Comparative views on research productivity differences between major social science fields in Vietnam: Structured data and Bayesian analysis, 2008-2018

Quan-Hoang Vuong [a]

Viet-Phuong La [a]

Thu-Trang Vuong [b]

Manh-Tung Ho [a,c]

Minh-Hoang Nguyen [a,c]

Manh-Toan Ho [a,d]

[a] SDAG Lab, Centre for Interdisciplinary Social Research, Phenikaa University, Yen Nghia, Ha Dong District, Hanoi, 100803, Viet Nam

[b] Sciences Po Paris, École doctorale, 27 Rue Saint-Guillaume, 75007 Paris, France

[c] Ritsumeikan Asia Pacific University, Graduate School of Asia Pacific Studies, Oita Prefecture, 874-8577, Japan

[d] A.I. for Social Data Lab, Vuong & Associates, 3/161 Thinh Quang, Dong Da District, Hanoi, 100000, Viet Nam

Abstract: Since Circular 34 from the Ministry of Science and Technology of Vietnam required the head of the national project to have project results published in ISI/Scopus journals in 2014, the field of economics has been dominating the number of nationally-funded projects in social sciences and humanities. However, there has been no scientometric study that focuses on the difference in productivity among fields in Vietnam. Thus, harnessing the power of the SSHPA database (http://sshpa.com/), a comprehensive dataset of 1,564 Vietnamese authors (854 males, 705 females) with 2,410 publications in the 2008 – 2018 period was extracted and analyzed. Various factors were considered including age, gender, new authors, leading authors, co-authorship, and Impact Factor. The findings suggest a high level of contribution from authors at the age of 40 – 44 in economics (858 publications) in a 12-years period, which is equivalent to the social medicine total output, and two times more than the total output of the education. Moreover, the presence and reinforcement of male researchers are still dominating in Economics and other fields, with the only exception of education. Despite the rapid rise in the number of Vietnamese lead authors, gender disparity among disciplines is an issue. Contrary to the strong international collaboration-oriented tendency in social medicine, economics, and other fields, educational authors are not open to international collaborating. Finally, most of the publications in economics belong to the group with JIF from 0 to 2, in contrast with the high number of social medicine publications with JIF from 2 to 5, which suggesting the field of economics is fulfilling the quantity, but still, need more quality publications.

Keywords: Social sciences and humanities; economics; social medicine; education; scientific productivity; SSHPA database.

Working Paper ISR-SDAG #20-01
Version 5; January 1, 2020

I. Introduction

From 2011 to 2017, Vietnam National Foundation for Science and Technology Development (NAFOSTED)—the equivalent of United States' National Science Foundation—funded 384 social sciences and humanities research projects, with an average budget of 745 million VND (~223,000 USD) per year, according to the data provided by NAFOSTED (Nafosted 2018). Within this period, roughly one out of four funded projects belonged to the discipline of economics (95/384). In 2014, Circular 37/2014/TT-BKHCN was introduced, which sets international publications in ISI/Scopus journals as the standard for the head of the project and the research result. Since then, the number of funded projects drastically decreased, from an average of 71 projects per year in 2011-2014 (of the total 283 projects) to only 23 projects per year in the 2015-2017 period. The number of funded projects in economics was 47 out of 68 projects in the 2015-2017 period (Nafosted 2018).

Even though there are some inconsistencies in these published figures from NAFOSTED, it is hard to deny that economics has been the pioneering figure in the new era of Vietnamese academia, in which scientific qualification must now adhere to international standards based on renowned databases such as ISI Web of Science (ISI WOS) and Scopus. However, there has been no scientometric study in Vietnam to compare scientific performance between fields, not even among the more prioritized natural sciences and STEM fields. Based on an original database on the researcher's productivity in social sciences and humanities (SSH) in Vietnam, this study aims to illustrate how economics is the leading discipline among Vietnam SSH.

II. Literature Review

Scientometric analyses focusing on Vietnam only stretched back less than ten years. Hien (2010) examined the total number of publications in international peer-reviewed journals per one million people, mean citation count, and the role of domestic researchers in peer-reviewed publications to compare the performance of 11 East and Southeast Asian countries. According to the study, Vietnam was one of the countries with low performance, with a high level of dependence on international authors, low output, and insignificant institutions. A year later, employing the Web of Science database on the period of 1991-2010, Nguyen and Pham (2011) suggested similar conclusions about Vietnam's scientific performance, with Vietnamese authors accounting for only 6% of the total output of Southeast Asian region. Another more recent study using Scopus data found the same results (Manh 2015). Notable, international collaboration accounted for 77% of the country's output. This share was reaffirmed once more using ISI WOS data in 2017 (Nguyen et al. 2017). Moreover, 90% of the Vietnamese researchers published articles as co-authors, and collaborated at least 13 times on average, mostly in a non-leading role. On a brighter note, the number of ISI-indexed publications from Vietnam has been rising by five times since 2009 (Adams et al. 2019); and international collaboration has helped to increase the quality and reputation of Vietnam science (Manh 2015; Nguyen et al. 2017; Q.-H. Vuong et al. 2018b).

Explorations of a dataset of 412 Vietnamese SSH researchers' scientific output in the 2008-2017 period have also provided significant preliminary results. Network analysis of the dataset suggested signs of low sustainability in the Vietnam SSH community, such as co-authorship network lacks information distribution, or a high level of reliance on a few highly connected members in the networks (T. M. Ho et al. 2017; T. Ho et al. 2017). Other studies examined the roles of collaboration, gender, age, regions, and first-authorship to the

productivity of Vietnamese SSH researchers. Results suggested no difference in productivity was associated with gender; however, first-authorship and seniority (age 40 – 50) appeared to have crucial contributions (Vuong et al. 2017; T.-T. Vuong et al. 2018). Finally, using the ordinary least squares method to analyze the institution aspect of the dataset, Q.-H. Vuong et al. (2018b) found that authors who are working at research institutions had much lower scientific output than authors who are affiliated with universities.

The difference between university-researcher and institution-researcher in Vietnam is striking because the investment from the government to higher education is relatively low; quality assurance is not an independent agency but an integral part of higher education; and the leadership and governance of higher education is still struggling between being controlled by the government and being fully autonomous (Salmi and Pham 2019). Nonetheless, in-depth interviews of 20 senior researchers in Vietnam SSH suggested that they had pressures and incentives to publish (Pham and Hayden 2019). However, there are struggles to publish internationally, while domestic scientific journals are usually low quality and lack credibility (Tran et al. 2019). Thus, while the scientific output from Vietnam is rising (Adams et al. 2019; T.-T. Nguyen et al. 2019) with the rise of specific fields in social sciences and humanities such as economics, or education (Vu et al. 2019; Le et al. 2019), there are specific demands to raise the quality of the environment (Salmi and Pham 2019; Phuong et al. 2015), as well as the system that will elevate it such as academic publishing (Vuong 2019b).

Early studies of Vietnamese scientific output had made use of the ISI WOS and Scopus data to provide a thorough overview. However, they did not explore specific fields and how the scientific output of each field contributes to the development of science in Vietnam. Meanwhile, even though the custom dataset of 412 Vietnamese social sciences and humanities researchers' scientific output can generate more profound results, its preliminary contributions still focused on the current state of Vietnam social sciences and humanities. Thus, this descriptive analysis, employing an expanded version of the large dataset, sets out to illustrate the output of the three main fields (economics, education, social medicine) and others, and how they shaped the past, present, and future of Vietnam SSH.

III. Methodology

A comprehensive dataset of scientific productivity of Vietnamese SSH researchers from 2008 – 2018 was extracted from the Social Sciences and Humanities Peer Awards (SSHPA) database, a homemade semi-automatic database that was built to record scientific productivity of Vietnamese SSH researchers. Details of the design logic, the architecture of the SSHPA database was thoroughly explained in (Q.-H. Vuong et al. 2018a). Eventually, the dataset (Publicly available on GitHub: https://github.com/sshpa/bayesvl/tree/master/LectureNotes/6.SSHPA/Data) contains records of 3,238 authors, in which 1,564 are Vietnamese (854 males, 705 females); 2,410 articles that were published in 1,171 journals (As of September 9, 2019, 23:43:01.040). Using descriptive approach, the productivity of Vietnamese SSH researchers in four main fields: economics (econ), education (edu), social medicine (med), and others (others) were analyzed based on various characteristics such as age, gender, new authors, leading authors, co-authorship, and Impact Factor. Moreover, the Bayesian approach was employed for data analysis for the data about age and gender.

Bayesian analysis was performed with the bayesvl package in R (La and Vuong 2019). The bayesvl package and R statistical software had been chosen for their potent capacity for generating graphics, diagnosing, and presenting research results from simulated data using Markov Chain Monte Carlo (MCMC) method. Moreover, the application of Bayesian statistics was also aimed at improving the research process and solving the problems posed by frequentist statistics, such as the plausibility of results, the reproducibility crisis, and the controversy related to interpreting the "p-value" (Vuong 2018, 2017).

The data analysis aims to answer the following questions:

Table 1. Research questions

Characteristics of the data	Questions
Age	• Which age group participated in most publications? • At which age do the authors have the highest productivity? • Is there any difference in the age of the author in each field?
Gender	• Is there any gender difference between the number of publications and researchers in each field? • What is the difference between the age of male and female researchers?
New Authors	• What are the variations in the number and the growth rate of new authors across research fields from 2010 to 2018? • What is the average age of new researchers in each field? • Is there any difference in the gender of new authors in each field?
Lead Authors	• Does the number of lead authors grow exponentially, and is there any difference among the fields? • What are the differences based on age, the gender of the lead authors, in comparison with other authors?
Co-authorship	• How is the number of articles distributed according to the number of co-authors among disciplines? • What is the difference in terms of the number of international collaborations among disciplines?
Journal Impact Factor	• How is the number of articles distributed by impact factor groups among disciplines? • How is the number of articles led by Vietnamese researchers distributed by impact factor groups among disciplines? • Does the impact factor among disciplines grow over the year? • Is the impact factor influenced by the age of Vietnamese lead authors among disciplines?

IV. Results

1. Age

This section aims to answer the following questions:

- Which age group participated in most publications?
- At which age do the authors have the highest productivity?
- Is there any difference in the age of the author in each field?

It should be noted that there are many authors in a single paper; thus, the data were based on both the number of authors and the number of papers in each age group as well as the average number of papers of an author in each age group.

1.1. Descriptive statistics

In Figure 1 and 2, the number of articles and authors in social medicine are equally distributed in each age group, while in other fields, these numbers are mostly focused on the group of researchers who are 30 to 45 years old. In economics, the age group of 35 – 39 and 40 – 44 occupied 48% of the total output of the field; including the age group of 30 – 34, the figure rose up to 65%. Consequently, the majority of authors were distributed in these age groups (444 authors), with 35 – 39 being the modal category. The high number of authors in this age group is also consistent with education and others.
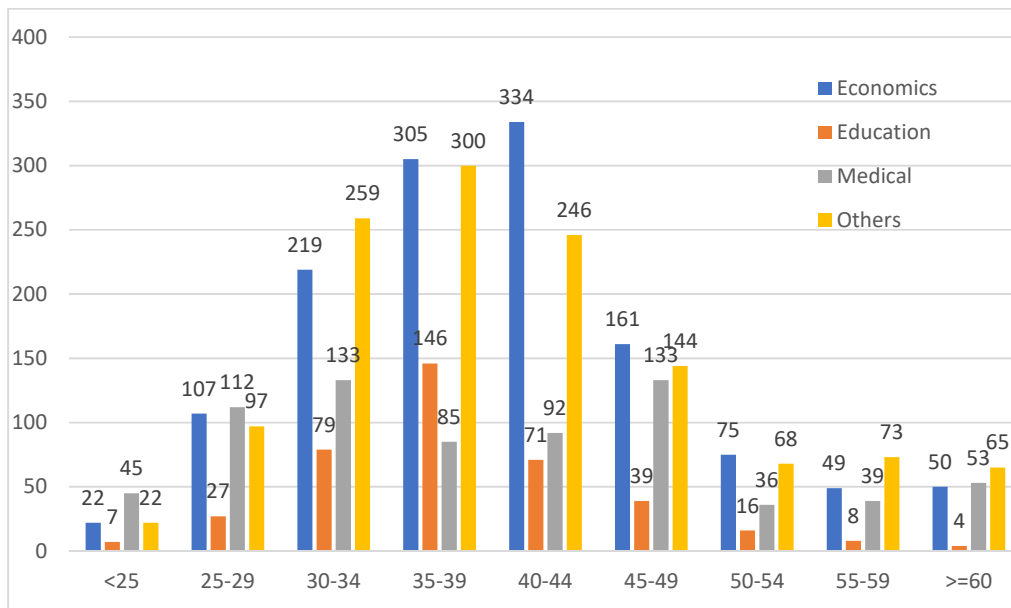
Figure 1. Number of articles in each age group

However, while the high number of authors in the age group of 35 – 39 in education and others is aligned with a high number of articles, the age group that produced the most papers in economics is 40 – 44.
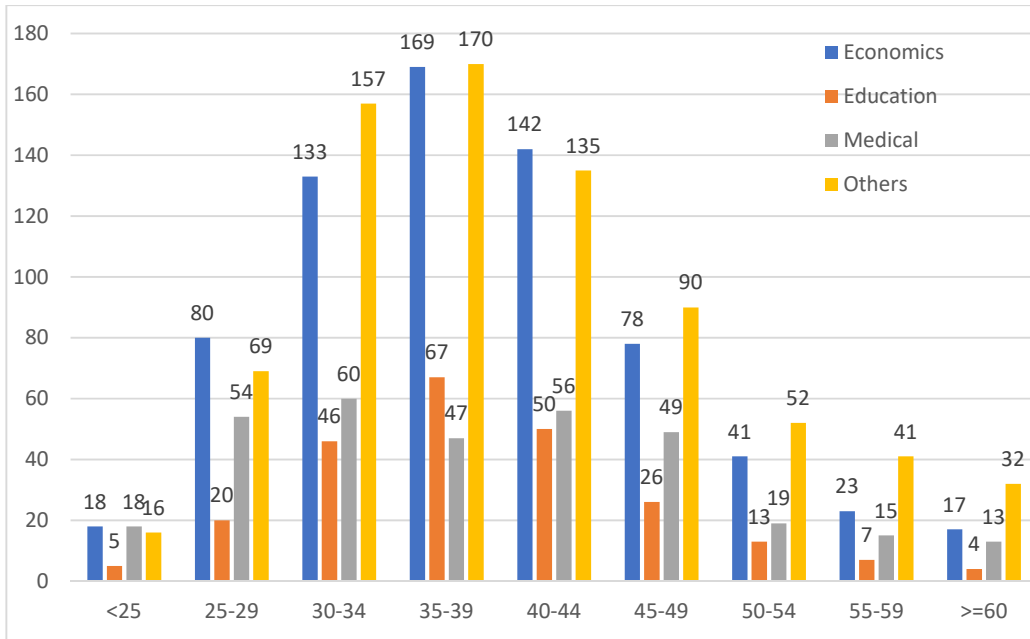
Figure 2. Number of authors in each age group

Figure 3 shows that 40 years old is the average age for researchers in each field. In recent years, the average age has been slowly declining towards 37-38 years old.
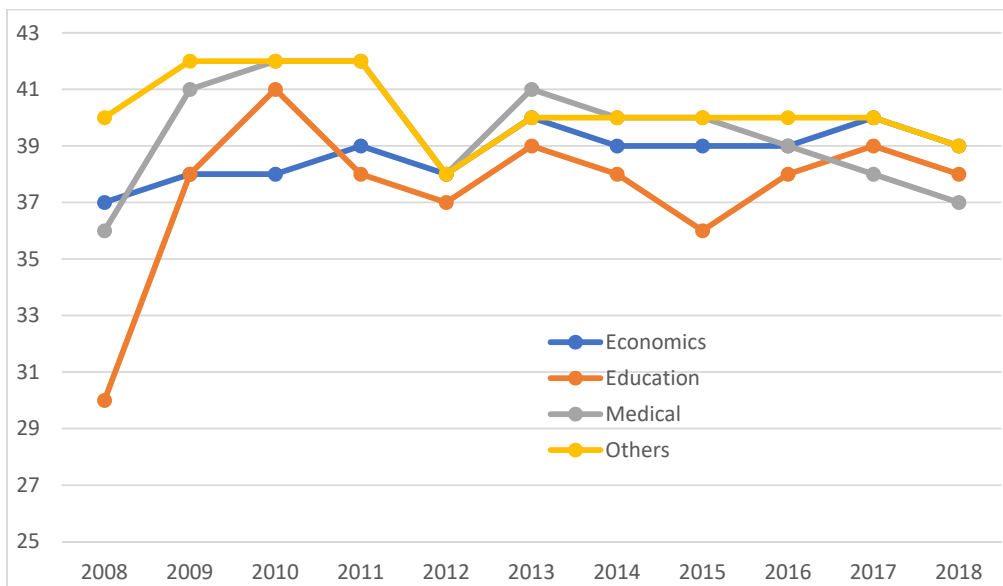

Figure 3. The average age of researchers in each field by year

Figure 4 suggests the accumulation of publications throughout the years. In economics, social medicine, and others, the average number of papers all peaked when authors are in their 60s or more. Notably, the peak of education researchers is in their 35-39, and then the number slowly declines.
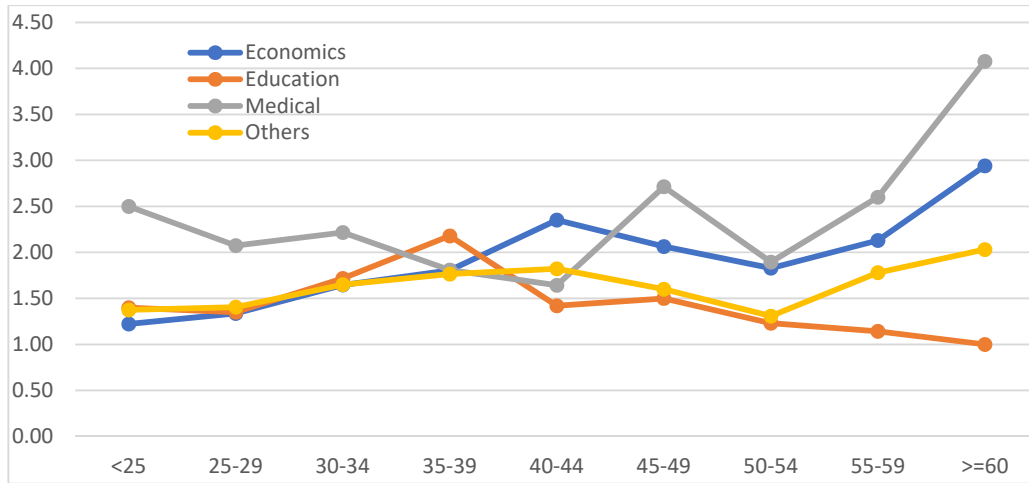
Figure 4. The average number of papers in each age group

1.2. Bayesian analysis

The Bayesian analytical model aimed at determining the association between age of authors in each field and the number of articles can be shown as:

$$O_{article} \sim \alpha_{age}[\alpha_{field}]$$

In which the outcome variable $O_{article}$ is the number of articles in each age group and each field, the dependent variable α includes age and field, which was distributed into two layers of hierarchy. There are four levels of the variable α_{field}: economics, educations, social medicine, and others. The variable α_{age} allocated authors into nine age cohorts of the five years, from 25 to 60+. As there are the same numbers of age cohorts for each field, the age variable eventually has 36 levels of values (4 fields × 9 age groups).

The Stan code is presented below:

```
data{
    // Define variables in data
    int<lower=1> Nobs;    // Number of observations (an integer)
    real article[Nobs];   // outcome variable
    int Nagenum;
    int<lower=1,upper=Nagenum> agenum[Nobs];
    int agenum2fieldnum[Nagenum];
    int Nfieldnum;
    int<lower=1,upper=Nfieldnum> fieldnum[Nobs];
}
transformed data{
    // Define transformed data
}
parameters{
    // Define parameters to estimate
    real<lower=0> sigma_article;
    real<lower=0> sigma_agenum;
    vector[Nagenum] u_agenum;
    real a0_fieldnum;
    real<lower=0> sigma_fieldnum;
    vector[Nfieldnum] u_fieldnum;
}
transformed parameters{
    // Transform parameters
    real mu_article[Nobs];
```

```
        real a_agenum[Nagenum];
        vector[Nfieldnum] a_fieldnum;
        // Varying intercepts definition
        for(k in 1:Nfieldnum) {
            a_fieldnum[k] = a0_fieldnum + u_fieldnum[k];
        }

        // Next level random intercepts
        for(k in 1:Nagenum) {
            a_agenum[k] = a_fieldnum[agenum2fieldnum[k]] + u_agenum[k];
        }

        for (i in 1:Nobs) {
            mu_article[i] = a_agenum[agenum[i]];
        }
}
model{
        // Priors
        sigma_agenum ~ normal(0,10);
        u_agenum ~ normal(0, sigma_agenum);
        a0_fieldnum ~ normal(0,10);
        sigma_fieldnum ~ normal(0,10);
        u_fieldnum ~ normal(0, sigma_fieldnum);

        // Likelihoods
        article ~ normal(mu_article, sigma_article);
}
generated quantities {
        // simulate data from the posterior
        real yrep_article[Nobs];
        // log-likelihood posterior
        vector[Nobs] log_lik_article;
        for (i in 1:num_elements(yrep_article)) {
          yrep_article[i] = normal_rng(mu_article[i], sigma_article);
        }
        for (i in 1:Nobs) {
          log_lik_article[i] = normal_lpdf(article[i] | mu_article[i],
sigma_article);
        }
}
```

The results of simulation are as follow:

```
Model Info:
  nodes:      3
  arcs:       2
  scores:     NA
  formula:    article ~ a_agenum[agenum]

Estimates:
Inference for Stan model: 360396df8dfc64b1ff2509fe87675927.
4 chains, each with iter=5000; warmup=2000; thin=1;
post-warmup draws per chain=3000, total post-warmup draws=12000.


             mean se_mean   sd 2.5%  25%  50%  75% 97.5% n_eff Rhat
a_agenum[1]  2.07    0.01 0.48 1.11 1.74 2.08 2.40  3.01  1573 1.00
a_agenum[2]  1.77    0.01 0.32 1.13 1.55 1.77 2.00  2.37  1533 1.00
a_agenum[3]  1.91    0.00 0.26 1.41 1.74 1.91 2.08  2.40  9800 1.00
a_agenum[4]  2.04    0.01 0.24 1.59 1.88 2.04 2.20  2.51  1139 1.00
a_agenum[5]  2.62    0.00 0.25 2.12 2.45 2.62 2.79  3.11  4441 1.00
a_agenum[6]  2.39    0.01 0.32 1.77 2.18 2.40 2.62  3.01  3031 1.00
```

a_agenum[7]	1.99	0.00	0.39	1.22	1.75	1.99	2.26	2.75	11168	1.00
a_agenum[8]	2.27	0.01	0.45	1.39	1.96	2.27	2.56	3.16	1560	1.00
a_agenum[9]	2.39	0.02	0.50	1.45	2.04	2.37	2.72	3.37	791	1.01
a_agenum[10]	2.24	0.01	0.56	1.15	1.86	2.24	2.59	3.36	1842	1.00
a_agenum[11]	2.01	0.01	0.46	1.08	1.73	2.02	2.31	2.91	5140	1.00
a_agenum[12]	1.98	0.02	0.40	1.22	1.72	1.97	2.23	2.82	274	1.01
a_agenum[13]	2.38	0.02	0.35	1.71	2.15	2.38	2.62	3.05	375	1.01
a_agenum[14]	1.92	0.01	0.38	1.17	1.67	1.92	2.18	2.64	2247	1.00
a_agenum[15]	1.98	0.01	0.43	1.11	1.70	2.00	2.27	2.81	6974	1.00
a_agenum[16]	2.03	0.01	0.50	0.98	1.69	2.04	2.37	2.98	6519	1.00
a_agenum[17]	1.97	0.01	0.55	0.85	1.62	1.99	2.36	2.98	2336	1.00
a_agenum[18]	2.20	0.01	0.56	1.05	1.86	2.19	2.56	3.29	10072	1.00
a_agenum[19]	2.56	0.03	0.51	1.59	2.21	2.53	2.89	3.58	390	1.01
a_agenum[20]	2.86	0.01	0.38	2.16	2.61	2.86	3.10	3.63	4226	1.00
a_agenum[21]	3.24	0.01	0.39	2.49	2.97	3.24	3.50	4.02	3718	1.00
a_agenum[22]	2.26	0.01	0.38	1.55	2.00	2.26	2.52	3.01	676	1.00
a_agenum[23]	2.08	0.01	0.35	1.37	1.85	2.09	2.33	2.77	4480	1.00
a_agenum[24]	3.31	0.01	0.43	2.48	3.00	3.30	3.60	4.19	2021	1.00
a_agenum[25]	2.12	0.01	0.47	1.20	1.83	2.12	2.43	3.05	5216	1.00
a_agenum[26]	2.67	0.01	0.50	1.74	2.33	2.64	2.99	3.72	3042	1.00
a_agenum[27]	3.02	0.01	0.59	1.96	2.62	2.99	3.37	4.28	2823	1.00
a_agenum[28]	2.26	0.01	0.47	1.33	1.96	2.26	2.56	3.20	8383	1.00
a_agenum[29]	1.88	0.00	0.33	1.23	1.67	1.87	2.10	2.52	8219	1.00
a_agenum[30]	1.92	0.00	0.24	1.45	1.76	1.92	2.08	2.41	7934	1.00
a_agenum[31]	2.08	0.00	0.23	1.61	1.93	2.08	2.24	2.53	6191	1.00
a_agenum[32]	2.14	0.01	0.26	1.66	1.96	2.14	2.32	2.66	1597	1.00
a_agenum[33]	1.99	0.01	0.31	1.37	1.77	1.98	2.20	2.58	448	1.01
a_agenum[34]	1.72	0.00	0.37	0.97	1.49	1.72	1.97	2.41	6220	1.00
a_agenum[35]	2.35	0.04	0.44	1.14	2.09	2.37	2.64	3.14	133	1.02
a_agenum[36]	2.22	0.01	0.42	1.44	1.93	2.22	2.51	3.06	1044	1.01
sigma_agenum	0.55	0.00	0.14	0.31	0.45	0.54	0.64	0.85	2026	1.00
a_fieldnum[1]	2.32	0.01	0.22	1.94	2.18	2.31	2.46	2.79	992	1.00
a_fieldnum[2]	2.10	0.00	0.22	1.62	1.97	2.12	2.26	2.51	4999	1.00
a_fieldnum[3]	2.32	0.01	0.21	1.94	2.18	2.31	2.45	2.76	982	1.00
a_fieldnum[4]	2.23	0.00	0.20	1.83	2.09	2.23	2.35	2.63	3075	1.00
a0_fieldnum	2.24	0.01	0.29	1.68	2.10	2.24	2.37	2.82	1172	1.00
sigma_fieldnum	0.36	0.01	0.46	0.03	0.11	0.23	0.43	1.48	1584	1.00

In which, the coefficients a_agenum is the coefficients of age in each respective field with 36 levels of value. The coefficient a_fieldnum is the coefficient for four levels of value of each field. According to the result of the simulation, Rhat is approximately 1, and n_eff is above 1000 samples, which suggests the results are good.

The MCMC chains are shown in Figure 5. Overall, all the chains are resembled, suggesting the autocorrelation phenomenon.
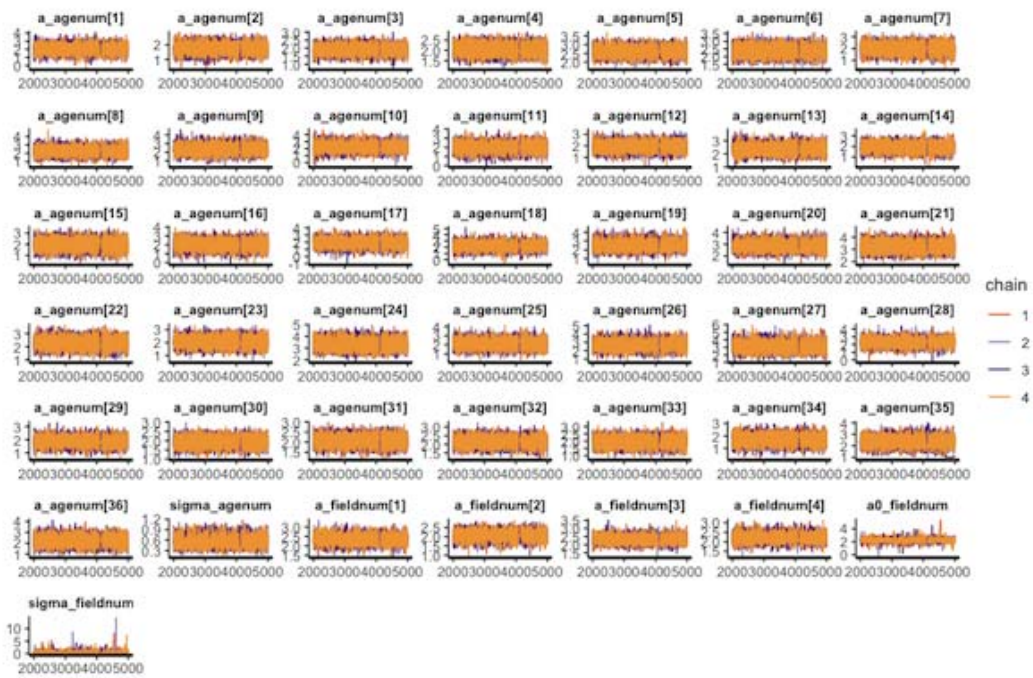
Figure 5. The MCMC chains for the Bayesian model of age.

In Figure 6, the productivity of each field is presented. The average lines in economics, education, and others are similar, around 2.1 articles. In social medicine, the average is higher, with 2.8 articles. In economics (Figure 6a), most of the authors who are under 40 have fewer publications than the average (the red line), except the age group of 50 - 54. The age group of 40 – 44 has been leading in productivity, while the age group of under 25 is aiming to surpass the average line.

In education (Figure 6b), most of the age groups have similar productivity, only the age groups of 35 – 39, older than 60, and under 25 are achieving better the average productivity. For social medicine (Figure 6c), the 45 – 40 age group is with the highest productivity, while in the others (Figure 6d), the 55 – 59 has the highest productivity.
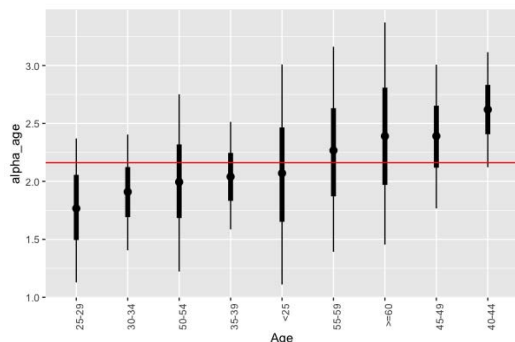
Fig 6a) Economics
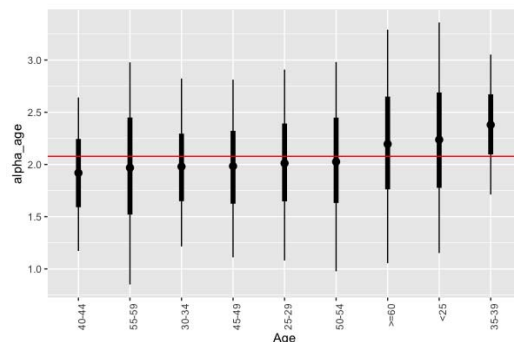
Fig 6b) Education
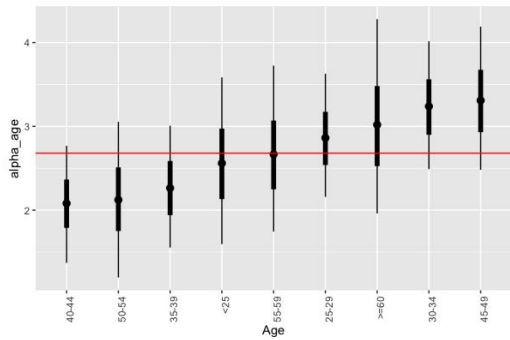
Fig 6c) Social Medicine   Fig 6d) Others
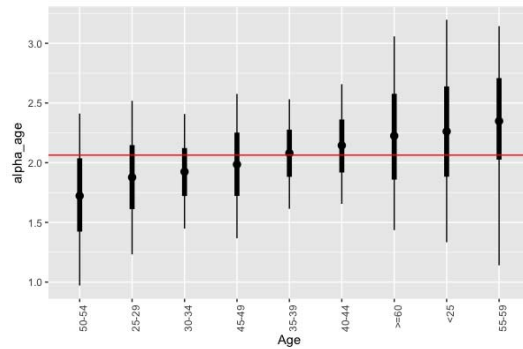
Figure 6. The productivity of researcher according to age and discipline

2. Gender

This section aims to answer the following questions:

- Is there any gender difference between the number of publications and researchers in each field?
- What is the difference between the age of male and female researchers?

2.1. Descriptive analysis

In Figures 7 and 8, males have a significantly higher quantity of publications as well as the number of researchers than females in economic and other fields. In the medical field, there was a slightly higher number of publications by males than females, but the number of female researchers exceeded that of male counterparts; therefore, gender disparity was not observed in the medical field. Interestingly, in the field of education, females surpassed males in terms of both the number of publications and researchers.
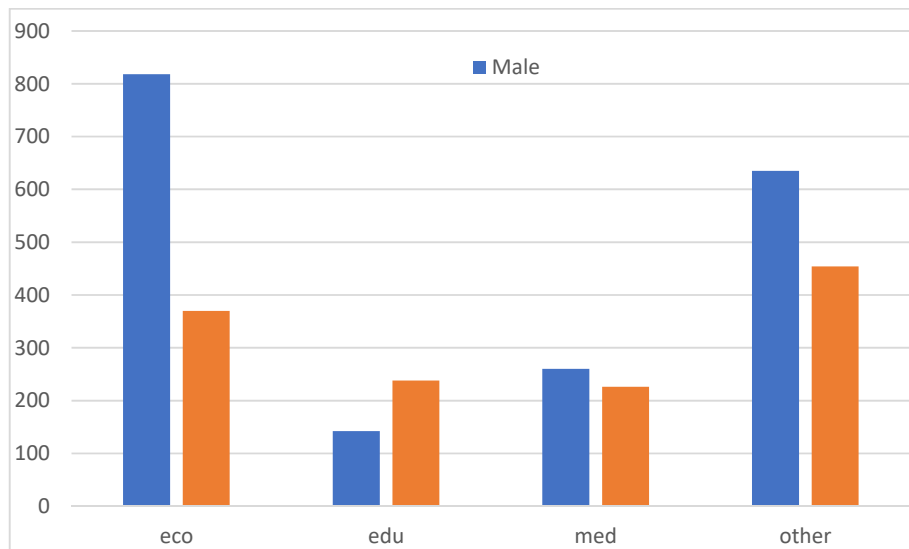
Figure 7: The total number of publications by gender and discipline
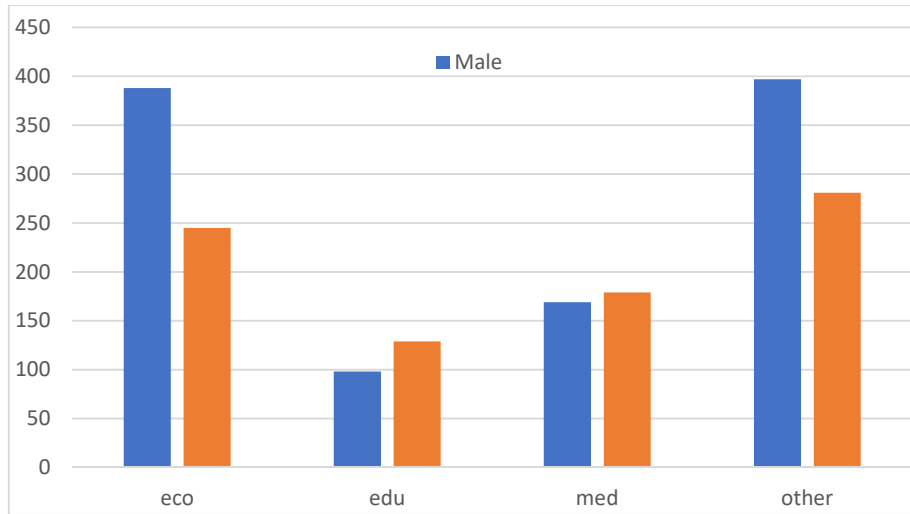
Figure 8: The number of researchers by gender and discipline

Table 2 shows that the average age of male researchers is relatively higher than the average age of female counterparts. In both genders, researchers in the economic field had the lowest average age, while those in other fields had the highest average age. However, as an author could be inputted many times in the database at different ages, the difference in the average age between male and female researchers is demonstrated more clearly from the longitudinal angle (see Figure A).

Table 2: Average age of male and female researchers by discipline

Field	Male	Female
eco	41	36
edu	41	36.8
med	41.8	37.3
other	42.4	37.4

Different from the increase of average age of researchers over year in the economic field, opposing trends between male and female were found in the educational, medical, and other fields from 2010 (see Figure A.1 to A.4). A rejuvenation tendency among male researchers was found in all fields, excluding Economics, whereas the average age of female researchers was getting higher.

2.2. Bayesian analysis

In the Bayesian analysis, the outcome variable is the number of publications. The model consists of two independent variables of gender – *"sex"*, and research field – *"field"* with two hierarchies as follows:

$$O_{article} \sim \alpha_{sex}[\alpha_{field}]$$
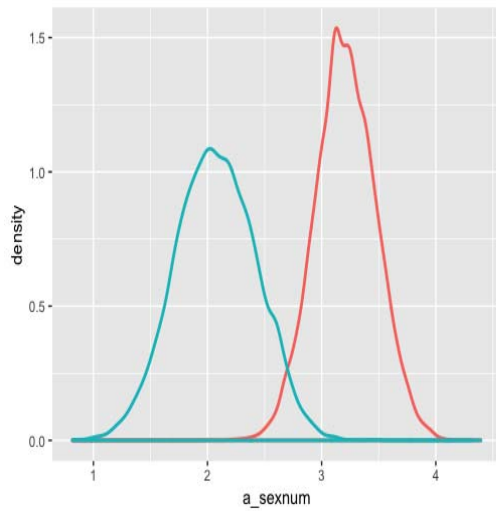
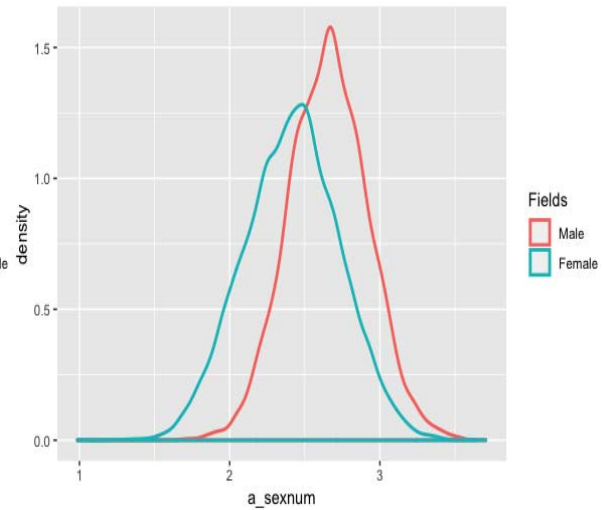Fig 9a): Economics
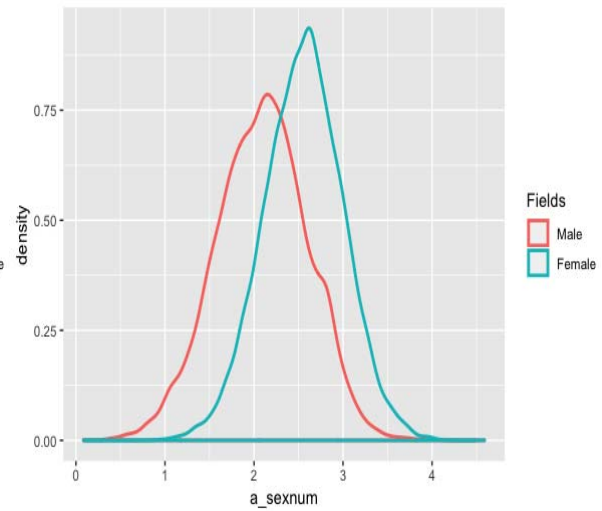
Fig 9b): Education
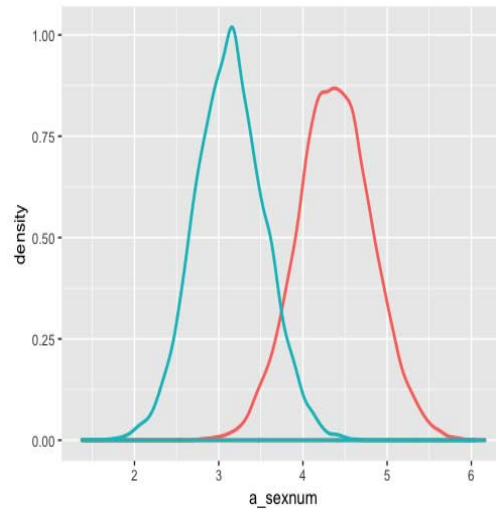
Fig 9c): Medicine

Fig 9d): Other fields

Figure 9: The article outcome according to gender and discipline

Figure 9 displays the distribution of posterior probabilities according to gender in each field. In the economic and medical fields, male researchers had a substantially higher probability of producing more articles than female researchers. Male researchers were also more productive than female counterparts in other fields, but the disparity was small (see Figure 9d). Only in the educational field did female researchers have slightly higher outcomes than male researchers (see Figure 9b).

3. New authors

A new author is the author first-time who appears in the SSHPA database in a specific year from 2008 to 2018. However, the data in 2008 is not accurate as all of the authors inputted are new authors, so the data from 2010 will provide higher accuracy.

This section aims to answer the following questions:

- What are the variations in the number and the growth rate of new authors across research fields from 2010 to 2018?
- What is the average age of new researchers in each field?

- Is there any difference in the gender of new authors in each field?

In general, from 2010 to 2018, the number of new authors among Social Sciences and Humanities in Vietnam grew rapidly, which hints at the significant increase in human resources that are capable of doing proper research (see Figure 10). During 2010 and 2018, the annual growth rate was approximately 21.82%.
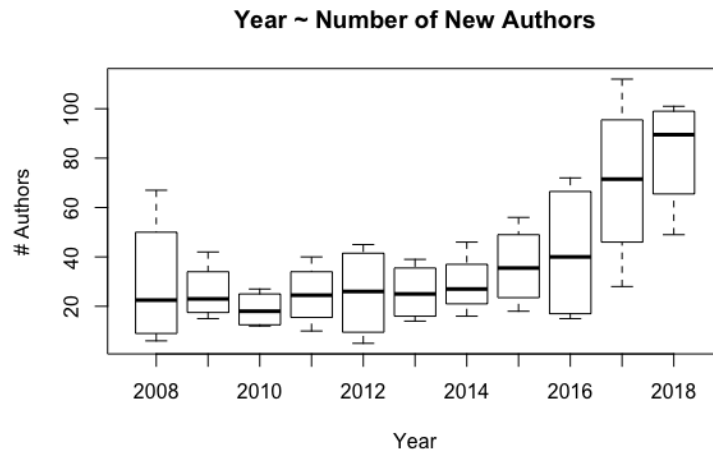

Year ~ Number of New Authors

Figure 10: The number of new authors in Social Sciences and Humanities during 2008-2018

Economic and other fields obtained the greatest and the second greatest increase in the number of new researchers, respectively, but Medicine and Education were the two fastest-growing fields regarding new researchers. Figure 11 presents the variations in the number of new authors during 2010 and 2018 in four categories: economics, education, social medicine, and other fields. The economics field observed the utmost increase of new authors with 497 people in total; other fields came after with 469 people in total. Nevertheless, in terms of annual growth rate, medical and educational fields held first and second ranks with 42.90% and 35.64%, respectively.
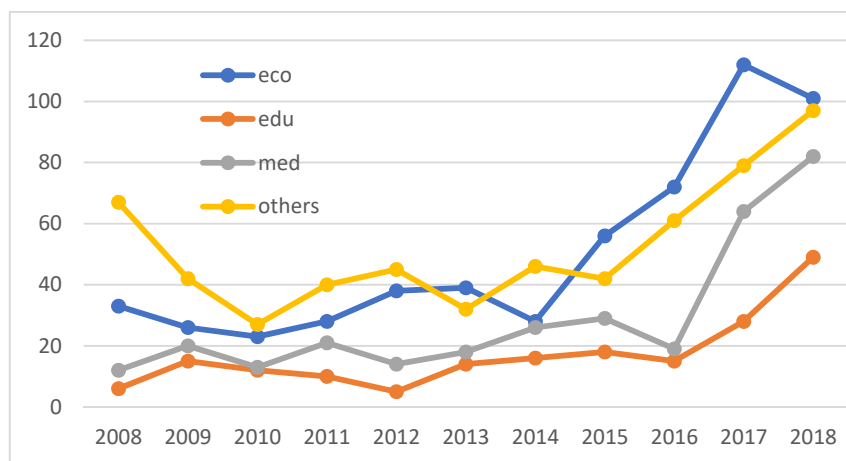
Figure 11: The number of new authors across disciplines during 2010 and 2018

Among fields, new authors from other fields had the highest average age at 39, while new authors in economic and medical fields had the lowest average age at 37. From Figure 12a to 12d, there is no clear increasing or

decreasing pattern across the field. However, the yearly average age of new authors in economic, educational, and other fields was relatively convergent to the aggregate average age, while in medical fields, the average age had an erratic fluctuation over the years.
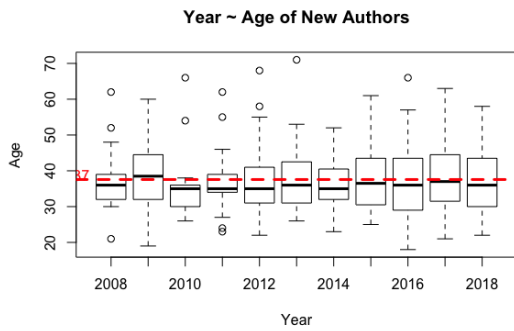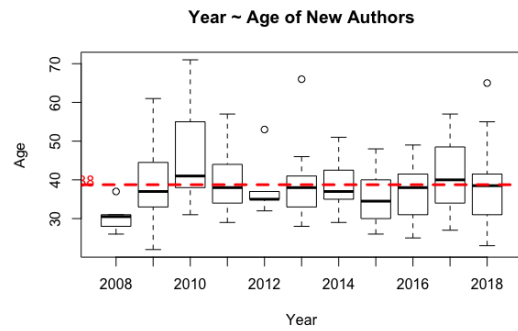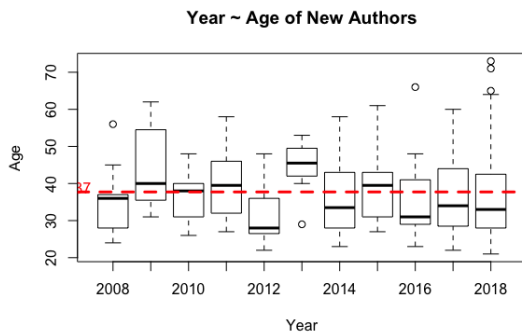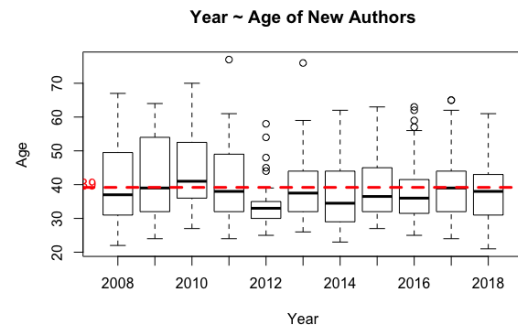

Fig 12a): Economics

Fig 12b): Education

Fig 12c): Medicine

Fig 12d): Other fields

Figure 12: The average age of new authors among discipline during 2008-2018

The number of new male and female researchers was on the rise in all fields from 2010 to 2018, but there were some differences across fields (see Figure 13a to 13d). On the one hand, the number of new female educational researchers surged and even doubled the number of male counterparts in 2013 and 2018 (see Figure 13b). The distribution of new male and female authors in the medical field was comparatively equal, and in some years – 2010, 2012, 2016, new female researchers exceeded male researchers (see Figure 13c).

On the other hand, in the economic and other fields, the number of new male researchers was dominant, despite the rapid growth rate of new female researchers.
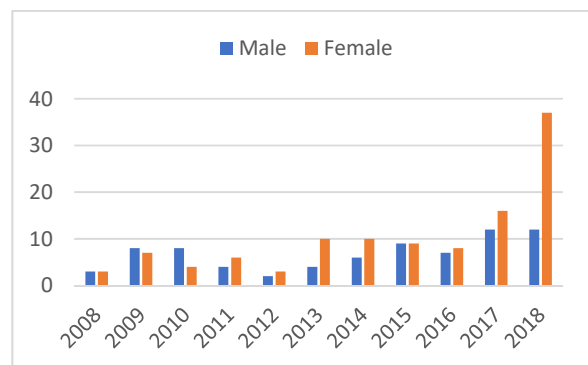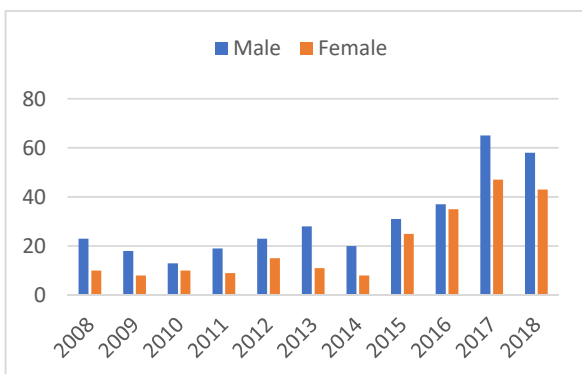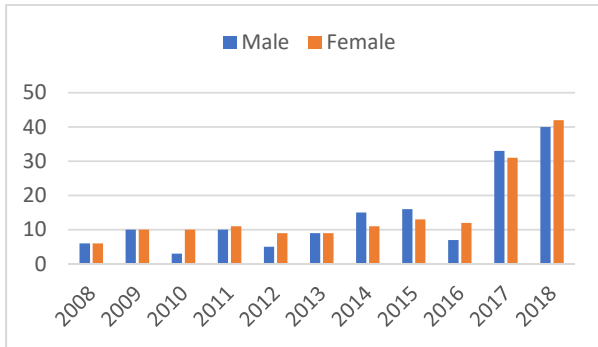
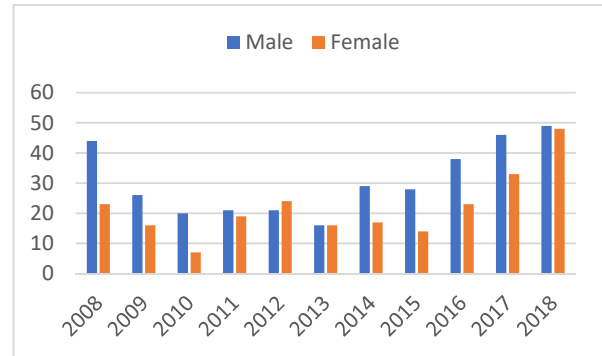Fig 13a): Economics

Fig 13b): Education

Fig 13c): Medicine

Fig 13d): Other fields

Figure 13: The number of new authors by gender and discipline during 2008-2018

As for the average age among new male and female researchers, new male researchers were found to be older than female counterparts in almost fields – economic, educational, and other fields (see Figure B). The average age of new male and female medical authors fluctuated over time and did not illustrate a clear pattern (see Figure A.3). Although new male researchers in educational and other fields were older than females, the differences were not substantial. It is remarkable that new female economic authors were much younger than males, but the age gap had gradually diminished since 2014 (see Figure A.1).

4. Lead authors

Lead authors are defined in this paper as authors who drastically outperform others in output; they are comprised of outliers in the dataset. The numbers of new leaders were computed on a yearly basis. A new 'leader author' is defined as someone who had broken into the ranks of performance outliers based on accumulated scientific output up to that year. This section aims to answer the following questions:

- Does the number of lead authors grow exponentially, and is there any difference among the fields?
- What are the differences based on age, the gender of the lead authors, in comparison with other authors?

The figure below presented the growth in the number of new authors per year from 2008 to 2018 (see Figure 14). Each box plot represents the distribution of the number of new lead authors by discipline. There is a clear upward trend with small fluctuations; the overall increase seemed exponential, picking up steeply around the year 2014.
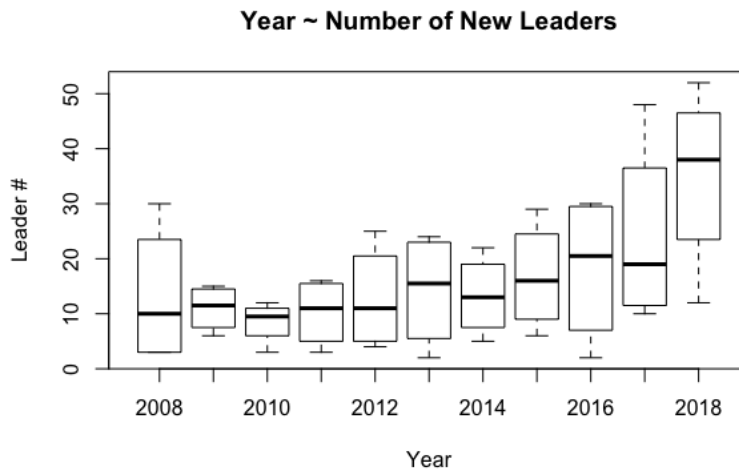
Figure 14. Number of new leading authors by year.

What is the profile of these outstandingly productive authors, and how did they differ from their more average counterparts? We first studied the mean age of lead authors, which was described in Figure 15.
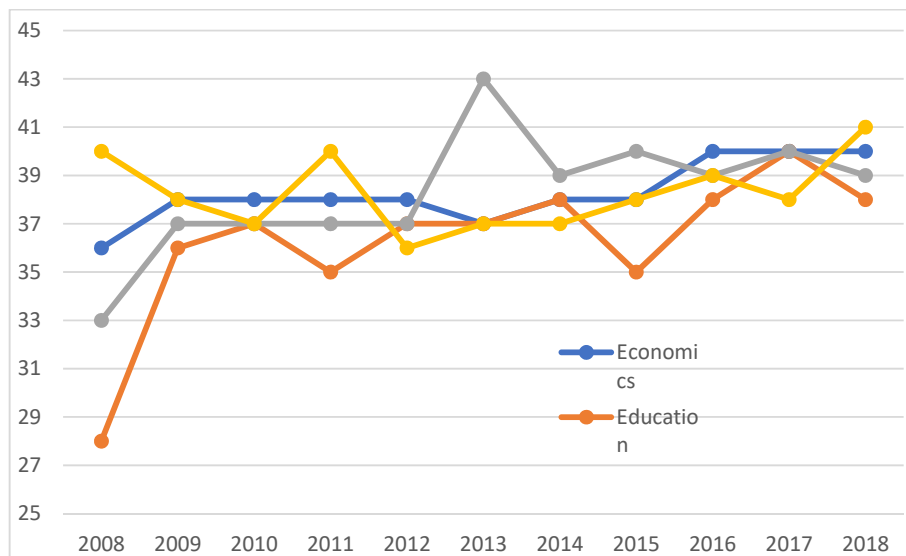

Figure 15. Mean age of leading authors by year and discipline.

The mean age of lead authors seemed to hover around mid-thirties to early forties. There is a very slight slope upwards, suggesting that authors were having their "boom" in productivity at an increasingly older age. As a general observation, lead authors seem to be the youngest in education and the oldest in economics, on average. The social medicine field, notably, showed an outlier in 2013, perhaps skewed by a particular individual.
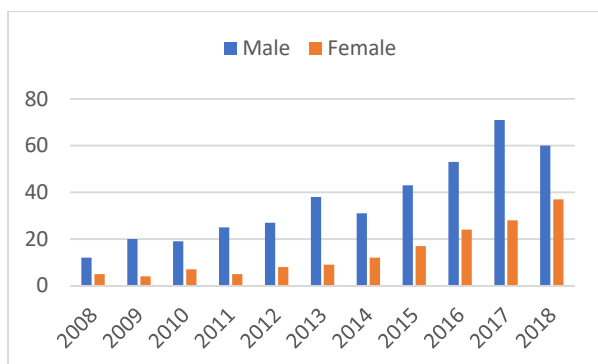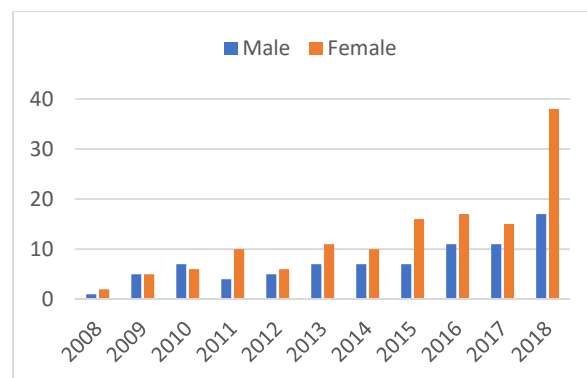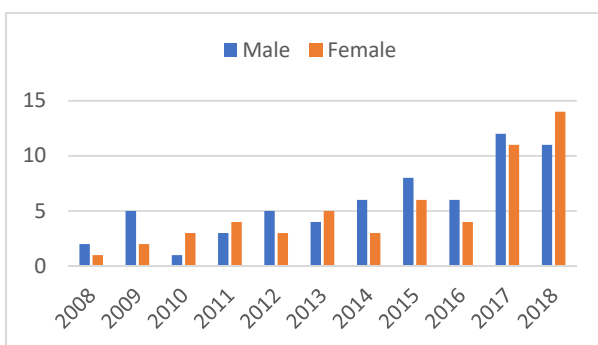
Fig 16a): Economics

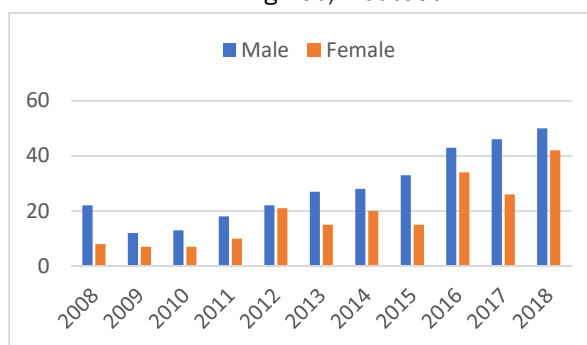
Fig 16b): Education

Fig 16c): Medicine

Fig 16d): Other fields

Figure 16. Number of leaders by gender and fields

Figure 16 provided a more detailed descriptive view of lead authors by gender in each discipline. As a general observation, there were more males than females, most noticeably in economics where the number of males nearly doubled that of females. The reverse could be observed in education, where women had consistently been dominating since 2011 and even had a drastic increase in 2018. The medical discipline, on the other hand, seemed to strike a delicate gender balance.

All this seemed to be a remarkable improvement in terms of gender balance in Vietnamese SSH; however, one must not overlook the fact that the majority of researchers remained in economics, the discipline with stark male dominance. Economics was also the discipline receiving the most investments and grants and regarded as the most substantial social science discipline, especially for their role in national growth. More on this in discussions.

5. Co-authorship

This section aims to answer the following questions:

- How is the number of articles distributed according to the number of co-authors among disciplines?
- What is the difference in terms of the number of international collaborations among disciplines?

The social medicine field was the most collaboration-oriented field among disciplines; papers were usually published by groups of co-authors ranging mostly from 2 to 11 authors (see Figure 17c). In economic and other fields, most of the papers were written by 1 to 5 (co-)authors, but the dominance of solo paper was witnessed in other fields (see Figure 17a and 17d, respectively). In contrast to the common collaborating trend, social medicine was relatively unique; papers in this field obtained a fewer number of co-authors than other disciplines (see Figure 17b).
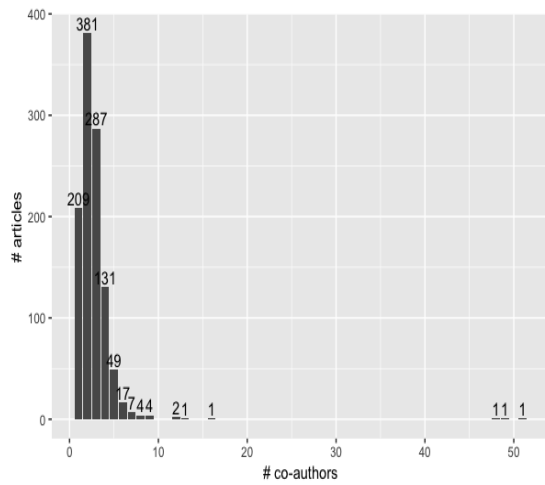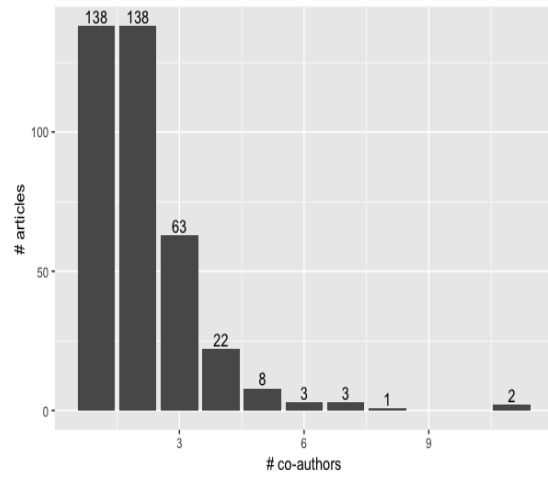
Fig 17a): Economics

Fig 17b): Education
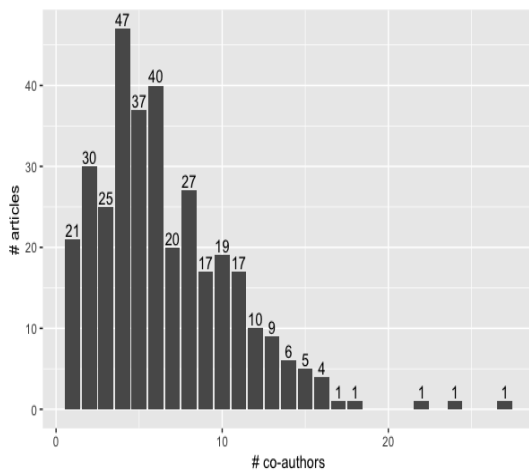
Fig 17c): Social medicine

Fig 17d): Other fields

Figure 17: The distribution of articles according to the number of co-authors in each field

Among disciplines, the economics obtained the highest number of international collaborations with 860 articles, and other fields came after with 966 articles (see Figure 18). However, the highest percentage of international collaboration was in the medical field, with 57.5%, whereas the educational field occupied the lowest percentage of internationally collaborated papers with merely 33.2%. These results hint at the preference for international collaboration in all fields, leading by medical field and excepting educational field.

Fig 18a): Economics

Fig 18b): Education
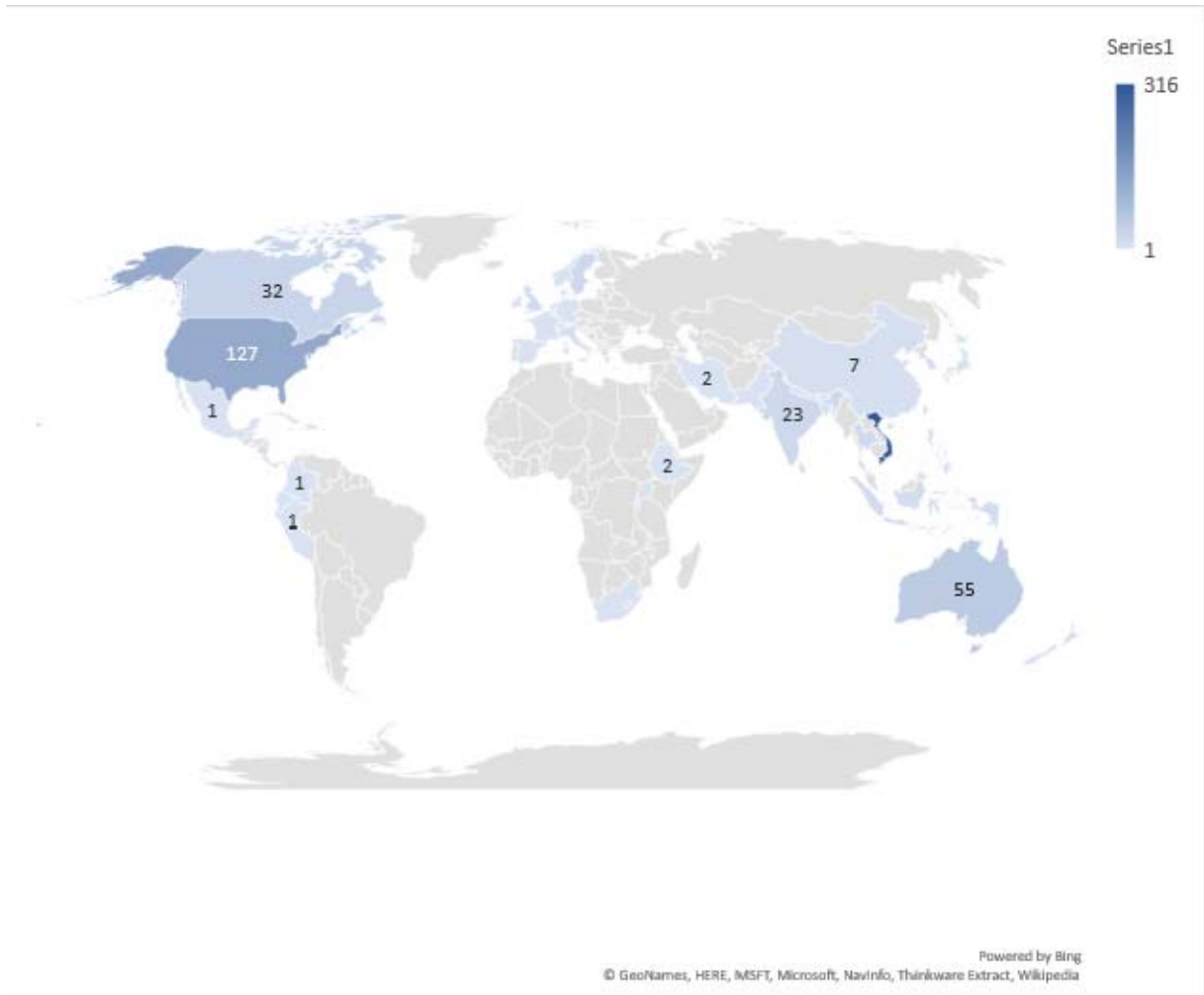
Fig 18c): Social medicine

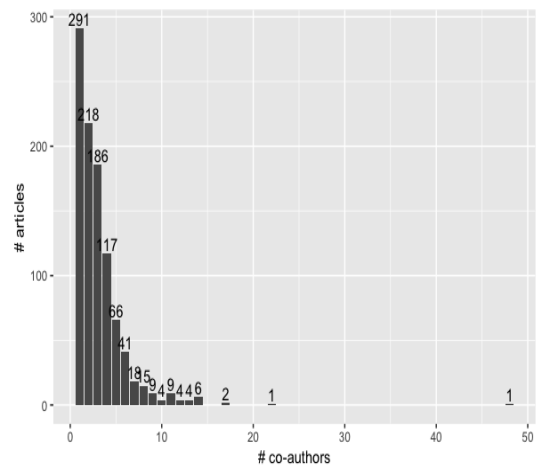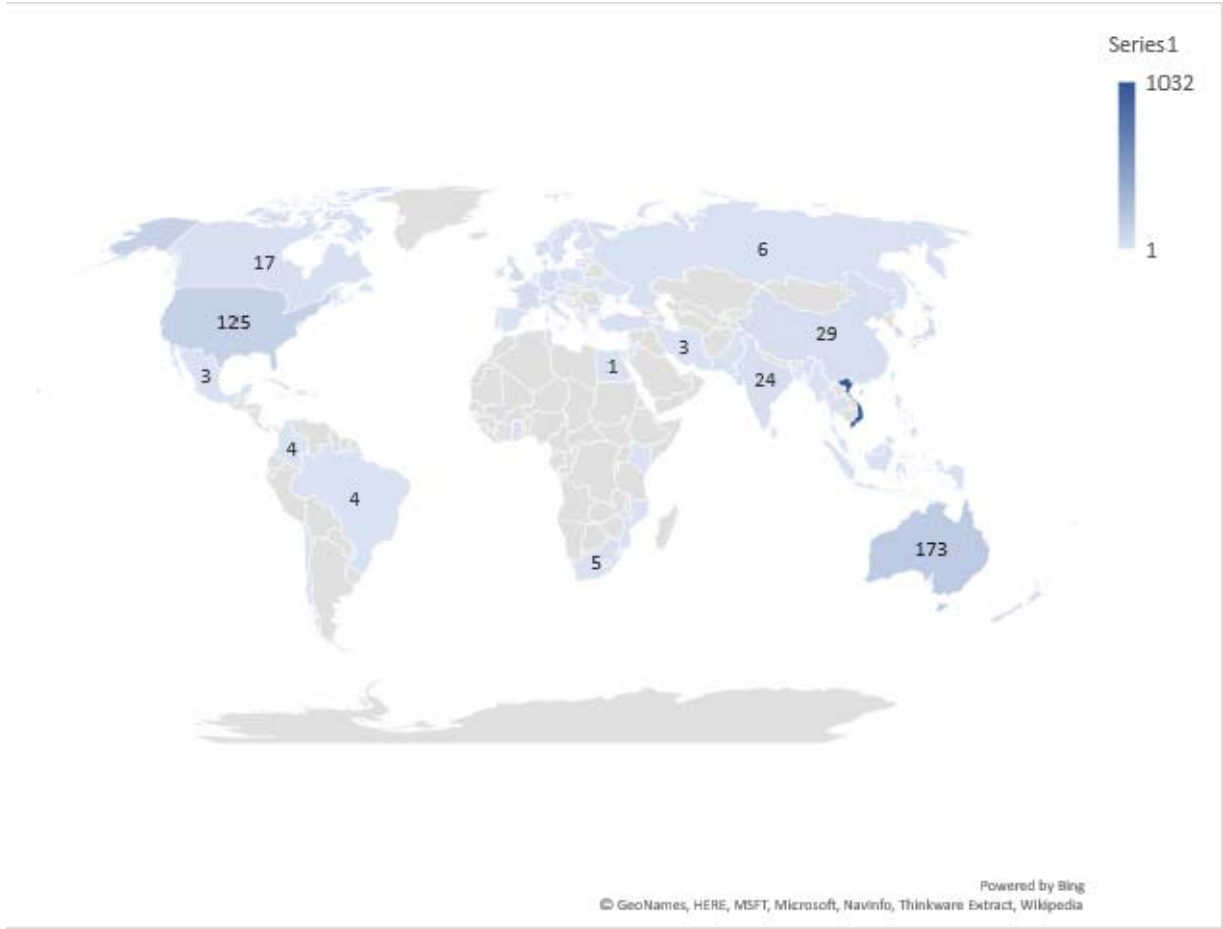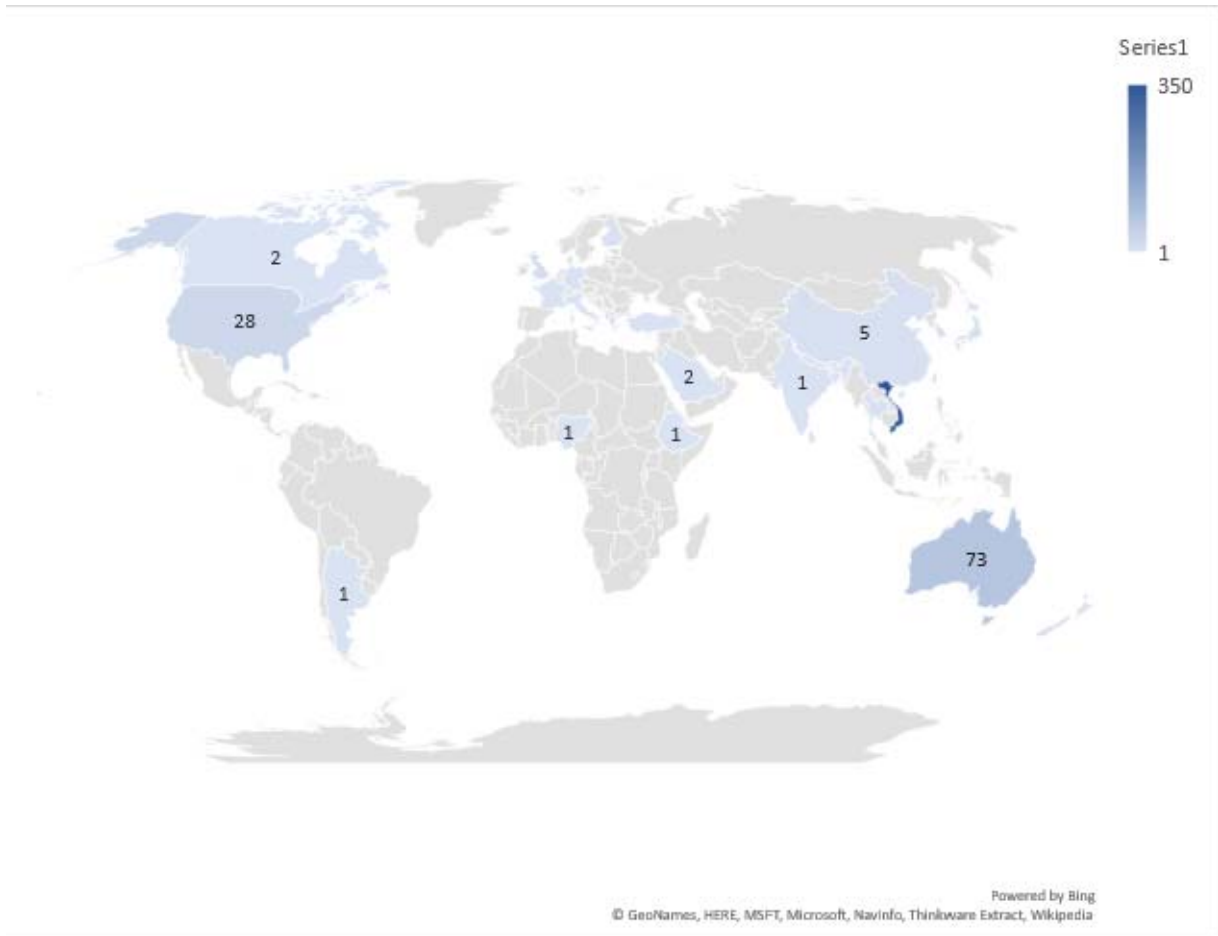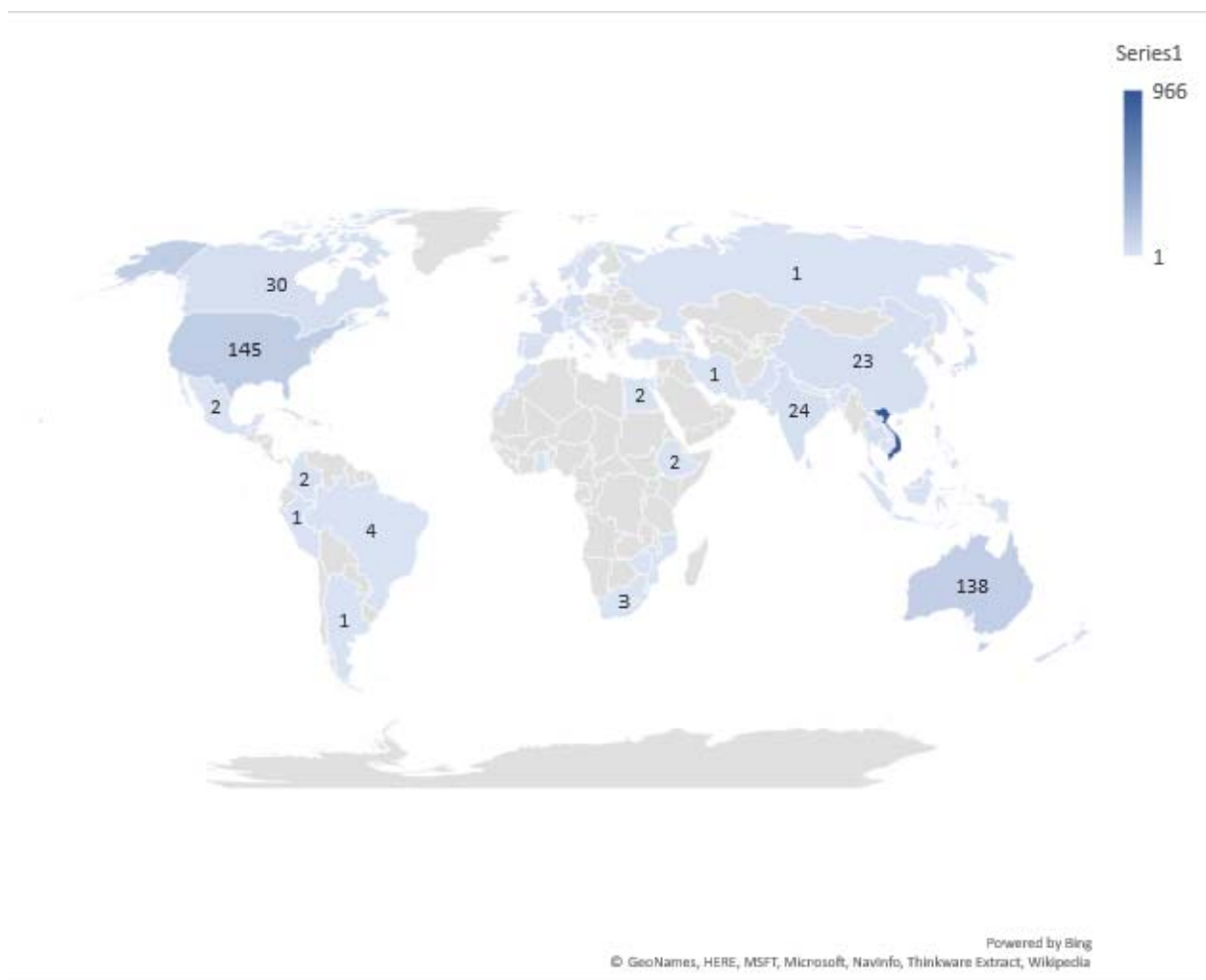Fig 18d): Other fields

Figure 18: The distribution of international collaborations by countries among disciplines

Compared to education and social medicine, economics and other fields had a remarkably higher number of international partners with 62 and 56 countries, respectively; most of the collaborations were with a wide range of countries around the world. Meanwhile, the collaboration network in education and social medicine was narrow, with only 28 and 40 countries, respectively; the collaborating countries were mainly in North America, Australia, Asia, and Europe. Australia and North America were two major collaborating partners across disciplines, which indicates the collaborating tendency with Western developed countries within the Vietnamese scientific community.

6. Journal Impact Factor

This section aims to answer the following questions:

- How is the number of articles distributed by impact factor groups among disciplines?
- How is the number of articles led by Vietnamese researchers distributed by impact factor groups among disciplines?
- Does the impact factor among disciplines grow over the year?
- Is the impact factor influenced by the age of Vietnamese lead authors among disciplines?

Table 3 presents the percentage of articles according to the impact factor group among disciplines. Educational papers published in journals without impact factor occupied 75.7% of the total articles, which was the highest proportion among the four disciplines. In contrast, articles in medical fields obtained relatively high impact factors, as 65.4% of total articles had an impact factor higher than 1. The impact factor of papers in economic and other fields was comparatively even that more than 90% of articles received impact factor less than 3.

Table 3: The number and percentage of articles according to impact factor group among disciplines

	=0		<=1		<=2		<=3		<=4		<=5		<=6		<=7		<=8		>8		Total
eco	602	59.5%	133	13.2%	156	15.4%	63	6.2%	39	3.9%	14	1.4%	1	0.1%	2	0.2%	1	0.1%	0	0.0%	1011
edu	259	75.7%	31	9.1%	35	10.2%	14	4.1%	1	0.3%	0	0.0%	2	0.6%	0	0.0%	0	0.0%	0	0.0%	342
med	97	30.8%	12	3.8%	62	19.7%	91	28.9%	30	9.5%	15	4.8%	5	1.6%	0	0.0%	1	0.3%	2	0.6%	315
other	516	55.1%	106	11.3%	143	15.3%	89	9.5%	40	4.3%	23	2.5%	11	1.2%	3	0.3%	1	0.1%	4	0.4%	936

Excluding articles published in journal without impact factor, a number of papers led by Vietnamese researchers in economic and educational field were mainly published in journals whose impact factor was less than or equal to 2, while those in medical field were mostly published in journals whose impact factor ranged from between 2 and 3 (see Figure 19). The number of articles published in journals whose impact factor was more than 5 dropped dramatically, especially in the educational field. Besides, there was merely one paper published in a journal with an impact factor of 8 or higher.
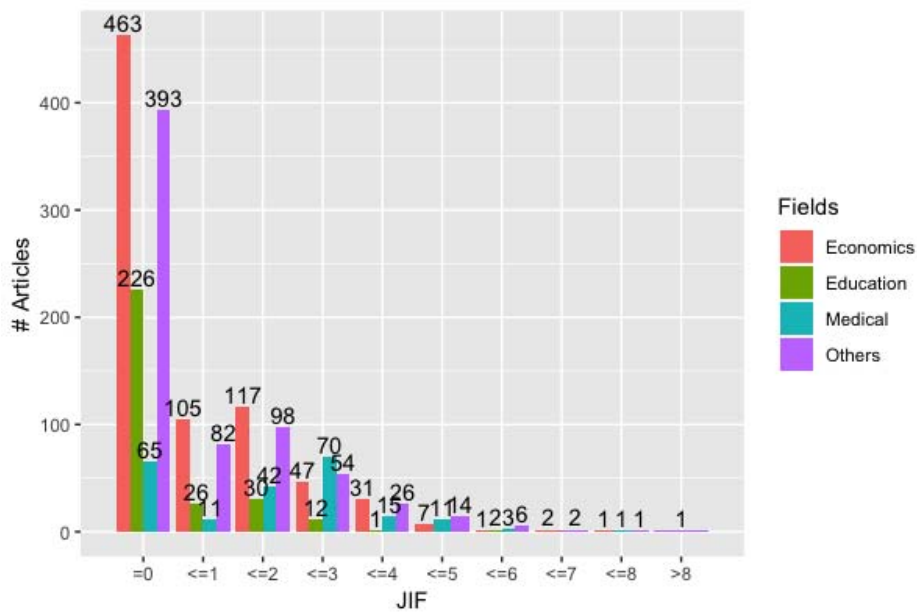
Figure 19: The number of articles led by Vietnamese researchers according to impact factor group among disciplines

In the economic and other disciplines, the average impact factor did not have a specific pattern and fluctuated erratically over time (see Figure C1 and C4). During the period between 2010 and 2018, the medical field observed a decreasing trend of average impact factor, whereas the education field witnessed an increasing

tendency of impact factor from 2014 to 2018 (see Figure C2 and C3). However, the medical field still obtained the highest average impact factor at 2.58, while other fields came after at 2.15.

The average age of Vietnamese lead author in the economics maintained almost unchanged as the impact factor rose (see Figure 20). In the educational field, there was a clear pattern that papers, which have been published in higher impact factor journals, were led by older Vietnamese authors. Similarly, other fields witnessed a similar pattern, although the pattern was not clear. On the contrary, younger Vietnamese lead authors were more likely to obtained publications in higher impact factor journals in the medical field.
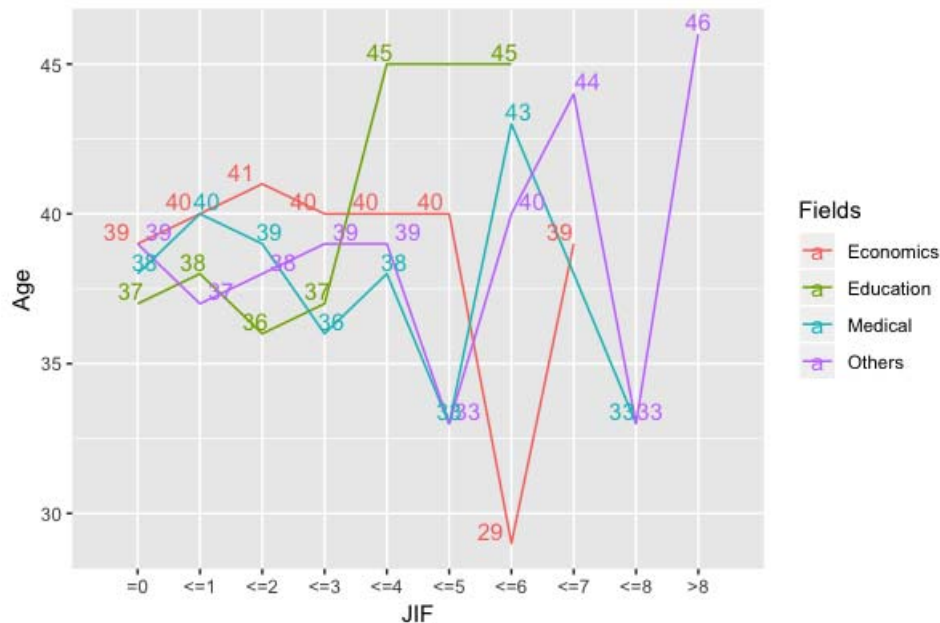

Figure 20: The average age of Vietnamese lead-authors plotted with impact factor group and discipline

V. Discussion

1. Age

The high number of authors from the age group of 30 – 49 and their high level of contribution are understandable as the authors in these age groups are at the beginning of their careers. A Ph.D. candidate is expected to finish his/her Ph.D. and enter the job market at early 30s. Thus, a researcher supposes to prove his/her skills through publications to have an advantage towards a position in a highly competitive job market (Donnelly et al. 2019). In Vietnam specifically, the introduction of Circular 37/2014/TT-BKHCN in 2014, which required all national projects to result in ISI/Scopus publications (Nafosted 2018), and Circular 08/2017/TT-BGDDT in 2017, which required PhD candidate to publish at least 2 articles in ISI/Scopus journals (H.-K. T. Nguyen et al. 2019), have pressured both doctoral candidates and established scientists to publish. The findings suggest most of Vietnamese SSH authors are following a standard career path. In the foreseeable future, these age groups will continue to have a crucial position in the development of Vietnam social sciences and humanities.

2. Gender

Our study found that the number of publications, as well as the number of researchers in economic and other fields, were dominated by males. This result underlines an excellent gender disparity in economics and other fields, thus, suggest policymakers pay more attention to inequalities that hamper the access and progress of women in science (Larivière et al. 2013). Interestingly, female researchers in education surpass their

counterparts in terms of number and output. This finding might root from the cultural context in Vietnam, in which education-related jobs are believed to be ideal for women (Larivière et al. 2013).

The average age of male researchers, according to our study, is higher than female researchers in all fields. It is also found that while the average age of male researchers tended to be lower over time in all fields, excluding Economics, female researchers' average age was getting higher. The finding hints at the low reinforcement rate of new young female researchers, thus, again, raises the concern about the inequalities that hinder women's access to science in Vietnam.

3. New authors
The current study found that during 2010 and 2018, economic and other fields had the greatest increase in the number of new researchers, but the fastest annual growth rate was perceived in medical and educational fields with 42.90% and 35.64%, respectively. This result emphasizes the leading role of Economics in the rapid development of Vietnamese social sciences in recent years (Vu et al. 2019). However, the position of Economics is shaken by the substantial growth rate of new researchers in the Medical and Educational fields. In the near future, the development of social sciences in Vietnam is expected to be fueled by scientific production from not only Economics but also Education and Medicine (Le et al. 2019).

Different from the dominant number of new male authors in economics and other fields, we also found that the number of new female authors almost equalized with their counterparts in the social medicine field and even overshadowed the number of new male authors in the educational field. This finding signals the narrowing gender disparity in the medical and educational fields. Nonetheless, given the ongoing imbalance between the reinforcements of male and female new authors in economic and other fields, policymakers are suggested to target economic and other fields for confining gender inequalities.

4. Lead authors
The findings suggest lead authors of Vietnamese SSH are often around mid-thirties to early forties, with the oldest in economics and youngest in education. In terms of gender, there are more male than female researchers, especially in economics. However, education is the field where the female had been outperforming their male counterparts. Previous studies using the dataset from 2008 - 2017 suggested Vietnamese SSH revolves around a number of highly connected individuals (T. Ho et al. 2017; T. M. Ho et al. 2017), and it is a good sign that the number of Vietnamese lead authors is also rising steadily, along with the number of publications.

5. Co-authorship
Among the four different disciplines, we found that Education is the least collaborating and international field. Unlike Education, authors in the other three fields prefer international collaboration and larger group cooperation. The diffusion of scientific co-authoring and international collaboration among disciplines of Social Sciences in Vietnam from 2008 to 2018 might result from the Vietnamese government's pursuit of science policies incentivizing strong research groups and international collaboration (Nhan 2017; H.-K. T. Nguyen et al. 2019). This result is aligned with the common world trend on the increase in the average number of co-authors, share of co-authored and international co-authored articles in Social Sciences (Henriksen 2016). Besides incentives given by the government, the pressure from 'publish or perish' can be another explanation for the rising co-authoring and international collaborating patterns. As Ph.D. students in Vietnam have been required to obtain at least two publications in international journals for qualification since 2017, the co-authoring tendency between supervisors and students has become more popular (Price et al. 2000; Fisher et al. 2013; Vuong 2019a).

6. Journal Impact Factor

The level of high Journal Impact Factor is subjective in each field. In economics, more than 50% of the field's output was in publications with no impact factor, while publications with JIF from 1 to 2 occupied 28.6% of the field's output. Articles in the educational field were observed to be published in relatively low impact factor journals; 75.7% of the total articles are issued in publications with no impact factor. In contrast, medical field related papers have the highest impact factor; the number of medical papers published in journals with an impact factor of 1 or above accounts for more than 65% of the total publications. However, the excessive gap between two fields might also be derived from the larger group of co-authors and a higher proportion of international collaboration, given the benefits when co-authoring and internationally collaborating, namely increasing visibility, lower cost, greater scale, and higher creativity (Wagner 2006; Wagner et al. 2001).

These figures should not be mistaken for the overall quality of the research practice in each field. However, to a certain degree, the number of publications with a high impact factor can suggest the maturity of the field because it is certainly not easy to publish in high ranking journals. Thus, social medicine, with its connection to the field of medicine, has frequently entered high ranking journals. Meanwhile, even though economics and education are leading with a high number of publications, most of these publications are in journals with a low impact factor. In general, as a nation's science needs both quantity and quality to advance further, Vietnam SSH still needs more quality publications, at least in terms of impact factor, to help the Vietnam scientific community establishes as a developed scientific community.

VI. Conclusion

Employing the power of the SSHPA database (http://sshpa.com/) to extract a comprehensive dataset of 1,564 Vietnamese SSH authors with 2,410 international publications in from 2008 to 2018, this article has analyzed and illustrated the scientific productivity of Vietnamese economics researchers in terms of age, gender, new authors, leading authors, co-authorship, and Impact Factor, comparing with their social medicine, education, and other fields counterparts. In 12 years, authors at the age of 40 – 44 in economics has contributed 858 publications, which is twice of the education's 397 articles, and equivalent to the social medicine total output. Moreover, male researchers are still the majorities in economics and other fields, with the only exception of education, suggesting gender disparity among disciplines is still an issue. Contrary to the strong international collaboration-oriented tendency in social medicine, economics, and other fields, educational authors are not open to international collaborating. Finally, most of the publications in economics belong to the group with JIF from 0 to 2, in contrast with the high number of social medicine publications with JIF from 2 to 5, which suggesting the field of economics is fulfilling the quantity, but still, need more quality publications.

This article still has certain limitations. Firstly, the analysis exclusively focuses on scientific productivity of social sciences and humanities in Vietnam, which limit the interest to Vietnamese scientists, science and education policymakers, and those who are researching Vietnamese science and education only. Moreover, the primary method of analysis is descriptive statistics. Future studies can employ more sophisticated methods to explore questions that relate to each field in Vietnamese social sciences and humanities.

References:

Adams, J., Pendlebury, D., Rogers, G., & Szomszor, M. (2019). Global Research Report – South and East Asia. *Global Research Report*: Institute for Scientific Information.

Donnelly, K., McKenzie, C. R. M., & Müller-Trede, J. (2019). Do Publications in Low-Impact Journals Help or Hurt a CV? *Journal of Experimental Psychology*. (article in press)

Fisher, B. S., Cobane, C. T., Vander Ven, T. M., & Cullen, F. T. (2013). How Many Authors Does It Take to Publish an Article? Trends and Patterns in Political Science. *PS: Political Science & Politics, 31*(4), 847-856, doi:10.2307/420730.

Henriksen, D. (2016). The rise in co-authorship in the social sciences (1980–2013). [journal article]. *Scientometrics, 107*(2), 455-476, doi:10.1007/s11192-016-1849-x.

Hien, P. D. (2010). A comparative study of research capabilities of East Asian countries and implications for Vietnam. [journal article]. *Higher Education, 60*(6), 615-625, doi:10.1007/s10734-010-9319-5.

Ho, T., Nguyen, H., Vuong, T., Dam, Q., Pham, H., & Vuong, Q. (2017). Exploring Vietnamese co-authorship patterns in social sciences with basic network measures of 2008-2017 Scopus data. *F1000Research, 6,* 1559, doi:10.12688/f1000research.12404.1.

Ho, T. M., Nguyen, H. K. T., Vuong, T.-T., & Vuong, Q.-H. (2017). On the Sustainability of Co-Authoring Behaviors in Vietnamese Social Sciences: A Preliminary Analysis of Network Data. *Sustainability, 9*(11), 2142.

La, V.-P., & Vuong, Q.-H. (2019). bayesvl: Visually Learning the Graphical Structure of Bayesian Networks and Performing MCMC with 'Stan'. The Comprehensive R Archive Network (CRAN). v.0.8.5.

Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. R. (2013). Bibliometrics: Global gender disparities in science. *Nature, 504*(7479), 211–213, doi:10.1038/504211a.

Le, T.-H. T., Pham, H.-H., La, V.-P., & Vuong, Q.-H. (2019). The faster-growing fields. In Q.-H. Vuong, & T. Tran (Eds.), *The Vietnamese Social Sciences at a Fork in the Road* (pp. 52–79). Warsaw, Poland: De Gruyter.

Manh, H. D. (2015). Scientific publications in Vietnam, as seen from Scopus during 1996–2013. [journal article]. *Scientometrics, 105*(1), 83-95, doi:10.1007/s11192-015-1655-x.

Nafosted (2018). *Quỹ Phát triển Khoa học và Công nghệ Quốc gia: 10 Năm Hình thành và Phát triển 2008 - 2018 [National Foundation for Science and Technology Development: 10 Years of Foundation and Development 2008 - 2010]*. Hanoi, Vietnam: NXB Khoa học và Kỹ thuật.

Nguyen, H.-K. T., Nguyen, T.-H. T., Ho, M.-T., Ho, M.-T., & Vuong, Q.-H. (2019). Scientific publishing: the point of no return. In Q.-H. Vuong, & T. Tran (Eds.), *The Vietnamese Social Sciences at a Fork in the Road* (pp. 143–162). Warsaw, Poland: De Gruyter.

Nguyen, T.-T., La, V.-P., Ho, M.-T., & Nguyen, H. K. T. (2019). Scientific publishing: a slow but steady rise. In Q.-H. Vuong, & T. Tran (Eds.), *The Vietnamese Social Sciences at a Fork in the Road* (pp. 33-51). Warsaw, Poland: De Gruyter.

Nguyen, T. V., Ho-Le, T. P., & Le, U. V. (2017). International collaboration in scientific research in Vietnam: an analysis of patterns and impact. [journal article]. *Scientometrics, 110*(2), 1035-1051, doi:10.1007/s11192-016-2201-1.

Nguyen, T. V., & Pham, L. T. (2011). Scientific output and its relationship to knowledge economy: an analysis of ASEAN countries. *Scientometrics, 89*(1), 107-117, doi:10.1007/s11192-011-0446-2.

Nhan, T. (2017). NAFOSTED đầu tư cho nhóm nghiên cứu mạnh [NAFOSTED funds strong research teams]. http://tiasang.com.vn/-khoa-hoc-cong-nghe/NAFOSTED-dau-tu-cho-nhom-nghien-cuu-manh--10876. Accessed November 15, 2019.

Pham, L. T., & Hayden, M. (2019). Research In Vietnam: The Experience Of The Humanities And Social Sciences. *Journal of International and Comparative Education (JICE)*, 27-40%@ 2232-1802, doi:10.14425/jice.2019.8.1.27.

Phuong, T. T., Duong, H. B., & McLean, G. N. (2015). Faculty development in Southeast Asian higher education: a review of literature. *Asia Pacific Education Review, 16*(1), 107-117, doi:10.1007/s12564-015-9353-1.

Price, J. H., Dake, J. A., & Oden, L. (2000). Authorship of Health Education Articles: Guests, Ghosts, and Trends. *American Journal of Health Behavior, 24*(4), 290-299, doi:10.5993/AJHB.24.4.5.

Salmi, J., & Pham, I. T. (2019). Academic Governance and Leadership in Vietnam: Trends and Challenges. *Journal of International and Comparative Education (JICE),* 103-118%@ 2232-1802, doi:10.14425/jice.2019.8.2.103.

Tran, T., Trinh, P.-T. T., Vuong, T.-T., & Pham, H.-H. (2019). The debates and the long-awaited reform. In Q.-H. Vuong, & T. Tran (Eds.), *The Vietnamese Social Sciences at a Fork in the Road* (pp. 17–32). Warsaw, Poland: De Gruyter.

Vu, T.-H., Tran, T., Hoang, P.-H., & Nguyen, M.-H. (2019). Economics: The trend-setting field. In Q.-H. Vuong, & T. Tran (Eds.), *The Vietnamese Social Sciences at a Fork in the Road* (pp. 80–97). Warsaw, Poland: De Gruyter.

Vuong, Q.-H. (2017). Open data, open review, and open dialogue in making social sciences plausible. http://blogs.nature.com/ scientificdata/2017/12/12/authors-corner-open-data-open-review-and-open-dialogue-in-making-social-sciences-plausible/. Accessed November 12, 2019.

Vuong, Q.-H. (2018). "How did researchers get it so wrong?" The acute problem of plagiarism in Vietnamese social sciences and humanities. *European Science Editing, 43*(3), 56-58, doi:10.20316/ese.2018.44.18003.

Vuong, Q.-H. (2019a). Breaking barriers in publishing demands a proactive attitude. *Nature Human Behaviour, 3*(10), 1034-1034, doi:10.1038/s41562-019-0667-6.

Vuong, Q.-H. (2019b). The harsh world of publishing in emerging regions and implications for editors and publishers: The case of Vietnam. *Learned Publishing, 32*(4), 314-324, doi:10.1002/leap.1255.

Vuong, Q.-H., Ho, M.-T., Vuong, T.-T., Napier, N. K., Pham, H.-H., & Nguyen, V.-H. (2017). Gender, age, research experience, leading role, and academic productivity of Vietnamese researchers in the social sciences and humanities: exploring a 2008-2017 Scopus dataset. *European Science Editing, 43*(3), 51-55, doi:10.20316/ESE.2017.43.006.

Vuong, Q.-H., La, V.-P., Vuong, T.-T., Ho, M.-T., Nguyen, H.-K. T., Nguyen, V.-H., et al. (2018a). An open database of productivity in Vietnam's social sciences and humanities for public use. [Data Descriptor]. *Scientific Data, 5*, 180188, doi:10.1038/sdata.2018.188.

Vuong, Q.-H., Napier, N. K., Ho, T. M., Nguyen, V. H., Vuong, T.-T., Pham, H. H., et al. (2018b). Effects of work environment and collaboration on research productivity in Vietnamese social sciences: evidence from 2008 to 2017 Scopus data. *Studies in Higher Education*, (article in press), doi:10.1080/03075079.2018.1479845.

Vuong, T.-T., Nguyen, H. K. T., Ho, T. M., Ho, T. M., & Vuong, Q.-H. (2018). The (In)Significance of Socio-Demographic Factors as Possible Determinants of Vietnamese Social Scientists' Contribution-Adjusted Productivity: Preliminary Results from 2008–2017 Scopus Data. *Societies, 8*(1), 3.

Wagner, C. S. (2006). International collaboration in science and technology: promises and pitfalls. In L. Box, & R. Engelhard (Eds.), *Science and Technology Policy for Development, Dialogues at the Interface*. London, UK: Anthem Press.

Wagner, C. S., Brahmakulam, I., Jackson, B., Wong, A., & Yoda, T. (2001). Science and Technology Collaboration: Building Capacity in Developing Countries? Santa Monica, United States: RAND Science and Technology.