchers are underrepresented and thus produced a lower number of papers in total compared to their male counterparts, a comparable amount of productivity was observed between the genders on average.

In terms of our second objective, we found that funding positively affected citation counts but we also observed that as career age increased the number of citations diminished slightly. Additionally, the results of the impact of gender on citation rate indicated that, on average, males and females received a similar number of citations and thus made a comparable scientific impact.

One of our most noteworthy insights of this thesis is that while female researchers published a comparable number of papers and got cited similarly to male researchers on average, they received lower amounts of funding – demonstrating that they are more efficient than their male counterparts.

This research contributed to the findings of previous gender and funding studies while combining the two factors and expanding the analysis to researchers in several fields in Canada.

Therefore, this research leads to the two main policy implications. First, it is suggested that NSERC shows greater policy concerns for underrepresentation of women in natural sciences and engineering and enhance gender-related funding strategies and initiatives. The appropriate gender-responsive policies would help attract and support women in these male-dominant research fields and to make advancements towards closing the gender gap. Second, we also suggest that funding agencies prioritize supporting early-career researchers, especially those who demonstrate high motivation and great potential for growth versus allocating more funding to senior researchers. A more balanced state of funding could accelerate innovation and give the early-career researchers the opportunity to gain recognition for their ideas and the potential for further collaboration with their peers in the scientific community.

## 7. Limitations and Future Work

There were some limitations during this analysis. First, the NSERC awardees dataset was lacking information such as values for certain attributes. It is worth noting that industrial funding for some researchers is much bigger than public funding. This could not be addressed in this study since we did not have access to the necessary funding data for those researchers. Second, we used the Scopus database to gather information about the publications by the NSERC awardees, which had shortcomings including missing authors and the limitations of an English-dominant platform. Publications in other languages e.g., French, were underrepresented (Okubo, 1997). Overall, this resulted in some missing information in the final Scopus dataset. The third major limitation was that Scopus data was less complete before 1996 (Ebadi and Schiffauerova, 2016). This led to a selection of a more recent period (1999-2018) for the regression analysis which covers most of the data. Although Scopus is considered a reliable source, other similar databases could be considered for upcoming studies. In the future, it is recommended that the focus of studies expands to comparing different funding programs, funding strategies and even other funding agencies besides NSERC. There is also an opportunity to focus on different research areas to see if there could be differences between them. In this study, linear regression analysis was primarily used which does not account for non-linearities in the variables. Other types of regression analysis that account for non-linear values should be considered. Another potential methodology for this kind of analysis includes machine learning algorithms for predicting where funding agencies could achieve the best results.

## Resources


1. Abramo, G., "'Angelo, C.A., Murgia, G. (2013). Gender differences in research collaboration. Journal of Informetrics, 7(4), 811-822. DOI: 10.1016/j.joi.2013.07.002
2. Ashkan Ebadi, Andrea Schiffauerova (2015). Bibliometric Analysis of the Impact of Funding on Scientific Development of Researchers, International Journal of Social, Education, Economics and Management Engineering Vol:9, No:5, 2015
3. Gök, A., Rigby, J. & Shapira, P. (forthcoming). The Impact of Research Funding on Scientific Outputs: Evidence from Six Smaller European Countries. Journal of the Association for Information Science and Technology. doi: 10.1002/asi.23406
4. Ashkan Ebadi, Andrea Schiffauerova (2015). How to become an important player in scientific collaboration networks Journal of Informetrics 9 (2015) 809–825
5. Belen Alvarez-Bornstein, Maria Bordons (2021). Is funding related to higher research impact? Exploring its relationship and the mediating role of the collaboration in several disciplines Journal of Informetrics 15 (2021) 101102
6. Ashkan Ebadi, Andrea Schiffauerova (2013). Impact of Funding on Scientific Output and Collaboration. A Survey of Literature Journal of Information & Knowledge Management, Vol. 12, No. 4 (2013) 1350037 DOI: 10.1142/S0219649213500378
7. Ebadi A, Schiffauerova A (2015) How to Receive More Funding for Your Research? Get Connected to the Right People! pLoS ONE 10(7): e0133061. doi: 10.1371/journal.pone.0133061
8. Wang J, Shapira P (2015) Is There a Relationship between Research Sponsorship and Publication Impact? An Analysis of Funding Acknowledgments in Nanotechnology Papers. pLoS ONE 10(2): e0117727. doi: 10.1371/journal. pone.0117727
9. Leila Tahmooresnejad, Catherine Beaudry, Andrea Schiffauerova (2015). The role of public funding in nanotechnology scientific production: Where Canada stands in comparison to the United States, Scientometrics 102:753–787 DOI: 10.1007/s11192-014-1432-2
10. Ashkan Ebadi, Andrea Schiffauerova (2016). How to boost scientific production? A statistical analysis of research funding and other influencing factors, Scientometrics 106:1093–1116 DOI: 10.1007/s11192-015-1825-x
11. Vincent Lariviere, Etienne Vignola-Gagne, Christian Villeneuve, Pascal Gelinas, Yves Gingras (2011). Sex differences in research funding, productivity, and impact: an analysis of Quebec university professors, Scientometrics DOI 10.1007/s11192-011-0369-y
12. Keyvan Vakili, Michael Blomfield (2016). The Research Money Can't Buy: The Impact of Funding on Scientists' Research Behavior
13. Shen, H. (2013). Ineimpact quantified: Mind the gender gap. Nature News, 495(7439), 22. https://doi.org/10.1038/495022a
14. Witteman, H. O., Hendricks, M., Straus, S., & Tannenbaum, C. (2019). Are gender gaps due to evaluations of the applicant or the science? A natural experiment at a national funding agency. The Lancet, 393(10171), 531–540. https://doi.org/10.1016/S0140-6736(18)32611-4
15. Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. R. (2013). Bibliometrics: Global gender disparities in science." Nature News, 504(7479), 211. https://doi.org/10.1038/504211a

16. Murray, D., Siler, K., Larivière, V., Chan, W. M., Collings, A. M., Raymond, J., & Sugimoto, C. R. (2019). Author-Reviewer Homophily in Peer Review. BioRxiv, 400515. https://doi.org/10.1101/400515
17. Zeng, X. H. T., Duch, J., Sales-Pardo, M., Moreira, J. A. G., Radicchi, F., Ribeiro, H. V., Woodruff, T. K., & Amaral, L. A. N. (2016). Differences in Collaboration Patterns across Discipline, Career Stage, and Gender. PLOS Biology, 14(11), e1002573. https://doi.org/10.1371/journal.pbio.1002573
18. Centra, J. A. (1983). Research productivity and teaching effectiveness. Research in Higher Education, 18(4), 379–389
19. Okubo, Y. (1997). Bibliometric indicators and analysis of research systems: Methods and examples. No. 1997/1. Paris: OECD Publishing
20. Payne, A. A., & Siow, A. (2003). Does federal research funding increase university research output? Advances in Economic Analysis and Policy, 3(1), 1–24
21. Jacob, B., & Lefgren, L. (2007). The impact of research grant funding on scientific productivity
22. Lawani, S. M. (1986). Some bibliometric correlates of impact in scientific research. Scientometrics, 9(1–2), 13–25
23. Moed, H. F. (2006). Citation analysis in research evaluation. Dordrecht: Springer
24. Adler, R., Ewing, J., & Taylor, P. (2009). Citation statistics. Statistical Science, 24(1), 1
25. Merton, R. K. (1968). The Matthew effect in science. Science, 159(3810), 56–63
26. Ubfal, D and A Maffioli (2011). The impact of funding on research collaboration: Evidence from a developing country. Research Policy, 40, 1269–1279
27. Merton, R. K. (1973). The sociology of science: Theoretical and empirical investigations. Chicago: University of Chicago press
28. Kyvik, S., & Olsen, T. B. (2008). Does the aging of tenured academic staff affect the research performance of universities? Scientometrics, 76(3), 439–455
29. Adams, J. D., Black, G. C., Clemmons, J. R., & Stephan, P. E. (2005). Scientific teams and institutional collaborations: Evidence from US universities, 1981–1999. Research Policy, 34(3), 259–285.
30. https://en.wikipedia.org/wiki/Natural_Sciences_and_Engineering_Research_Council
31. https://en.wikipedia.org/wiki/Scopus
32. Anahita Hajibabaei, Andrea Schiffauerova, Ashkan Ebadi (2022) Gender-specific patterns in the artificial intelligence scientific ecosystem
33. Anahita Hajibabaei, Andrea Schiffauerova, Ashkan Ebadi (2022) Women and key positions influence collaboration networks: analyzing central scientists' profiles in the artificial intelligence ecosystem through a gender lens
34. https://en.wikipedia.org/wiki/SPSS
35. Andrea Schiffauerova, Ashkan Ebadi, 2016 Gender differences in Research Output, Funding and Collaboration
36. Sa´ C, Cowley S, Martinez M, Kachynska N, Sabzalieva E (2020) Gender gaps in research productivity and recognition among elite scientists in the U.S., Canada, and South Africa. pLoS ONE 15(10): e0240903. https://doi.org/10.1371/journal. pone.0240903

37. Holman L, Stuart-Fox D, Hauser CE (2018) The gender gap in science: How long until women are equally represented? pLoS Biol 16(4): e2004956. https://doi.org/10.1371/journal. pbio.2004956
38. Shin H, Kim K, Kogler DF (2022) Scientific collaboration, research funding, and novelty in scientific knowledge. pLoS ONE 17(7): e0271678. https://doi.org/10.1371/journal.pone.0271678
39. P. Dorta-Gonzalez ´ and M.I. Dorta-Gonzalez 2023 Citation differences across research funding and access modalities
40. Mike Thelwall, Subreena Simrick, Ian Viney, Peter Van den Besselaar 2023 What is research funding, how does it influence research, and how is it recorded? Key dimensions of variation
41. Reza Sattari, Jung Bae, Enrico Berkes, Bruce A. Weinberg 2022 The ripple effects of funding on researchers and output
42. Rachel Heyard, Hanna Hottenrot 2021 The value of research funding for knowledge creation and dissemination: A study of SNSF Research Grants
43. Coleman, J. S., & Lazarsfeld, P. F. (1981). Longitudinal data analysis. New York: Basic Books