

How to boost scientific production? A statistical analysis of research funding and other influencing factors

Ashkan Ebadi¹  · Andrea Schiffauerova^{1,2}

Received: 16 July 2015 / Published online: 4 January 2016
© Akadémiai Kiadó, Budapest, Hungary 2015

Abstract This paper analyzes the impact of several influencing factors on scientific production of researchers. Time related statistical models for the period of 1996 to 2010 are estimated to assess the impact of research funding and other determinant factors on the quantity and quality of the scientific output of individual funded researchers in Canadian natural sciences and engineering. Results confirm a positive impact of funding on the quantity and quality of the publications. In addition, the existence of the Matthew effect is partially confirmed such that the rich get richer. Although a positive relation between the career age and the rate of publications is observed, it is found that the career age negatively affects the quality of works. Moreover, the results suggest that young researchers who work in large teams are more likely to produce high quality publications. We also found that even though academic researchers produce higher quantity of papers it is the researchers with industrial affiliation whose work is of higher quality. Finally, we observed that strategic, targeted and high priority funding programs lead to higher quantity and quality of publications.

Keywords Statistical analysis · Funding · Research output · NSERC · Canada

Introduction

Research requires appropriate amount of investment enabling researchers to hire skillful manpower, as students or research assistants, to purchase the required equipment, tools, or to be able to cooperate with other experts in the field. Hence, research is often expensive

✉ Ashkan Ebadi
a_ebad@encs.concordia.ca; ashkan.ebadi@gmail.com

¹ Concordia Institute for Information Systems Engineering (CIISE), Concordia University, Montreal, QC, Canada

² Department of Engineering Systems and Management, Masdar Institute of Science and Technology, Abu Dhabi, United Arab Emirates

(Gulbrandsen and Smeby 2005). Billions of dollars are being annually spent on the research and development (R&D) activities through the federal funding agencies. Universities, colleges, and research institutes are the key players in knowledge production (Gulbrandsen and Smeby 2005). In 2013, the Natural Sciences and Engineering Research Council of Canada (NSERC), as one of the major Canadian federal granting agencies,¹ invested more than one billion dollars in research by funding more than 29,000 students, over 11,000 university professors, and about 2400 Canadian-based companies (NSERC 2013). However, better access to the funding resources can make prominent researchers more productive, bringing them gradually more credibility and resources. This process is called *credibility cycle* in the literature (Latour and Woolgar 1979).

On one hand, substantial financial resources are annually invested in promoting scientific activities with the aim of advancing scientific development. On the other hand, researchers and their projects are highly dependent on funding as one of the main research drivers. This highlights the importance of an effective funding allocation procedure in a way that the available money is efficiently distributed among the most competent scientists. Therefore, the evaluation of researchers' performance in regards to the amount of funding that they have received as well as the assessment of existing funding allocation strategies is essential. This can help the decision makers to either set new strategies or modify the current ones in order to support the research activities and scientific development. However, evaluating the relation between the research input (e.g. research funding) and the quantity (e.g. number of publications) and quality/impact (e.g. number of citations) of the research output has been a challenging issue for policy makers where a number of techniques (e.g. bibliometrics, statistical analysis) have been used for this purpose (King 1987).

Several studies analyzed different aspects of the relation between funding and scientific production at various levels.² In an early case study performed by McAllister and Narin (1983) for the National Institute of Health (NIH), the relation between NIH's funding and number of publications of the U.S. medical schools was investigated. Using bibliometric indicators, they found a quite strong relationship between the funding and the number of papers published. Payne and Siow (2003) analyzed the impact of federal funding on scientific production of 74 research universities. Employing a regression analysis on a panel data set spanning from 1972 to 1998, they investigated the effects of funding on the articles publication and patents registration by the researchers. Their results show a small positive impact of funding on the number of patents while the effect on the number of articles is relatively higher (\$1 million leads to 11 more articles and .2 more patents). In an econometric study, Huffman and Evenson (2005) used a panel data for 48 U.S. states from 1970 to 1999 to evaluate the relation between funding composition and agricultural productivity. According to their results, federal competitive grant funding had a significant negative impact on the productivity of public agricultural researchers. Jacob and Lefgren (2011) analyzed the effectiveness of government expenditures in R&D. Their database contained researchers who were funded by NIH in 1980–2000 and they used OLS regression to perform the analysis. According to their results, NIH grants had a small positive impact on the publication rate such that the receipt of a grant of roughly \$1.7 million in a given year may lead to only about one additional publication over the next 5 years. This positive impact was higher for postdoctoral fellows.

¹ For more information about NSERC see http://www.nserc-crsng.gc.ca/index_eng.asp.

² For more information, see the review by Ebadi and Schifffauerova (2013).

A number of researchers analyzed the impact of financial investment on scientific production at cross-country level (e.g. Leydesdorff and Wagner 2009; Crespi and Geuna 2008). Shapira and Wang (2010) investigated the impact of nanotechnology funding. They used Thomson Reuter's database for the period of August 2008 to July 2009 and used simple bibliometric indicators to give a general picture of countries which are working in nanotechnology field. They argued that as an impact of large investment that has been made, China is getting closer to the U.S. in terms of the number of publications but Chinese papers still have lower quality in comparison with the Americans and Europeans.

A few other studies investigated the effect of funding on the quality of scientific output. Peritz (1990) focused on the citation impact of funded and unfunded research in the field of economics and found that even if both funded and unfunded research works are published in a high-impact journal, the funded research will be more cited. In another study, Lewison and Dawson (1998) focused on the biomedical field and observed a positive relation between number of the funding bodies and the quality of the publications measured by the mean impact factor of the journals the research was published in. In a more recent study, Sandström (2009) used bibliometrics to analyze the relation between funding and research output in Sweden and found no relation between funding and quality of publications.

Funding can also have an indirect effect on the scientific productivity. One example could be the stimulation of scientific collaboration (Ebadi and Schiffauerova 2015c). According to Adams et al. (2005) public funding had a significant positive impact on the team size of the researchers affiliated with the top 110 American universities. Collaboration itself can influence the scientific output. De Solla Price and Beaver (1966) studied the publications and collaboration of 592 researchers and found a good correlation between their collaboration and scientific output. Lee and Bozeman (2005) did a statistical survey study of 443 academic researchers in the United States. According to their results, the number of peer-reviewed journal articles of the researchers is significantly related to the number of their collaborators. In another study, Martín-Sempere et al. (2002) analyzed the impact of intramural and extramural collaboration on productivity. They found that researchers who belong to no scientific group show lower productivity and they tend to collaborate less internationally.

The evaluation of research performance in Canada has started attracting the attention of the policy makers recently. In Canada, scientific articles have been recognized as the main output of researchers and universities (Godin 2003) and bibliometrics has been mostly used for scientific evaluation purposes. Gingras (1996), in a report to the Program Evaluation Committee of NSERC, discussed the feasibility of bibliometric evaluation of the funded research. Following his study, a few other Canadian researchers used bibliometrics for analyzing the funding impact (e.g. Godin 2003; Campbell et al. 2010; Campbell and Bertrand 2009) that mostly found a positive relation between funding and productivity. Thorsteinsdóttir (2000) studied and compared external research collaboration in two small regions, Iceland and Newfoundland (Canada). Using bibliometric analysis, he found that apart from getting access to financial resources, researchers in the mentioned regions do collaborate to share the research material and equipment needed in order to be able to work in the wider scientific world.

In three recent studies, Beaudry and her colleagues focused on the scientific production of the Canadian researchers working in biotechnology and nanotechnology fields. Beaudry and Clerk-Lamallice (2010) included network structure variables in their regression model, as well as grant and contract amounts of researchers. According to their results, contracts had no negative effect on the publication output of researchers working in biotechnology. However, a positive effect of funding and strong network position on the scientific output

was observed. Using a similar model, Beaudry and Allaoui (2012) added number of patents and age of the researchers variables to their model and assessed the impact on the scientific output of researchers in the field of nanotechnology. Although they found a positive effect of public funding on scientific publications, the effect of private funding on scientific output was nonexistent. In the latest study, Tahmooresnejad et al. (2015) used a similar model to the two previously mentioned studies to make a comparison between Canada and the United States in terms of the impact of public grants and scientific collaboration on the output of researchers working in nanotechnology. Their results suggest a positive impact of both network position and funding on the scientific output.

As discussed, although most of the studies in the literature have found a positive relation between funding and scientific output regardless of intensity of the relation (e.g. Godin 2003; Payne and Siow 2003; Jacob and Lefgren 2007), there also exist some studies that found no significant relation (e.g. Beaudry and Allaoui 2012³; Carayol and Matt 2006) or even a negative impact (e.g. Huffman and Evenson 2005). Hence, results are inconsistent which might be mainly due to the different scopes of the research and the datasets that were used. Apart from research funding, other factors may also play a role in achieving higher scientific productivity. The rate and quality of the past publications of a researcher might have a significant influence on his/her next year scientific productivity, as productive researchers are expected to at least maintain their level of scientific production. Another example would be the impact of the career age on a researcher's productivity since more senior researchers in general are expected to be more productive (Merton 1973; Kyvik and Olsen 2008). Other reasons for their higher expected productivity would be the better accessibility to financial and expertise sources since senior researchers are in general more reputable and have a more established collaboration network. We note that almost all of the previous studies have focused on a very limited scope (e.g. one scientific field, one university) and used only a few of the influencing factors for the analysis (e.g. funding, group size). This could be mainly due to the fact that increasing number of independent variables in a limited dataset augments the risk of having correlation among them.

The main contribution of this paper consists in further and comprehensive investigation of the inter-relations between the scientific output and several influencing factors. For this purpose, we focused on the performance of all the NSERC funded researchers within the period of 1996 to 2010. In addition, a unique data gathering procedure was used that increased the accuracy of the results. “Data and methodology” section presents the detailed information in this regard. The large target dataset enabled us to evaluate the inter-relations for the Canadian researchers who are working in natural sciences and engineering, by including several variables of different types (e.g. funding, past productivity, collaboration, career age, etc.) in the regression model. To our knowledge this is the first study that comprehensively investigates the individual scientific production of researchers covering various scientific fields and at a country level. The remainder of the paper proceeds as follows: “Data and methodology” section presents the data, methodology, and the model; “Results” section presents the empirical results and interpretations; “Conclusion” section concludes; and “Limitations and future works” section discusses the limitations.

³ They found no impact of private funding but positive impact of public funding.

Data and methodology

Data

Three large datasets, which are the databases of funding, funded researchers' publications, and articles' quality/impact, were integrated in this research. Since public funding is the main input source for the university research in Canada (Niosi 2000), it is NSERC, the main public funding agency in natural sciences and engineering, which was selected as the focal funding organization for this research. The main reasons for choosing NSERC was its role as the main federal funding organization in Canada, and the fact that almost all the Canadian researchers in natural sciences and engineering receive at least a basic research grant from NSERC (Godin 2003). In addition, one of our other motivations was the availability of NSERC funding data to the public. Moreover, full names of researchers (both first and last names) are listed in NSERC that helped us to perform the entity disambiguation and integrate the funding and publication data. Elsevier's Scopus⁴ was selected as the source of scientific publications. It provided us with the necessary data on the articles, e.g. co-authors, their affiliations, year of publication. The body of the articles was searched for the support of NSERC and all the papers that had acknowledged the support were extracted for the period of 1996 to 2010. This was a crucial step in gathering more accurate data since the common procedure in the similar studies is extracting the funded researchers' data and then gathering all the articles that were published by those researchers. This must have resulted in an over-estimation of the number of articles, as researchers usually use several sources of funding. The acknowledgement-based search was grounded in the assumption that all the NSERC grantees acknowledge the source of funding in the article.⁵ The reason for selecting the time interval from 1996 to 2010 was lower coverage of Scopus before 1996 (e.g. lack of citation data). In total, 120,439 articles authored by 36,124 distinct authors from 1996 to 2010 were collected.

Having collected the funding and publications databases, the next crucial step was integrating them. We did several preprocessing tasks on the collected data such as correcting special characters, parsing affiliations, and detecting the research area. Detecting the research domain of the funded researchers helped us to perform the entity disambiguation more accurately. In particular, we used Latent Dirichlet Allocation (LDA) technique⁶ to extract keywords from the title of the articles and detect the research domain of the funded researchers. After performing the preprocessing stage, a crucial step was matching authors in the publications database with the funded researchers in the funding database. Here, we faced with two particular problems: (1) The case of having various but almost similar name formats, for example, "Alan Smith", "A. Smith" and "A. C. Smith". We needed to determine if they are all pointing to the same author, and (2) The movement of researchers, that is if "Alan Smith" in "McGill University" is the same person as "Alan Smith" in "University of Toronto". To address the mentioned problems, we designed and coded a semi-automatic machine learning entity disambiguation system. We had the

⁴ Scopus is a commercial database of scientific articles that has been launched by Elsevier in 2004. It is now one of the main competitors of Thomson Reuter's Web of Science.

⁵ According to NSERC policies and regulations, researchers are required to acknowledge the source of funding in their publications. We held more than 30 interviews with randomly selected funded researchers and almost all of them approved such assumption and confirmed that they do acknowledge NSERC in their papers.

⁶ Machine learning technique for topic modeling, first introduced by Blei et al. (2003).

advantage of availability of clean NSERC data which contained the full names and affiliations of the funded researchers. In addition, our Scopus publication dataset contained the current and past affiliations of authors. We defined the similarity measure based on various factors including name of researcher, his/her affiliations (including past affiliations) and research area, and used it to perform the entity disambiguation between funding and publication databases. A JAVA program was coded implementing the Nearest Neighbor algorithm. The assignment procedure was designed semi-automatic due to the difficulties and complexities in entity disambiguation task which might harm the quality of the output. Thus to minimize the error margins, the program asked the user to confirm the match for the cases with the similarity score lower than a pre-defined threshold. At the end of this stage, the same Scopus-id was assigned to the matched records in funding and publication databases.

We used SCImago⁷ for collecting the journal rankings information of the collected articles. SCImago does not provide the impact factor data before 1999; hence we considered 1999 data for the articles published in the period of 1996 to 1999. For the rest of the articles we used the ranking of the journal in the year that the article was published in. SCImago was selected for three main reasons. First, it provides yearly data of the journal rankings that enabled us to perform a more accurate analysis, since we considered the rankings of the journal in the year that an article was published not its impact in the current year. Second reason was the high compatibility of SCImago with our publications database as it is powered by Scopus. And lastly, it is an open access resource with high coverage and quality. It covers more journals than the popular Web of Science and provides a wider variety of countries and languages (Falagas et al. 2008). In the next section, the models and variables are discussed.

Model specification and variables

This paper investigates the impact of some influencing variables on the quantity and quality of publications of the funded researchers. The models and variables that were used for each of the estimations are presented in the following sections. STATA 12⁸ data analysis and statistical software was used to estimate the models.

Quantity of the publications model

Since the purpose of this article is to study the impact of funding, past productivity related variables and some other determinant factors on the scientific productivity of the funded researchers, we consider the number of articles in a given year as the dependent variable (*noArt*). This measure has been widely used in the literature as a proxy of the scientific productivity (e.g. Centra 1983; Okubo 1997). Our dependent variable is therefore a count measure. Hausman et al. (1984) proposed the Poisson model for a count measure. Although the best matching regression model is Poisson, in reality it is rare to satisfy the Poisson assumption on the actual distribution of a natural phenomenon, because most of the time, an over-dispersion or under-dispersion is detected in the sample data. This causes the Poisson model to underestimate or overestimate the standard errors and thus results in misleading estimates for the statistical significance of variables (Coleman and Lazarsfeld 1981). According to Hausman et al. (1984), in order to obtain robust standard errors

⁷ www.scimagojr.com.

⁸ For more information see: <http://www.stata.com/stata12/>.

correcting the estimates, binomial regression can be employed. Therefore, we use negative binomial regression to estimate the number of papers published in a given year by an individual. The regression model in the reduced form is estimated as follows:

$$noArt_i = f(avgFund3_{i-1} + avgIf3_{i-1} + noArt_{i-1} + avgCit3_{i-1} + avgTeamSize_i + careerAge_i + dProvince_i + dInst_i + dFundProg_i) \tag{1}$$

In the model, $avgFund3_{i-1}$ is the average amount of funding that researcher has received over the past 3 years. In the literature, three-year (e.g. Payne and Siow 2003) or 5 year (e.g. Jacob and Lefgren 2007) time windows have been considered for the funding to take effect. We considered both for our model and found that the three-year time window is better suited based on the correlations observed and the intensity of the relation. We also considered a 3 year time window and calculated the average impact factor of the journals in which the author has published articles ($avgIf3_{i-1}$) as a proxy for the quality/impact of his/her papers. As another measure for the quality of the papers, we added $avgCit3_{i-1}$ variable to the model that is the average citations for the articles in the past 3 years. Although both mentioned indicators reflect the quality/impact of a publication, they slightly differ. Impact factor-based indicators mainly reflect the respectability of a journal by authors and reviewers. However, citation counts mainly indicate the importance and the impact of a work within a scientific community. Like all other indicators, the mentioned proxies have some drawbacks, hence, we decided to include both of them. $AvgTeamSize_i$ represents the average number of co-authors in an author’s papers in a given year. We also considered the past productivity of the funded researcher represented by $noArt_{i-1}$ in the model, because we assumed that productive researchers will more likely remain productive. If a researcher has been productive before and has produced high quality publications it can be assumed that this trend will continue. In general, older researchers can be more productive (Merton 1973; Kyvik and Olsen 2008) due to several factors e.g. the better access to the funding and expertise sources, more established network, better access to modern equipments, etc. Hence as a proxy for the career age of the researchers, we included a control variable named $careerAge_i$ representing the time difference between the date of their first article in the database and the given year.

Dummy variables of different types were also added to the model. The dummy variable $dProvince_i$ represents different Canadian provinces and checks for the location impact of a researcher on his/her productivity. Since most of the high ranking Canadian universities are located in Ontario, Quebec, British Columbia, and Alberta provinces, and considering the universities as the key players in scientific activities, it is expected that researchers from the mentioned four provinces show higher average productivity. Moreover, researchers were categorized according to their affiliation as academic or non-academic (industrial) and another dummy variable ($dInst_i$) was added to the model to reflect the impact of the type of affiliation of a researcher on productivity. Since the research activities are mainly performed at the universities, academic researchers are expected to have higher productivity than their industrial counterparts. And to see the impact of different NSERC funding programs, $dFundProg$ was included in the model. It is expected that better targeted and limited scope programs that allocate funding to high priority projects result in higher productivity of researchers.

Quality of the publications model

To investigate the impact on the quality of funded researchers’ papers, we considered the average amount of citations for all the articles of a funded researcher in year i as the

dependent variable ($avgCit_i$). It is argued in the literature that citations counts can be considered as one of the measures of the quality of publications (e.g. Lawani 1986; Moed 2006). Although some problems are associated with citation counts as a quality proxy of the publications (e.g. self citations, negative citations), it is still a common indicator and is widely used as it provides useful information about the publications' quality (Adler et al. 2009). The following regression model (reduced form) is used:

$$avgCit_i = f(avgFund3_{i-1} + avgArt3_{i-1} + avgIf3_{i-1} + avgCit3_{i-1} + avgTeamSize_i + careerAge_i + dProvince_i + dInst_i + dFundProg_i) \quad (2)$$

The definition of the variables (both independent and dummies) are the same as the ones for model (1) except for $avgArt3_{i-1}$ that is the average number of publications for a funded researcher in the period of $[i-1, i-3]$, if the research has been funded in year i . For both quantity and quality models, two datasets were used: (1) including all the researchers and (2) excluding students. Since NSERC funding also covers university students, we used both datasets to be able to compare the results. In addition, by excluding students from the data, we focused more on the professional researchers and their performance as the main purpose of the analysis.

Results

Results of the analyses are presented in two sections. In the first section, the results of visualization analysis and descriptive statistics are shown. The second section discusses the results of the regression analysis.

Visualization and descriptive analysis

Data visualizations are used to find some preliminary patterns in the data. Figure 1 shows the trend of funding over the examined period. We adjusted the amount of total funding based on the constant Canadian dollar in 2003 to remove the general effects of expenditure increase (inflation). As it can be seen, a significant raise is observed from 2001 to 2007. After 2007, the trend of inflation adjusted total funding is almost constant maintaining its level at around \$900 million.

In Fig. 2 we analyzed the trend of average inflation adjusted funding invested in each article produced. The figure reveals that the cost of articles has been on average decreasing after 1999. In other words, the increase in the number of articles have been much more than the increase in the amount of funding, hence, it seems that especially after 1999 researchers have continuously and significantly produced more publications.

We applied visualization techniques in order to explore geographical and career age aspects of funding and productivity of the researchers. In the following figures of this section, number of articles per year (Y -axis) and total funding per year (X -axis) are normalized to a value between 0 and 1. As expected, a considerable share of articles and funding belong to Ontario, Quebec, British Columbia, and Alberta (Fig. 3). In addition, as seen the share of researchers who are located in the Prince Edward Island or the Canadian territories is negligible as there is no reddish circle in the figure. We divided the funded researchers into three categories (junior, middle, and senior), each with several levels based on their career age defined in “Data and methodology” section. In Fig. 3, the size of the circles represents the career age. As it can be seen, interestingly, it seems that not only the

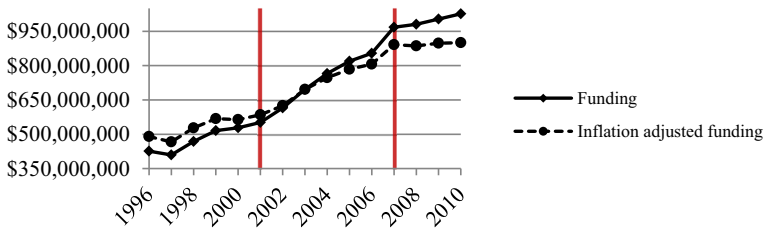


Fig. 1 Trend of total funding and inflation adjusted funding, 1996–2010

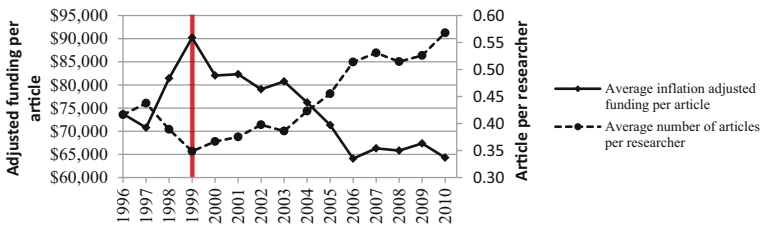


Fig. 2 Average inflation adjusted funding versus articles per researcher, 1996–2010

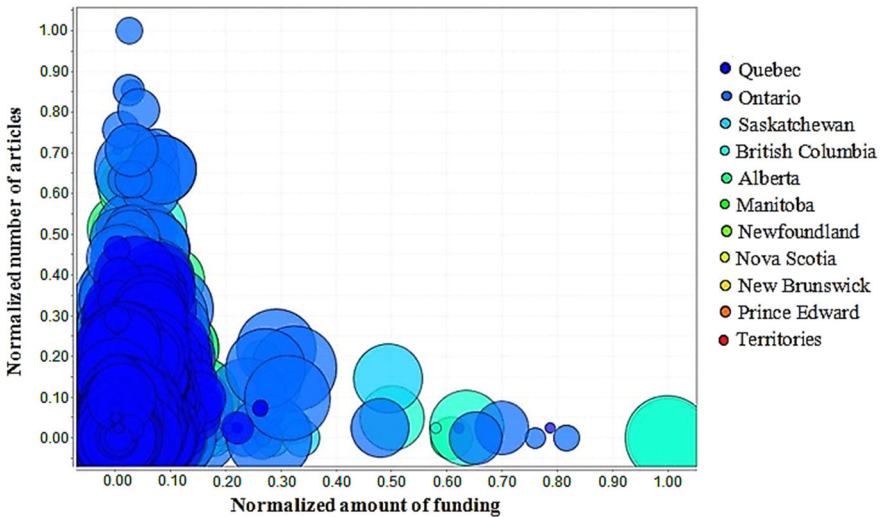


Fig. 3 Funding versus number of articles in Canadian provinces according to the career age as circle sizes, 1996–2010

researchers from the mentioned provinces have been more productive but also the senior researchers are more located in these provinces, as the figure is dominated by blue color. From the figure nothing can be said about the relation between funding and productivity while it seems that funding is a bit biased towards senior researchers.

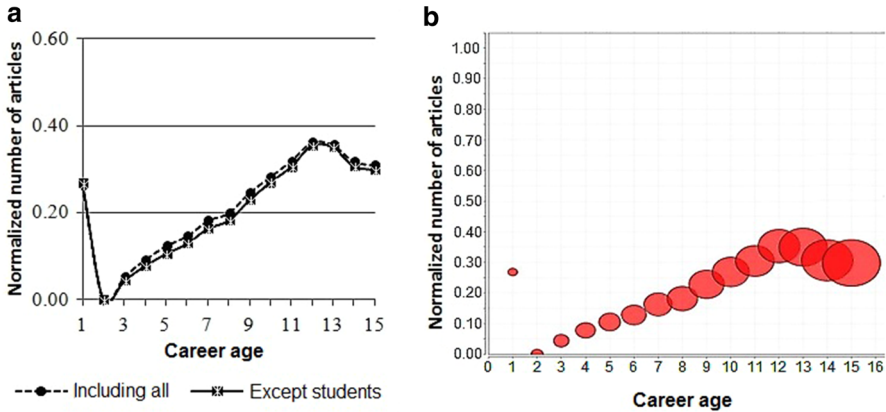


Fig. 4 **a** Career age versus normalized number of publications. **b** Career age, normalized number of publications, and funding as circle sizes

The career age of researchers is used as the control variable and was calculated such that it was set as 1 for the researchers who started publishing in 2010. Career age increases respectively as we move backward in the time axis. Figure 4a shows the interaction of the career age variable with the number of articles. The number of articles was normalized to a value between 0 and 1 by dividing the number of articles by the highest number of articles. We considered two cases, where one is including all the funded researchers while we excluded the students⁹ in the second case. As seen, both curves have exactly the same trend that indicates a positive relation between age of the researcher and his/her productivity. In other words, it seems that as the career age of the researcher grows, his/her productivity also increases and peaks at a certain age which is highly dependent on the discipline (Beaudry and Allaoui 2012). The observed relation between age and productivity is in line with Lehman (1953) and Lee and Bozeman (2005). In addition, the curves imply non-linear effects for which we will consider a quadratic variable in our regression. We added the funding data to the analysis, which is represented in Fig. 4b as the size of the circles. The figure is suggesting that there is a positive relation between age and funding until a certain age while a negative relation is observed afterwards. Hence, from Fig. 4 it can be said that career age of the researchers and the amount of funding allocated to them might have a positive relation with their number of publications until a certain age. Since the relation follows an inverted U shape it can be said that mid-career researchers are the most efficient ones in producing articles based on their available amount of funding.

Statistical analysis

In this section, the regression results are presented and discussed for the quantity and quality of the publications and the factors that affect them. In each of the following sections, as discussed before, two models are analyzed; one includes all the researchers (complete model) and one which does not include the students (student excluded model).

⁹ The NSERC database originally contains both scholarships and grants.

Quantity of the publications

Quantity of publications, complete model

Before running the regression model, we first analyzed the associations between dependent and independent variables. We considered all the combinations of the lags for the variables in the model and used the ones that yielded the most robust results. This is similar to the approach of Schilling and Phelps (2007), and Beaudry and Allaoui (2012). According to Table 1, the absolute value of all the correlation coefficients is lower than .37, which indicates that the degree of linear correlation among the selected variables is very weak.

Apart from the explanation given in “Data and methodology” section in regard to the use of negative binomial predictor for our model, we also tested Poisson model and found that Poisson model does not fit to our data because the goodness of fit Chi squared test was statistically significant. Hence, we employed negative binomial regression on our data to estimate the impact of the considered factors on the scientific productivity of the funded researchers measured by the number of articles in a year. Table 2 shows the result of the regression including all the independent, interaction, and dummy variables.

As it can be seen the average amount of researcher’s funding in the past years has a significant and relatively high positive impact on the scientific production of the researcher. This is in accordance with several studies (e.g. Arora and Gambardella 1998; Boyack and Börner 2003; Godin 2003; Zucker et al. 2007; Beaudry and Allaoui 2012) who found that larger amount of funding will result in higher number of published papers. We used the average impact factor of the journals in which a researcher published his/her articles in the past 3 years as a proxy for the quality of his/her work (*avgIf3*). As it can be observed in Table 2, higher quality of the papers of a researcher in the past 3 years increases the number of published articles. This was quite expected since researchers who published in higher quality journals can have in general higher reputation. Higher reputation can bring higher amount of funding that may enable the researcher to expand his/her activities through finding new partners, working on new projects, etc. with an aim to increase the overall productivity. This relation is also confirmed by the positive overall impact of the career age of the researchers that will be discussed later.

According to the model, past productivity (*noArt1*) of a researcher has also a positive effect on the number of publications. This is also expected since it is more probable that a researcher with higher productivity attracts more funds that in turn might result in higher number of publications. In addition, it is more likely that a productive researcher at least maintains his/her level of productivity in the coming year. Moreover, according to the results the average team size of the researchers (*avgTeamSize*) positively influences their

Table 1 Correlation matrix, complete quantity model

Variable	<i>noArt_t</i>	<i>ln_avgFund3_{t-1}</i>	<i>ln_avgIf3_{t-1}</i>	<i>noArt1_{t-1}</i>	<i>avgTeamSize_t</i>	<i>careerAge_t</i>
<i>noArt_t</i>	1.0000					
<i>ln_avgFund3_{t-1}</i>	.2891	1.0000				
<i>ln_avgIf3_{t-1}</i>	.0603	.0804	1.0000			
<i>noArt1_{t-1}</i>	.3655	.2462	.0706	1.0000		
<i>avgTeamSize_t</i>	.0804	.0032	.0780	.0346	1.0000	
<i>careerAge_t</i>	.1577	.3540	.0409	.1987	-.0026	1.0000

Table 2 Negative binomial regression, the complete model

<i>noArt</i>	Coef.	SE	Z	$P > z $	[95 % Conf. interval]	
<i>ln_avgFund3</i>	.2490258***	.0046262	53.83	.000	.2399587	.2580929
<i>ln_avgIf3</i>	.0330018***	.0061728	5.35	.000	.0209032	.0451003
<i>noArt1</i>	.0378271***	.0004333	87.29	.000	.0369778	.0386764
<i>avgTeamSize</i>	.0026179***	.0001516	17.27	.000	.0023208	.002915
<i>careerAge</i>	-.0297705***	.004317	-6.90	.000	-.0382317	-.0213093
<i>careerAge</i> ²	.0023998***	.0002767	8.64	.000	.0018474	.0029322
Interaction variables						
<i>teamXage</i>	-.000142***	.0000169	-8.40	.000	-.0001752	-.0001089
Affiliations dummy variable						
<i>dAcademia</i>	.2486067***	.030134	8.25	.000	.1895452	.3076683
Provinces dummy variables						
<i>dQuebec</i>	-.0794855***	.0108013	-7.36	.000	-.1006557	-.0583154
<i>dB Columbia</i>	-.0475211***	.0127599	-3.72	.000	-.0725299	-.0225122
<i>dAlberta</i>	.0765243***	.0132953	5.76	.000	.0504661	.1025826
<i>dSaskatchewan</i>	.005897	.0235165	.25	.802	-.0401944	.0519884
<i>dNBrunswick</i>	-.087087***	.0302696	-2.88	.004	-.1464143	-.0277598
<i>dManitoba</i>	.0101386	.0241655	.42	.675	-.037225	.0575022
<i>dNFoundland</i>	-.0472348	.0315774	-1.50	.135	-.1091253	.0146558
<i>dPEdward</i>	-.0771565	.0792196	-.97	.330	-.232424	.078111
<i>dNScotia</i>	-.0693328***	.021324	-3.25	.001	-.111127	-.0275386
Funding programs dummy variables						
<i>dStrategic</i>	.0838271***	.0176068	4.76	.000	.0493183	.1183359
<i>dTools</i>	.0281414	.0249109	1.13	.259	-.0206831	.076966
<i>dCollaborative</i>	-.0115878	.0175147	-.66	.508	-.0459161	.0227404
<i>dIndustrial</i>	.0680523**	.026957	2.52	.012	.0152176	.1208871
<i>dStudent</i>	-.2530735***	.0188543	-13.42	.000	-.2900273	-.2161197
<i>dOther</i>	-.004898	.011799	-.42	.678	-.0280236	.0182275
<i>_cons</i>	-2.85391***	.0536381	-53.21	.000	-2.959038	-2.748781
<i>ln(alpha)</i>	-.6738737***	.013409			-.700155	-.6475925
<i>alpha</i>	.5097302	.006835			.4965084	.5233041
Likelihood-ratio test of $\alpha = 0$			Chibar2(01) = 1.5e + 04	Prob \geq chibar2 = .000		

* $p < .10$, ** $p < .05$, *** $p < .01$, number of observations: 84,048

productivity. Larger scientific team size can enable researchers to better distribute the work among the team members. It would be also possible to work on more or larger projects. Hence, in general we can assume that larger teams have better access to scientific resources (e.g. manpower, equipments, and finance) which will help them to increase the scientific productivity. Although there are also some disadvantages of working in larger teams (e.g. coordination costs), according to our dataset the overall impact of team size on the productivity of the funded researchers is positive. This is in line with some other studies that also found a positive relation between the team size and scientific output (e.g. Ebadi and Schiffauerova 2015a; Plume and van Wiejen 2014).

The career age of the funded researchers, that we employed as the control variable, has an overall positive impact on the number of publications. We first considered the model without the quadratic term that resulted in a positive coefficient for the career age variable (.0062546). As explained in Fig. 4a, non-linear effects were observed for the career age of the researchers. We added the quadratic term to see the curvature of the relationship. Hence, the predictive effect of the researchers' career age is represented by $\beta_1 \text{careerAge} + \beta_2 \text{careerAge}^2$ which is increasing over the range of the career age. Therefore, number of publications increases with the career age of the researchers. From the regression result and the discussion, it seems that our dataset partially verifies the existence of the Matthew effect (Merton 1968) in a sense that higher number of publications of high quality in the past brings more reputation to a researcher, which may result in securing more research funding, which in turn attracts even more funding for the researcher in the future.¹⁰ Although both career age and team size have positive impact on the number of publications, interestingly the interaction variable has a negative and significant effect. This may imply that as the career age of the researchers increases, working in larger team reduces their productivity.

In order to dig deeper into where the NSERC funding has had a stronger effect in terms of the number of publications, we included dummy variables in the regression representing the institution type and Canadian provinces. We also considered dummy variables for NSERC funding programs to compare the impact of the programs. The institution type dummy variable (*dAcademia*) takes value 1 if the funded researcher is affiliated with an academic institution, and 0 if his affiliation is non-academic. According to Table 2, academic funded researchers are significantly different from the non-academic ones and are producing around 25 % (.249) more than the non-academic researchers. Analysis of the provinces dummy variables reveals that the funded researchers from Quebec, British Columbia, Alberta, New Brunswick, and Nova Scotia are significantly different from the ones who reside in Ontario, which was the omitted dummy variable. Interestingly, among the mentioned provinces only the coefficient of Alberta dummy variable is positive (.0765243) which shows higher productivity of Alberta's funded researchers.

To analyze the effect of different NSERC funding programs, we categorized the programs into seven main categories: discovery grants,¹¹ strategic projects, collaborative grants, student scholarships, tools, industry grants, and other programs (the ones that do not belong to any of the other mentioned six types). These names are used in the rest of the paper to address different NSERC funding programs. We considered the discovery grants as the omitted variable. From the analysis, it is observed that the effects of strategic, industrial, and student scholarships are significantly different from the discovery grants program while the effect is only negative for the student scholarships (−.2530735). According to the definition of these grants, the results are quite expected since discovery grants were not extremely competitive during the studied period and most of the Canadian researchers used to receive it. However, the industrial and strategic grants are more targeted, while a lower productivity and a lower level of funding are expected in general for students in comparison with professional researchers.¹² Specifically for the strategic project grants, the aim is to improve the scientific development in selected high-priority areas

¹⁰ The rich get richer.

¹¹ NSERC discovery grants program mainly supports the long term research projects. For more information, see: http://www.nserc-crsng.gc.ca/Professors-Professeurs/Grants-Subs/DGIGP-PSIGP_eng.asp.

¹² As an example of the students' contribution, please refer to Larivière (2012) who particularly studied the research activities of the Ph.D. students and their contribution to the advancement of knowledge.

Table 3 Correlation matrix, student-excluded quantity model

Variable	$noArt_i$	$ln_avgFund3_{i-1}$	ln_avgI3_{i-1}	$avgCit3_{i-1}$	$noArt1_{i-1}$	$avgTeamSize_i$	$careerAge_i$
$noArt_i$	1.0000						
$ln_avgFund3_{i-1}$.2984	1.0000					
ln_avgI3_{i-1}	.0305	.1072	1.0000				
$ln_avgCit3_{i-1}$.1248	.1434	.4084	1.0000			
ln_noArt1_{i-1}	.4478	.2799	.0328	.0249	1.0000		
$avgTeamSize_i$.0856	.0269	.0982	.0665	.0488	1.0000	
$careerAge_i$.1602	.3378	.0193	.2495	.1787	-.0058	1.0000

that influences Canada's economic and societal position. In the next section we remove the students' data and analyze the model for the rest of the funded researchers.

Quantity of publications, students-excluded model

We removed the students' data and performed the regression for the rest of the researchers. For this purpose, we labeled a researcher as "student" in a year whenever his/her highest average grant was coming from one of the student funding programs in that year. Moreover, to better account for the quality of the work of the funded researchers, we also considered the average number of citations of their articles in the past 3 years ($avgCit3_{i-1}$). The correlation matrix presented in Table 3 shows a weak linear correlation degree among the considered variables.

The results of the negative binomial regression for the student-excluded model of the number of publications are shown in Table 4. As seen, average journal impact factor ($avgI3$) has a significant negative impact on the quantity of the publications, while average citations ($avgCit3$) have a positive effect. The size of the effect for the both mentioned factors is almost the same. As mentioned earlier, these proxies slightly differ although they can be both considered as a measure of the quality of publications. Impact factor mainly reflect the respectability of a journal by authors and reviewers, however, average number of citations mainly indicate the importance and the impact of a work within a scientific community. Hence, it can be said that the quality of the papers of the professional scientists¹³ measured by the average number of citations in the past 3 years influences the number of publications positively. The citation-based proxy seems to be a better measure for evaluating the quality of the professional researchers' papers. According to the regression results, researchers with high amounts of funding who publish relatively low quality papers in high quality journals might possess a relatively lower rate of publications in comparison with their counterparts who produce high quality works. These papers would not be highly cited, which justifies the negative coefficient of $avgI3$. In addition, from Tables 2 and 4 it seems that students are one of the key factors in making the impact of the journal factor positive, as the coefficient is positive in the complete model while it becomes negative in the students-excluded one. Hence, it can be said that there is a positive relation between students' productivity and their publishing preferences.

¹³ We call *Researcher* in this paper anybody who is funded by NSERC while *Professional Researchers* are only the ones that are not counted as students.

Table 4 Negative binomial regression, student-excluded (professional) model

<i>noArt</i>	Coef.	SE	<i>z</i>	<i>P</i> > <i>z</i>	[95 % Conf. interval]	
<i>ln_avgFund3</i>	.1948294***	.0056854	34.27	.000	.1836861	.2059726
<i>ln_avgIf3</i>	−.1066736***	.0091611	−11.64	.000	−.124629	−.0887182
<i>ln_avgCit3</i>	.1163277***	.0053229	21.85	.000	.105895	.1267603
<i>ln_noArt1</i>	.5851529***	.0071044	82.36	.000	.5712285	.5990772
<i>avgTeamSize</i>	.0022361***	.0002179	10.26	.000	.0018089	.0026632
<i>careerAge</i>	.066107***	.005784	11.43	.000	.0547707	.0774434
<i>careerAge</i> ²	−.0038581***	.0003453	−11.17	.000	−.0045349	−.0031813
Interaction variables						
<i>teamXage</i>	−.0000895***	.0000239	−3.74	.000	−.0001364	−.0000426
Affiliations dummy variable						
<i>dAcademia</i>	.2690371***	.0377299	7.13	.000	.195088	.3429863
Provinces dummy variables						
<i>dQuebec</i>	−.058617***	.0127778	−4.59	.000	−.083661	−.033573
<i>dB Columbia</i>	−.0349328**	.0150169	−2.33	.020	−.0643653	−.0055003
<i>dAlberta</i>	.0656808***	.015411	4.26	.000	.0354758	.0958857
<i>dSaskatchewan</i>	.0126419	.0273714	.46	.644	−.0410051	.0662888
<i>dNBrunswick</i>	−.0552766	.0363791	−1.52	.129	−.1265783	.0160251
<i>dManitoba</i>	.0250575	.0283492	.88	.377	−.030506	.0806209
<i>dNFoundland</i>	−.0408263	.039075	−1.04	.296	−.1174118	.0357592
<i>dPEdward</i>	−.0479439	.0933412	−.51	.608	−.2308893	.1350015
<i>dNScotia</i>	−.0525115**	.0250681	−2.09	.036	−.1016441	−.0033788
Funding programs dummy variables						
<i>dStrategic</i>	.0642019***	.0195962	3.28	.001	.025794	.1026098
<i>dTools</i>	.0047521	.0287229	.17	.869	−.0515438	.0610479
<i>dCollaborative</i>	−.0310454	.0200215	−1.55	.121	−.0702868	.008196
<i>dIndustrial</i>	.0729073**	.0302634	2.41	.016	.0135921	.1322224
<i>dOther</i>	−.0377083***	.0135293	−2.79	.005	−.0642253	−.0111914
<i>_cons</i>	−2.535194***	.0658431	−38.50	.000	−2.664244	−2.406144
<i>ln(alpha)</i>	−1.058564***	.019625			−1.097028	−1.0201
<i>alpha</i>	.3469538	.006809			.3338618	.3605591
Likelihood-ratio test of alpha = 0			Chibar2(01) = 6349.65	Prob ≥ chibar2 = .000		

* *p* < .10, ** *p* < .05, *** *p* < .01, number of observations: 43,514

Other interesting finding is the high impact of a researcher’s past productivity (*noArt1*) on the number of publications. Hence, not only the quality of the works in the past plays an important role in higher productivity, but also the rate of publications is a major sign of the productive researchers. The career age of the researchers is also showing a positive impact, while the quadratic term (*careerAge*²) affects negatively. Hence according to the curvature of the relationship, although our study covers 15 years from 1996 to 2010, it can be predicted that around 18 years after the start of the work of a NSERC funded researcher¹⁴ his/her scientific productivity starts to decline. Therefore, mid-career NSERC funded researchers seem to be more productive. This finding is in line with Cole (1979), Wray

¹⁴ Measured by the date of his first publication that is available on SCOPUS.

(2003), Wray (2004), Kyvik and Olsen (2008), and Beaudry and Allaoui (2012) who also found the higher scientific productivity of mid-career aged researchers.

Other estimated factors including the dummy variables are showing the same effect as the ones predicted by the complete model. The only exception is for the dummy variable of New Brunswick province that becomes no longer significant in the students-excluded model, indicating that there is no significant difference between New Brunswick researchers and the omitted province of Ontario. In the next two sections, we estimate the impact of the influencing factors on the quality of the researchers' papers.

Quality of the publications

Quality of the publications, complete model

In this section, relation between the selected influencing factors and quality of the publications is investigated. Number of citations was considered as the quality proxy of the publications. Again, both complete and students-excluded models are studied. The correlation matrix of the considered variables is presented in Table 5, which reports a very weak linear correlation for most of the variables. The absolute value of the correlation coefficients is less than .4.

Since in the quality of papers model the dependent variable (*avgCit*) is not a count measure, we used multiple regression analysis for estimating the impact of the considered factors on the quality of the papers of the NSERC funded researchers. According to Table 6, all the independent variables significantly influence the quality of the papers measured by average number of citations. As expected, past funding (*avgFund3*) has a positive impact on the quality of the papers. This is interesting since in the literature mainly no relation is found between funding and quality of the works (e.g. Godin 2003; Payne and Siow 2003; Tahmooresnejad et al. 2015¹⁵). The past productivity (*avgArt3*) and the quality of the past works of a funded researcher also positively affect the average citations received by his/her papers in the current year. Hence, this is implicitly confirming that productive researchers who published high quality works in the past will likely continue producing high quality papers in the future. As expected, researchers who get involved in larger scientific teams also produce higher quality papers. Hence, it is likely that scientists benefit from the collaboration to increase the quantity and quality of their scientific output through their involvement in larger research teams, where they can have better access to resources (Katz and Martin 1997; Melin 2000; Beaver 2001; Heinze and Kuhlmann 2008), to expertise (Katz and Martin 1997; Thorsteinsdóttir 2000), and to funding (Beaver 2001; Heinze and Kuhlmann 2008). Through scientific collaboration, researchers interact with each other and can criticize team members' work (or duties). This internal referring may result in a higher quality publication thus authors will be more cited (Salter and Martin 2001; Lee and Bozeman 2005; Adams et al. 2005). Another interesting point is the negative relation observed between the career age of the funded researcher and the quality of his/her work. This means that as the career age of the researcher increases he/she produces on average lower quality papers. This can be caused by several factors, e.g. lower motivation, or higher reputation in a way that the papers are published but not necessarily highly cited, etc. This indicate that senior researchers who are highly funded may publish more in high ranking journals but their works are less cited. Also, as the career age of a researcher increases, larger team sizes also influence the quality of papers negatively

¹⁵ They only found a positive impact in the United States.

Table 5 Correlation matrix, complete quality model

Variable	<i>noArt_i</i>	<i>ln_avgFund3_{i-1}</i>	<i>avgArt3_{i-1}</i>	<i>avgIff3_{i-1}</i>	<i>avgCit3_{i-1}</i>	<i>avgTeamSize_i</i>	<i>careerAge_i</i>
<i>avgCit_i</i>	1.0000						
<i>ln_avgFund3_{i-1}</i>	.0536	1.0000					
<i>avgArt3_{i-1}</i>	.0833	.3965	1.0000				
<i>avgIff3_{i-1}</i>	.1394	.1198	.1910	1.0000			
<i>avgCit3_{i-1}</i>	.1845	.0680	.1002	.3028	1.0000		
<i>avgTeamSize_i</i>	.0614	.0016	.0197	.0551	.0307	1.0000	
<i>careerAge_i</i>	.0111	.3548	.3630	.0802	.0986	-.0114	1.0000

Table 6 Regression results, complete quality model

<i>avgCit</i>	Coef.	SE	<i>t</i>	<i>P > t </i>	[95 % Conf. interval]	
<i>ln_avgFund3</i>	.3537779***	.0245286	14.42	.000	.3057021 .4018537	
<i>avgArt3</i>	.2763136***	.0170422	16.21	.000	.2429111 .309716	
<i>avgIff3</i>	.4535849***	.0189189	23.98	.000	.4165041 .4906657	
<i>avgCit3</i>	.1170045***	.0024088	48.57	.000	.1122833 .1217258	
<i>avgTeamSize</i>	.008713***	.0005863	14.86	.000	.0075639 .009862	
<i>careerAge</i>	-.1470174***	.0214365	-6.86	.000	-.1890326 -.1050023	
<i>careerAge²</i>	.0084294***	.0014472	5.82	.000	.0055928 .011266	
Interaction variables						
<i>teamXage</i>	-.000511***	.0000798	-6.40	.000	-.0006674 -.0003546	
Affiliations dummy variable						
<i>dAcademia</i>	-.6036918***	.1373958	-4.39	.000	-.8728855 -.3342981	
Provinces dummy variables						
<i>dQuebec</i>	-.1449621**	.0571347	-2.54	.011	-.2569452 -.032979	
<i>dBcolumbia</i>	.3221105***	.0682374	4.72	.000	.1883662 .4558549	
<i>dAlberta</i>	-.3136267***	.0737739	-4.25	.000	-.4582226 -.1690309	
<i>dSaskatchewan</i>	-.4159663***	.1246093	-3.34	.001	-.6601986 -.1717339	
<i>dNBrunswick</i>	-.7473366***	.1499116	-4.99	.000	-1.041161 -.4535121	
<i>dManitoba</i>	-.6972798***	.1289009	-5.41	.000	-.9499237 -.4446359	
<i>dNfoundland</i>	-.6921608***	.1653934	-4.18	.000	-1.016329 -.3679923	
<i>dPEdward</i>	1.723699***	.4082086	4.22	.000	.9236167 2.523782	
<i>dNScotia</i>	-.2261751**	.1106283	-2.04	.041	-.443005 -.0093452	
Funding programs dummy variables						
<i>dStrategic</i>	.2671592***	.1027175	2.60	.009	.0658344 .4684839	
<i>dTools</i>	.3709673***	.1403731	2.64	.008	.095838 .6460966	
<i>dCollaborative</i>	.0247293	.0964148	.26	.798	-.1642423 .2137008	
<i>dIndustrial</i>	.1869615	.1425337	1.31	.190	-.0924023 .4663254	
<i>dStudent</i>	2.069167***	.0765411	27.03	.000	1.919148 2.219187	
<i>dOther</i>	1.059779***	.0652033	16.25	.000	.9319813 1.187576	
<i>_cons</i>	-.5555149**	.270778	-2.05	.040	-1.086236 -.0247941	

* *p* < .10, ** *p* < .05, *** *p* < .01, number of observations: 111,994

(*teamXage*). Hence, from the results it seems that young researchers who work in large teams are more likely to produce high quality publications.

Analyzing the dummy variable of the institution type reveals that funded researchers who are affiliated with industry are producing on average higher quality papers, measured by the average number of citations. Regarding the provinces, all the Canadian provinces dummy variables are significantly different from Ontario, which is the omitted dummy variable. The coefficient is negative for all the provinces except for British Columbia and Prince Edward. However, nothing can be concluded about the funded researchers located in Prince Edward province since the number of articles, number of researchers, and the total amount of funding is much lower there in comparison with other provinces. Hence, on average researchers located in all the provinces may produce fewer publications with higher quality than their counterparts in Ontario. The only exception is for British Columbia where it seems that its researchers produce the highest quality publications. We omitted the discovery grants dummy variable for analyzing the impact of different NSERC funding programs. As it can be seen, *dStrategic*, *dTools*, *dStudent*, and *dOther* are significantly and positively different from the omitted program. This finding was expected for the strategic funding programs but not expected for the student programs. In general, it can be said that limited scope of a funding program with more narrowly defined targets (e.g. strategic funding programs) can result in higher quality papers. On the other hand, one may not expect a direct positive impact of very general programs like discovery grants since they cover almost all the funded researchers.

Quality of the publications, student-excluded model

In this section, we use the same variables and do the same analysis on the student-excluded data. Table 7 reports the linear correlations among the considered variables. The absolute value of the correlation coefficients is less than .38, which is the correlation between the past average productivity (*avgArt3*) and past average funding (*avgFund3*). We continue with the multiple regression analysis on the data.

Table 8 shows the regression results for the student-excluded quality model. The signs of the resulted variables are exactly the same as the ones in the complete quality model and the coefficients are almost the same as well. Hence, the justifications that were presented in the previous section hold. The only difference is for the career age variable (*careerAge*) and industrial programs dummy variable (*dIndustrial*). According to Table 8, the dummy variable for the industrial funding programs is showing significantly different impact (with

Table 7 Correlation matrix, student-excluded (professional) quality model

Variable	<i>noArt_i</i>	<i>ln_avgFund3_{i-1}</i>	<i>avgArt3_{i-1}</i>	<i>avgIf3_{i-1}</i>	<i>avgCit3_{i-1}</i>	<i>avgTeamSize_i</i>	<i>careerAge_i</i>
<i>avgCit_i</i>	1.0000						
<i>ln_avgFund3_{i-1}</i>	.0998	1.0000					
<i>avgArt3_{i-1}</i>	.1078	.3775	1.0000				
<i>avgIf3_{i-1}</i>	.1601	.1175	.1885	1.0000			
<i>avgCit3_{i-1}</i>	.2149	.0792	.1047	.3035	1.0000		
<i>avgTeamSize_i</i>	.0627	.0326	.0237	.0542	.0329	1.0000	
<i>careerAge_i</i>	.0454	.2842	.3381	.0706	.1072	-.0033	1.0000

Table 8 Regression results, student-excluded quality model

<i>avgCit</i>	Coef.	SE	<i>t</i>	<i>P</i> > <i>t</i>	[95 % Conf. interval]	
<i>ln_avgFund3</i>	.3490946***	.024293	14.37	.000	.3014807	.3967086
<i>avgArt3</i>	.25321***	.0158916	15.93	.000	.2220627	.2843574
<i>avgI3</i>	.4613971***	.0189419	24.36	.000	.4242712	.498523
<i>avgCit3</i>	.1299262***	.0024331	53.40	.000	.1251573	.134695
<i>avgTeamSize</i>	.0113608***	.0007585	14.98	.000	.0098742	.0128474
<i>careerAge</i>	−.0105686	.0205416	−.51	.607	−.0508299	.0296926
<i>careerAge</i> ²	.0004639	.0013778	.34	.736	−.0022366	.0031645
Interaction variables						
<i>teamXage</i>	−.0007191***	.000088	−8.17	.000	−.0008916	−.0005465
Affiliations dummy variable						
<i>dAcademia</i>	−.576718***	.129011	−4.47	.000	−.8295779	−.323858
Provinces dummy variables						
<i>dQuebec</i>	−.1447395***	.0555324	−2.61	.009	−.2535824	−.0358967
<i>dB Columbia</i>	.3213612***	.0669523	4.80	.000	.1901354	.4525869
<i>dAlberta</i>	−.3296383***	.0725502	−4.54	.000	−.4718357	−.1874409
<i>dSaskatchewan</i>	−.3856221***	.1182055	−3.26	.001	−.6173034	−.1539409
<i>dNBrunswick</i>	−.692624***	.1435	−4.83	.000	−.9738822	−.4113658
<i>dManitoba</i>	−.5918489***	.1233531	−4.80	.000	−.8336195	−.3500783
<i>dNfoundland</i>	−.6801265***	.1585457	−4.29	.000	−.9908742	−.3693789
<i>dPEdward</i>	1.950506***	.3946359	4.94	.000	1.177024	2.723987
<i>dNScotia</i>	−.2155348**	.1080504	−1.99	.046	−.4273122	−.0037574
Funding programs dummy variables						
<i>dStrategic</i>	.2788403***	.0947333	2.94	.003	.0931642	.4645164
<i>dTools</i>	.3724214***	.1292664	2.88	.004	.1190607	.625782
<i>dCollaborative</i>	.0372067	.0888975	.42	.676	−.1370313	.2114447
<i>dIndustrial</i>	.2645677**	.1315162	2.01	.044	.0067976	.5223378
<i>dOther</i>	1.024165***	.0612311	16.73	.000	.9041533	1.144178
<i>_cons</i>	−.9746751***	.2648914	−3.68	.000	−1.493859	−.4554912

* *p* < .10, ** *p* < .05, *** *p* < .01, number of observations: 99,216

the coefficient of .26) in comparison with the omitted dummy variable of the discovery grants. The career age of the NSERC funded researchers in the student-excluded model has become insignificant hence it seems that the possible negative effect of the career age is duplicate and has been explained by the other variables of the model.

Conclusion

In this paper, we investigated the impact of funding and other influencing factors like scientific team size and past productivity on the quantity and quality of the publications of the funded researchers. All the four regression models confirmed the significant positive impact of funding on the productivity of the researchers. The positive relation between funding and the rate of publications has been also confirmed in the work of other scholars,

e.g. Arora and Gambardella (1998), Boyack and Börner (2003), Payne and Siow (2003), Jacob and Lefgren (2007), Zucker et al. (2007), and Beaudry and Allaoui (2012). However, to our knowledge the studies that used statistical analysis to assess the quality of the publications of the funded researchers in a large scope are limited. Payne and Siow (2003) performed an econometric study focused on 74 universities and observed no significant relation between funding and research quality. The dataset was limited and old (covering the years of 1972 to 1998). In another recent study, Tahmooresnejad et al. (2015) performed a cross country analysis between the US and Canada and observed a positive impact of funding on the quality of nanotechnology publications just in the United States. We extended the previous two mentioned studies by focusing on the Canadian researchers who are active in all the natural sciences and engineering disciplines and by using a large dataset to analyze the inter-relations at the individual level of researchers. Since a large scope has been used in our research, the observed significant positive relation between funding and quality of the publications could be of interest.

Although our results confirm that higher level of funding may result in higher scientific performance, the financial resources are limited and we cannot simply increase the level of funding for everybody in order to boost scientific output. Instead, allocation strategies have to be well designed and continuously revised. One should note that it may be too early to come up with policy recommendations and further research is required. However, based on our results, we make several suggestions concerned with the allocation strategy which should lead into more efficient funding support.

First we need to start with the most expected and obvious one. Our results confirm that the past productivity of a funded researcher in terms of both quantity and quality of his/her publications is one of the important factors that positively affects the rate and quality of his/her publications. Since past productivity was found to be positively related to the future scientific productivity, a focus on funding of researchers with high quality publication record is advised. Supporting highly productive eminent researchers as one of the main criteria has already been the strategy of many funding agencies. Apart from confirming the validity of this commonly applied strategy we have brought some interesting insights into the various factors playing role in the funded researchers' performance, and based on these we derived some specific and less usual implications for funding allocation policies:

One of the most interesting findings concerns the impact of the career age on productivity of a funded researcher. For the quantity of the publications model it has been observed that mid-career NSERC funded researchers seem to be more productive which is in line with the work of other scholars like Cole (1979), Wray (2003), Wray (2004), Kyvik and Olsen (2008), and Beaudry and Allaoui (2012). In addition, it was observed that the career age negatively affects the quality of published works, which means that as the career of the researchers progresses they tend to produce on average lower quality papers. We also found that highly funded researchers who on average publish relatively low quality papers in high ranking journals, have lower average rate of publication. To the best of our knowledge this is the first paper that highlights such a relation. This relation may reflect an effect of high reputation enjoyed by some senior researchers which may enable them to get some of their works of lower quality published in high ranking journals. Given these results and considering the fact that funding is usually more biased towards senior researchers (Ebadi and Schiffauerova 2015b), we need to implicitly highlight the importance of more equal funding distribution among young and senior researchers, both with excellent scientific profiles. This brings us to the second suggestion for the funding allocations, which is to "give chance" to younger researchers with an evidence of a great potential as opposed to

keep funding senior researchers whose scientific performance is already “beyond the zenith”.

On a similar note, our results suggest that team size has positive impact on the research productivity. The researchers who get involved in larger scientific teams also produce higher quality papers. Hence, it is likely that scientists benefit from the collaboration to increase the quantity and quality of their scientific output. We also made an interesting observation as to the career age of the team members. As the career age of researchers increases, working in larger team reduces their productivity and the quality of their papers. Hence, from the results it seems that young researchers who work in large teams are more likely to produce high quality publications as opposed to the older team members. Therefore, we suggest that the partnership grants or team funding programs have special focus on funding young promising researchers who may play an instrumental role in developing teams which produce high quality scientific output.

We have also looked into the impact of the institution type of the funded researchers' affiliations on the quantity and quality of their scientific publications. We found that although the NSERC funded researchers who were affiliated with academic institutions were more productive in terms of the number of publications, the papers of the NSERC funded but industry-affiliated researchers were of higher quality (measured by the average number of citations). This may reflect the “publish or perish” environment existing in the academic world in which academic researchers are forced to publish as many papers as possible, while the quality of their works may be sometimes lacking. Industrial researchers funded from public financial sources, on the other hand, are not rushed into publishing papers in great quantities, but tend to focus on the quality of their papers which consequently then receive on average higher citations. Supporting grants which involve industrial researchers or include industrial researchers in the academic teams has thus become our fourth suggestion. The advantages of academia-industry collaboration have been long known to funding agencies, but this work has brought a direct evidence of the unique performance of publically funded industrial researchers and their impact on increasing the quality of the scientific output.

Finally, we compared the impact of different NSERC funding programs on scientific output of the funded researchers to find out which program yields the highest productivity. It was found that strategic programs which are of high priority and narrower scope are the most influencing funding programs based on the results in all the four estimated models. Hence, our fifth suggestion for boosting the scientific productivity is defining well targeted priority funding programs and/or allocating funding to researchers to work on high priority projects instead of continuously supporting broadly defined and less efficient granting programs such as discovery grants.

Limitations and future work

We were exposed to some limitations in this paper. First, we selected SCOPUS for gathering information about the NSERC funded researchers' articles. Since SCOPUS and other similar databases are English biased, hence, non-English articles are under-represented (Okubo 1997). Secondly, since SCOPUS data is less complete before 1996, we limited the time interval to 1996–2010 for our analysis. Another inevitable limitation related to the data was the spelling errors and missing values. Although SCOPUS is

confirmed in the literature to have a good coverage of articles, as a future work it would be recommended to focus on other similar databases to compare and confirm the results.

Different scientific disciplines follow different patterns in publishing articles, collaborating with other researchers, or even getting and allocating grants to the projects/researchers. Hence to better examine scientific productivity and efficiency, a future work direction could be assessing the impact of funding on the rate of publications for different scientific disciplines separately. In addition, other funding councils can be considered as the source of funding data. This kind of analyses, and comparing the efficiency of different funding organizations may help the decision makers to set the best funding allocation strategy. Meanwhile, fractional counting of publications can be also considered in future works.

References

- 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.
- Adler, R., Ewing, J., & Taylor, P. (2009). Citation statistics. *Statistical Science*, *24*(1), 1.
- Arora, A., & Gambardella, A. (1998). *The impact of NSF support for basic research in economics*. Economics Working Paper Archive at WUSTL.
- Beaudry, C., & Allaoui, S. (2012). Impact of public and private research funding on scientific production: The case of nanotechnology. *Research Policy*, *41*(9), 1589–1606.
- Beaudry, C., & Clerk-Lamallice, M. (2010). *Grants, contracts and networks: What influences biotechnology scientific production?* (pp. 16–18). London: Danish Research Unit for Industrial Dynamics (DRUID) Conference.
- Beaver, D. D. (2001). Reflections on scientific collaboration (and its study): Past, present, and future. *Scientometrics*, *52*(3), 365–377.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *The Journal of Machine Learning Research*, *3*, 993–1022.
- Boyack, K. W., & Börner, K. (2003). Indicator-assisted evaluation and funding of research: Visualizing the influence of grants on the number and citation counts of research papers. *Journal of the American Society for Information Science and Technology*, *54*(5), 447–461.
- Campbell, D., & Bertrand, F. (2009). *Bibliometrics as a performance measurement tool for the evaluation of research: The case of Canadian forest service*. Science-Metrix, 2009 Annual CES Conference.
- Campbell, D., Picard-Aitken, M., Côté, G., Caruso, J., Valentim, R., Edmonds, S., & Bastien, N. (2010). Bibliometrics as a performance measurement tool for research evaluation: The case of research funded by the national cancer institute of Canada. *American Journal of Evaluation*, *31*(1), 66–83.
- Carayol, N., & Matt, M. (2006). Individual and collective determinants of academic scientists' productivity. *Information Economics and Policy*, *18*(1), 55–72.
- Centra, J. A. (1983). Research productivity and teaching effectiveness. *Research in Higher Education*, *18*(4), 379–389.
- Cole, S. (1979). Age and scientific performance. *American Journal of Sociology*, *84*, 958–977.
- Coleman, J. S., & Lazarsfeld, P. F. (1981). *Longitudinal data analysis*. New York: Basic Books.
- Crespi, G. A., & Geuna, A. (2008). An empirical study of scientific production: A cross country analysis, 1981–2002. *Research Policy*, *37*(4), 565–579.
- De Solla Price, D. J., & Beaver, D. (1966). Collaboration in an invisible college. *American Psychologist*, *21*(11), 1011.
- Ebadi, A., & Schiffauerova, A. (2013). Impact of funding on scientific output and collaboration: A survey of literature. *Journal of Information and Knowledge Management*, *12*(04), 1350037.
- Ebadi, A., & Schiffauerova, A. (2015a). On the relation between the small world structure and scientific activities. *PLoS ONE*, *10*(3), e0121129.
- Ebadi, A., & Schiffauerova, A. (2015b). How to receive more funding for your research? get connected to the right people! Accepted in *PLoS ONE*.
- Ebadi, A., & Schiffauerova, A. (2015c). How to become an important player in scientific collaboration networks? *Journal of Informetrics*, *9*(4), 809–825.

- Falagas, M. E., Kouranos, V. D., Arencibia-Jorge, R., & Karageorgopoulos, D. E. (2008). Comparison of SCImago journal rank indicator with journal impact factor. *The FASEB Journal*, 22(8), 2623–2628.
- Gingras, Y. (1996). Bibliometric analysis of funded research. A feasibility study.
- Godin, B. (2003). *The impact of research grants on the productivity and quality of scientific research*. No. 2003. INRS Working Paper.
- Gulbrandsen, M., & Smeby, J. (2005). Industry funding and university professors' research performance. *Research Policy*, 34(6), 932–950.
- Hausman, J. A., Hall, B. H., & Griliches, Z. (1984). Econometric models for count data with an application to the patents-R&D relationship. *Econometrica*, 52(4), 909–938.
- Heinze, T., & Kuhlmann, S. (2008). Across institutional boundaries? Research collaboration in German public sector nanoscience. *Research Policy*, 37(5), 888–899.
- Huffman, W. E., & Evenson, R. E. (2005). *New econometric evidence on agricultural total factor productivity determinants: Impact of funding composition*. Iowa State University, Department of Economics, Working Paper, 3029.
- Jacob, B., & Lefgren, L. (2007). The impact of research grant funding on scientific productivity.
- Jacob, B. A., & Lefgren, L. (2011). The impact of research grant funding on scientific productivity. *Journal of Public Economics*, 95(9), 1168–1177.
- Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy*, 26(1), 1–18.
- King, J. (1987). A review of bibliometric and other science indicators and their role in research evaluation. *Journal of Information Science*, 13(5), 261–276.
- Kyvik, S., & Olsen, T. B. (2008). Does the aging of tenured academic staff affect the research performance of universities? *Scientometrics*, 76(3), 439–455.
- Larivière, V. (2012). On the shoulders of students? The contribution of PhD students to the advancement of knowledge. *Scientometrics*, 90(2), 463–481.
- Latour, B., & Woolgar, S. (1979). *Laboratory life: The social construction of scientific facts*. Princeton: Princeton University Press.
- Lawani, S. M. (1986). Some bibliometric correlates of quality in scientific research. *Scientometrics*, 9(1–2), 13–25.
- Lee, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science*, 35(5), 673–702.
- Lehman, H. C. (1953). Age and achievement.
- Lewis, G., & Dawson, G. (1998). The effect of funding on the outputs of biomedical research. *Scientometrics*, 41(1), 17–27.
- Leydesdorff, L., & Wagner, C. (2009). Macro-level indicators of the relations between research funding and research output. *Journal of Informetrics*, 3(4), 353–362.
- Martín-Sempere, M. J., Rey-Rocha, J., & Garzón-García, B. (2002). The effect of team consolidation on research collaboration and performance of scientists. Case study of Spanish university researchers in geology. *Scientometrics*, 55(3), 377–394.
- McAllister, P. R., & Narin, F. (1983). Characterization of the research papers of US medical schools. *Journal of the American Society for Information Science*, 34(2), 123–131.
- Melin, G. (2000). Pragmatism and self-organization: Research collaboration on the individual level. *Research Policy*, 29(1), 31–40.
- Merton, R. K. (1968). The Matthew effect in science. *Science*, 159(3810), 56–63.
- Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. Chicago: University of Chicago press.
- Moed, H. F. (2006). *Citation analysis in research evaluation*. Dordrecht: Springer.
- Niosi, J. (2000). *Canada's national system of innovation*. Montreal: McGill-Queen's Press-MQUP.
- NSERC. (2013). Report on plans and priorities, 2013–2014. Retrieved from http://www.nserc-crsng.gc.ca/NSERC-CRSNG/Reports-Rapports/RPP-PPR/2013-2014/index_eng.asp.
- Okubo, Y. (1997). *Bibliometric indicators and analysis of research systems: Methods and examples*. No. 1997/1. Paris: OECD Publishing.
- Payne, A. A., & Siow, A. (2003). Does federal research funding increase university research output? *Advances in Economic Analysis and Policy*, 3(1), 1–24.
- Peritz, B. C. (1990). The citation impact of funded and unfunded research in economics. *Scientometrics*, 19(3–4), 199–206.
- Plume, A., & van Wiejen, D. (2014). Publish or perish? The rise of the fractional author. *Trends Journal of Sciences Research*.
- Salter, A. J., & Martin, B. R. (2001). The economic benefits of publicly funded basic research: A critical review. *Research Policy*, 30(3), 509–532.

- Sandström, U. (2009). Research quality and diversity of funding: A model for relating research money to output of research. *Scientometrics*, *79*(2), 341–349.
- Schilling, M. A., & Phelps, C. C. (2007). Interfirm collaboration networks: The impact of large-scale network structure on firm innovation. *Management Science*, *53*(7), 1113–1126.
- Shapira, P., & Wang, J. (2010). Follow the money. *Nature*, *468*(7324), 627–628.
- Tahmooresnejad, L., Beaudry, C., & Schifffauerova, A. (2015). The role of public funding in nanotechnology scientific production: Where Canada stands in comparison to the United States. *Scientometrics*, *102*, 753–787.
- Thorsteinsdóttir, O. H. (2000). External research collaboration in two small science systems. *Scientometrics*, *49*(1), 145–160.
- Wray, K. B. (2003). Is science really a young man's game? *Social Studies of Science*, *33*(1), 137–149.
- Wray, K. B. (2004). An examination of the contributions of young scientists in new fields. *Scientometrics*, *61*(1), 117–128.
- Zucker, L. G., Darby, M. R., Furner, J., Liu, R. C., & Ma, H. (2007). Minerva unbound: Knowledge stocks, knowledge flows and new knowledge production. *Research Policy*, *36*(6), 850–863.