Effects of gender, age, research experience and leading role on academic productivity of Vietnamese researchers in the social sciences and humanities: exploring a 2008-2017 Scopus dataset

Quan-Hoang Vuong, Tung Manh Ho, Nancy K. Napier, Thu-Trang Vuong and Hiep Hung Pham

Academic productivity has been studied by scholars all around the world for many years. However, in Vietnam, this topic has thus far been under-researched. This research therefore aims to better understand the correlations between gender, age, research experience, the leading role of corresponding authors and the total numbers of their publications, in the specific realm of social sciences. The study employs a Scopus dataset during 2008-2017, containing publication profiles of 410 Vietnamese researchers. Contrary to a range of previous studies, the results indicate that among accomplished social scientists, males have not been more productive or proficient than females with respect to academic publications (β_{male} = -0.179, \ p = 0.60). On the other hand, the proficient skills and broad vision of corresponding authors have proved to exert a rather strong influence on their sheer number of papers (ρ = 0.832). Older age and longer research time also contribute to more success in their academic careers (β_{age\geq50} = 0.950, \ p<0.05; \ β_{research time}=0.042, \ p<0.05).

Keywords: Social science publications, corresponding author, career age, age, gender
JEL Classifications: I23, O38, P36

CEB Working Paper N° 17/014
April 2017
Effects of gender, age, research experience and leading role on academic productivity of Vietnamese researchers in the social sciences and humanities: exploring a 2008-2017 Scopus dataset

Quan-Hoang Vuong
Université Libre de Bruxelles, Centre Emile Bernheim
50 Ave. Franklin D. Roosevelt, Bruxelles B-1050, Belgium
Email: qvuong@ulb.ac.be

Tung Manh Ho
Vietnamese Academy of Social Sciences, Institute of Philosophy
59 Lang Ha, Ba Dinh, Hanoi, Vietnam
Email: tungmanhho@gmail.com

Nancy K. Napier
Boise State University, College of Business and Economics
1910 University Drive, Boise, ID, USA
Email: nnapier@boisestate.edu

Thu-Trang Vuong
Campus Européen de Dijon, Sciences Po Paris
14 Avenue Victor Hugo, 21000 Dijon, France
Email: thutrang.vuong@sciencespo.fr

Hiep Hung Pham
Department of International Business Administration, Chinese Culture University
55 Hwa Kang Road, Yangming Shan 11114, Taipei, Taiwan R.O.C.
Email: phamhunghiep@gmail.com

This version: Manuscript v5
Date: April 16, 2017
Abstract

Academic productivity has been studied by scholars all around the world for many years. However, in Vietnam, this topic has thus far been under-researched. This research therefore aims to better understand the correlations between gender, age, research experience, the leading role of corresponding authors and the total numbers of their publications, in the specific realm of social sciences. The study employs a Scopus dataset during 2008-2017, containing publication profiles of 410 Vietnamese researchers. Contrary to a range of previous studies, the results indicate that among accomplished social scientists, males have not been more productive or proficient than females with respect to academic publications ($\beta_{\text{male}} = -0.179$, $p = 0.60$). On the other hand, the proficient skills and broad vision of corresponding authors have proved to exert a rather strong influence on their sheer number of papers ($\rho = 0.832$). Older age and longer research time also contribute to more success in their academic careers ($\beta_{\text{age} \geq 50} = 0.950$, $p < 0.05$; $\beta_{\text{research time}} = 0.042$, $p < 0.05$).

Keywords: Social science publications, corresponding author, career age, age, gender

JEL codes: I23, O38, P36

Introduction

Research studies on different factors affecting scientific productivity have been carried out for many years and continue to attract academic attention. Gender is one of the most studied among potentially influential factors on productivity. Generally, most previous results showed that male scientists outperformed female [1-9]. However, findings varied by fields [10]. For instance, in natural sciences, namely chemistry, Long (1992) unveiled that the quantitative difference of publications and citations between sexes (men outperforming women) increased during the first decade of the career, but were reversed later [1]. Also in chemistry, Reskin (1978) found out that though men tended to publish more papers than women, the difference remained small [3]. In physics, Mairese & Pezzoni (2015) indicated that the number of publications by female authors was about one third that of male authors in average, which is a considerable difference [5]. Conversely, in social sciences, several recent studies suggested that while men had been outperforming women in terms of publications and citations in previous generations, there seems to be a change in the younger generation of researchers, as gendered performance differences began to disappear. If performance differences still exist at all in our case, it would be young female researchers who outperformed young male researchers, especially in developed societies [6,7]. The reasons for this difference were raised over various aspects, but they can
be distilled into three main reasons: publication issues, marriage, and specialization. Women’s lower productivity in terms of number of publications was largely due to their frequent non-publishing works [1,5]. In many cases, child-rearing was also an obstacle impeding them from their research career [2-4]. Besides, women specialized less and changed jobs more often than men, thereby losing out on increasing their productivity in a specific field [3]. However, when other factors were taken into account, it turned out that female researchers appeared to be as productive as their male colleagues, even more productive in some cases [5]. In effect, the average citations per publication of female authors were higher than those of male authors, by 1.5 times [1].

Beside gender, age and research experience (sometimes called ‘career age’ in this study) are also two potentially influential elements affecting differences in academic/scientific productivity. Age is associated with both the quantity and quality of scientific output [12-18]. Young female researchers were, on an average, almost half a year younger than their male counterparts [6]. In a study in six different fields, Cole (1979) found that scientists’ reached their highest productivity twice during their life: one between their thirties and forties, and one at 50 [12].

Another indicator often mentioned by authors on this subject was career age, which was defined as the number of years since their first publication. Allison & Stewart (1974) and Jones et al. (1991) proved that there was a significant positive association of career age and life-time research productivity [19-20]. Also, the sheer number of publications peaked at two different points: the first at 5-10 years since the beginning of the scientist’s career, and the second when they were close to retirement [16,21]. There was a decline either mid-career or during the latter part of their careers; however, it could be observed that productivity was higher for those who were in their retirement years but had yet to leave academic employment [21]. Furthermore, age at first publication and the number of publications before Ph.D. also exerted important effects upon both decade and lifetime productivity [3,22-23]. Notably, male scientist’s productivity would be significantly higher than female in the fifth year after Ph.D [3]. Yet, the differences could also differ across various sectors [24]. In the social sciences, productivity remained more or less at the same level in all age groups [14]. This phenomenon could be attributed to the fact that knowledge production in these fields occurred at a slower pace; thus, researchers could be productive throughout their careers [14]. On the other side, Wagner-Döbler (1995) believed that if the number of scientists with different ages of career were standardized, it could be seen immediately that senior scientists contribute to the same extent as younger scientists, not only with regard to the frequency of publication but also with regard to especially influential contributions. He suggested that studies on scientific productivity should distinguish between a psychological or anthropological perspective and an account dealing with the structure and intensity of participation in a scientific discipline in the course of its development. The first aspect could be illuminated, for
example, by measuring the speed of publication of scientists, the second by the structure of participation or output in a period of time [25].

In the world, Northern America, Oceania and Western Europe were regions with the highest numbers of paper produced per citizen. Of which, Western Europe had the largest total number of publications, followed by Northern America with a contribution of 38.50% and 31.39% in the world’s total publications, respectively. In Asia, research activities remained at a humble figure, 15.16% [26]. In Vietnam, the rate of growth in scientific output was 17% per annum, and international collaboration was about 77% of the total output, with Japan and America being the biggest collaborating countries [27,28]. Remarkably, three-quarters of the growth was associated with international collaborations rather than purely domestic production, and internationally coauthored papers received twice the average citation as domestic papers [28]. In Southeast Asia, Vietnamese scientific publications only accounted for 0.6% in the total, ranked the 4th among the countries [29]. In the group of six favored emerging markets countries – CIVETS, Vietnam also had the fourth position in total publications, after by Turkey, South Africa and Egypt; and third position in numbers of citations per paper, after South Africa and Indonesia [30]. However, the data also indicated that Vietnam is in the growth phase of building up research capacity [28].

In this study, the factors of age, gender, and career age are considered in relation to the sheer number of publications of Vietnamese social scientists. Moreover, the study also provides some insights on the importance of the researcher’s role as a corresponding author in the works. A number of concluding remarks are drawn in hopes of inspiring and contributing to upcoming studies.

**Materials and Methods**

Data on the number of international journal articles published by Vietnamese social science scholars is collected from online databases, conducted by Vuong & Associates team from January to April, 2017. The survey was performed under the license of V&A/03/2017 (March 15, 2017).

Selected subjects were those who have Vietnamese nationality, have at least 01 publication based on Vietnamese circumstances, or use data primarily concerning Vietnam for the research. After the collection of data, a valuable dataset of 410 authors was gathered along with their number of published articles as indexed in Scopus in the 2008-2017 period.

From the full dataset, information is extracted to serve this study, focusing on details such as whether the scholar is a sole author, leading author, or co-author... For each article published as a sole author, the researcher gain an 1 in “au.solo” (and if the researcher has no publication as a single author, “au.solo” would be 0). Similarly, an article published as the corresponding author (or having his/her name as the first position in the author group) will get the researcher 1 unit of count in “au.key”; and,
in the same logic, a co-author who didn’t lead will gain 1 in “au.coll” for every co-publication. In addition, the number of all collaborating authors and the number of Vietnamese authors of which, the number of years in research, field of study, age, gender, etc are also included in the data.

The raw data were entered in MS Excel, then processed and converted to CSV format. The CSV file was next analyzed in the statistical software R (3.3.1), which reported statistical results as well as relevant graphs.

This paper is set to resolve two main problems: (1) the impact of age and the number of articles in which the authors taking the leading role on the total number of publications; and (2) the influence of the number of articles in which the authors taking the leading role, career age, and gender on the total number of publications.

To answer these two questions, the method employed is multi-variable linear OLS regression with the general model as follows:

\[ Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k \]

The condition is that \( k \) independent variables \( X_i \) must have the same sample size \( n \) as dependent variable \( Y \). \( Y \) is a numerical variable, while \( X_i \) can be numerical or categorical [31]. After being processed in R, the results provide the values of \( \beta_i \), which represent the linear impact of \( X_i \) on \( Y \) (with \( Y \) being the total number of publications in this research). The statistical significance of predictor variables in the model is determined based on \( z \)-value and \( p \)-value; with \( p<0.05 \) being the conventional level of statistical significance required for a positive result.

Results

Descriptive statistics

The data collected provide personal information such as age and gender, as well as professional details: total research time, the role of the scientist in a scientific work, and the total number of the publications since their graduation of master courses.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Min</th>
<th>Min</th>
<th>Max</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>19</td>
<td>42.10</td>
<td>72</td>
<td>9.15</td>
</tr>
<tr>
<td>Career age</td>
<td>2</td>
<td>15.05</td>
<td>64</td>
<td>8.76</td>
</tr>
<tr>
<td>Total number of publications</td>
<td>1</td>
<td>3.6</td>
<td>63</td>
<td>5.89</td>
</tr>
<tr>
<td>Number of papers in leading role</td>
<td>0</td>
<td>1.77</td>
<td>60</td>
<td>4.24</td>
</tr>
</tbody>
</table>
Of 410 authors observed, 62.2% are male and 37.8% female with the age ranges from 19 to 72 years old. The dominant group of age is 35-45, accounting for 57.56% (Fig.1a). The largest number of publications is 63 (Fig.1b). The average number of publications of male authors is a little higher than that of female: 3.84 and 3.21 papers, respectively (Fig.1c).
More than half of the scientists (over 58%) were from Northern provinces of Vietnam. Some overseas Vietnamese authors were also on the survey (nearly 17%) and the scientific output of this group displayed the strongest dispersion (Fig. 1d). The shortest time in research was 2 years with 1 publication, while the longest was 64 years. 339 out of 410 authors have 20 years of experience or less in their career. It is worth noting that, paradoxically, those with an experience of more than 40 years often had a very low output (Fig. 1e).

In terms of the key role of authors, the statistics showed that almost all authors held key position in 10 papers or less; only 2% of them played a critical role in more than 10 works. To ensure equity between authors who wrote multiple articles but not hold the key position in any article and authors who published solo or who were leading in numerous publications, the data set is further calculated by the factor of conversion. This helps to find the differences between the authors, who were pioneering the development of academia in Vietnam, and the authors who merely played a collaborating role in scientific works.

**Estimation results**

**Result 1: Effects of age group and the leading role on authors’ numbers of publications**

The sheer number of publications (“ttlitems”) is examined in its relation to age group (“age_gr”) and the number of papers in which the scientists hold the key role (“au.key”). While
“au.key” is a continuous variable, “age_gr” is a categorical one, divided into 4 groups: less than 30 years old (“less30”), from 30 to less 40 (“b3040”), from 40 to less 50 (“b4050”), and 50 and older (“g50”).

Results obtained after processing the data in R are displayed in the following table:

<table>
<thead>
<tr>
<th>“ttlitems”</th>
<th>Intercept</th>
<th>“au.key”</th>
<th>“age_gr”</th>
<th>“age_gr”</th>
<th>“age_gr”</th>
<th>“age_gr”</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta_0$</td>
<td>$\beta_1$</td>
<td>$\beta_2$</td>
<td>$\beta_3$</td>
<td>$\beta_4$</td>
<td></td>
</tr>
<tr>
<td>“less30”</td>
<td>1.115***</td>
<td>1.145***</td>
<td>-0.390</td>
<td>0.682c</td>
<td>0.950*</td>
<td></td>
</tr>
<tr>
<td>[4.132]</td>
<td>[29.898]</td>
<td>[-0.399]</td>
<td>[1.843]</td>
<td>[2.128]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘c’; z-value in square brackets; baseline category for: “age_gr” = “b3040”. Residual standard error: 3.258 on 405 degrees of freedom. F-stat = 232.7 on 4 and 405 df, p-value: < 2.2×10^-16. Adj. $R^2 = 0.6938$.

From Table 2, it can be seen that most of the coefficients are statistically significant with $p<0.1$. Therefore, the relationships between these above variables are affirmed. The regression equation is provided in (Eq.1) as follows:

$$ttlitems = 1.115 + 1.145\times au.key - 0.390\times age_grless30 + 0.682\times age_grb4050 + 0.950\times age_grg50 \quad (Eq.1)$$

From (Eq.1) the total number of publications of an author aged 40 and being key author in 2 papers, for example, would be calculated as follows:

$$1.115 + 1.145\times2 - 0.390\times0 + 0.682\times1 + 0.950\times0 = 4.087.$$ 

**Result 2: Effects of authors’ leading role in publications, career age, and gender on numbers of publications:**

In this model, the factors taken into account are “au.key”, research time (“restime”) and gender (“sex”). Through a similar procedure, the results are obtained and reported in Table 3.

<table>
<thead>
<tr>
<th>“ttlitems”</th>
<th>Intercept</th>
<th>“au.key”</th>
<th>“restime”</th>
<th>“sex”</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>$\beta_1$</td>
<td>$\beta_2$</td>
<td>$\beta_3$</td>
<td>$\beta_4$</td>
</tr>
<tr>
<td>“less30”</td>
<td>1.115***</td>
<td>1.145***</td>
<td>-0.390</td>
<td>0.682c</td>
</tr>
<tr>
<td>[4.132]</td>
<td>[29.898]</td>
<td>[-0.399]</td>
<td>[1.843]</td>
<td>[2.128]</td>
</tr>
</tbody>
</table>

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘c’; z-value in square brackets; baseline category for: “restime” = “b3040”. Residual standard error: 3.258 on 405 degrees of freedom. F-stat = 232.7 on 4 and 405 df, p-value: < 2.2×10^-16. Adj. $R^2 = 0.6938$.
Based on Table 3, it can be observed that with \( p < 0.05 \) only “au.key” and “restime” are statistically significant, meanwhile “sex” does not show any influence on “ttlitems”. The regression equation is:

\[
\text{ttlitems} = 1.028 + 1.155 \times \text{au.key} + 0.042 \times \text{restime} - 0.179 \times \text{sexM} \quad (\text{Eq.2})
\]

Applying (Eq.2) is applied to calculate the number of publications of a female scientist having 15 years in research career, the result is showed as: \( 1.028 + 1.155 \times 2 + 0.042 \times 15 - 0.179 \times 0 = 3.968 \).

**Discussion**

Several preliminary findings on the productivity of social scientists in Vietnam can be inferred from these figures and estimation results.

The first model quantifies the impact of the number of papers in which the scientists play the most crucial role on the total numbers of their publications. Adjusted \( R^2 \) being 0.6938 suggests that the influential factors in (Eq.1) explain more than 69% for the change of these numbers. Moreover, F test is also conducted with null-hypothesis \( H_0 \) stating that all coefficients are simultaneously equal to 0. The results of the test show that \( df_1=4, df_2=405 \), and \( p=2.2 \times 10^{-6} \), therefore \( H_0 \) is rejected. In other words, the coefficients in the model are not simultaneously equal to 0 and the model is valid.

Observing the estimate coefficients in (Eq.1), the coefficient of “au.key” has a positive value and the magnitude is 1.145. The numerical value implies a proportional relationship between the number of publications in which the author led and the total number of their publications. In detail, when other variables are being controlled, each increase in the number of papers in which the author held the critical role boosts total number of publications to rise by 1.145.
Regarding the influence of age, Table 2 presents that the absolute values of the estimated coefficients increase along with the increase of age, with $\beta_2 = -0.390$ for the group of less than 30 years old and $\beta_4 = 0.950$ for the ones aged 50 or older. Therefore, the older the researcher is, the larger the impact of their age on the number of publications. Taking into account the signs of these coefficients, it can be remarked that older age likely result in more publishing activity.

The effect of career age and gender is displayed in model 2 (Table 3). The role of a scientist in a research work is once again entered into the model in order to emphasize the importance of this factor. Concerning $R^2$, when the variable “age_gr” is replaced by “restime” and “sex”, the value of $R^2$-adjusted is 69.35, a negligible decrease compared to its value in (Eq.1). The negligible reduction of $R^2$-adjusted also implies that additional independent variables have an insignificant effect on the dependent variable.

Furthermore, the dataset for this study is tested for the Pearson correlations among pairs of variables. For example, the test with the null hypothesis (H0) that the correlation between “au.key” and “ttlitesm” is 0. The test is performed in R, with the result showing a correlation coefficient of 0.832, with $p = 2.2 \times 10^{-16}$. This means that H0 is rejected [32], and there exists a strong relationship between the number of papers in which the scientist is the corresponding author and their total number of publications.

Similarly, other correlations are computed following [32], namely $\rho(ttlitesm,restime) = 0.078$ ($p = 0.114$), $\rho(ttlitesm,age) = 0.088$ ($p = 0.076$), as shown in (Fig.2).
The estimate coefficients of “au.key” and “restime” in (Eq.2) have positive values ($\beta_1=1.155$, $\beta_2=0.042$), suggesting that longer research time and more publications in which the author hold the key position could result in a larger volume of lifetime publication. However, the magnitude of “restime” is rather small, which shows that the effect of career age is quite limited.

In addition, the estimate coefficient of “sex” in Table 3 is $\beta_3=-0.179$, $p=0.6$. It can be remarked that gender is not associated with scientific productivity in terms of number of publications.

**Conclusion**

Based on the analysis, it can be concluded that the sheer numbers of publications of Vietnamese social scientists were influenced by research time, age and the number of papers in which they were corresponding authors. Older age and longer research time are correlated with more scientific output. Scientists with between 15 to 25 years of research experience published the largest numbers of papers. That being said, there are a number of authors aged over 40 having much fewer publications compared to younger researchers. The reason to this might be that the previous generation of research scientists in Vietnam had less access to the global scientific community, due to a lack of communication prior to the arrival of Internet or even the socio-economic reforms (*Doi Moi*) in the 1980s for the oldest authors. A large portion of their contributions ought to have been published without being indexed by Scopus.
The author’s initiative and active attitude also contribute to their success in scientific career, with the correlation between the number of publication in which the researcher was a key author and the total number of publications of said researcher being 0.832. Corresponding authors are those who come up with the ideas, take control of the project and are responsible for the contents and the quality of the paper. Therefore, they are required to have not only proficient skills in their specialized field but also a broad vision on other social problems, and a certain degree of leadership if they work in a group. These qualities either in turn refine their productive capacities, or are generally related to the researcher’s own diligence. As a result, their works tend to be qualitatively and quantitatively ampler.

In contrast to the remark of male scientific outperformance in earlier studies [1-9], the findings in this study indicate that there is no productive difference between two genders in Vietnam. The empirical numbers of publications of male and female scientists are 3.835 and 3.213, respectively. The result essentially means that marital and parental responsibilities of women no longer hinder their scientific productivity, at least in modern Vietnam. While this may sound optimistic, especially in the context of global efforts towards gender equality, it must still be noted that the number of female scientists remains substantially smaller than that of male scientists. Higher productivity does not mean that women are now free to devote themselves to science; it only means that women who remained in the field are being as active and productive as men in their scientific contribution.

Acknowledgements

The authors would like to thank V&A research team for efficient research assistance during the process of collecting raw data and preparing the clean data set for this analysis, specifically Dam Thu Ha, Do Thu Hang, Nguyen Thi Phuong, and Do Phuong Ngoc.

References

3. Reskin BF. Scientific productivity, sex, and location in the institution of science. American Journal of Sociology 1978;83(5):1235-43. DOI: https://doi.org/10.1086/226681


7. van Arensbergen P, van der Weijden I, van den Basilar P. Gender differences in scientific productivity: a persisting phenomenon?. *Scientometrics* 2012;93(3):857-68. DOI: [https://doi.org/10.1007/s11192-012-0712-y](https://doi.org/10.1007/s11192-012-0712-y)


25. Wagner-Döbler R. Where has the cumulative advantage gone? Some observations about the frequency distribution of scientific productivity, of duration of scientific participation, and of speed of publication. *Scientometrics* 1995;32(2):123-32. DOI: https://doi.org/10.1007/BF02016890


