

HOMOPHILY AMONG RESEARCHERS: AN EMPIRICAL STUDY

Sameer Kumar
Asia-Europe Institute, University of Malaya, Kuala Lumpur, Malaysia
Email: sameer@um.edu.my

Abstract

Homophily or assortativity is a phenomenon where entities show preference to associate with others due to some kind of similarity. Using an online questionnaire, we enquired with a group of researchers (n = 112) and asked them if they had showed preference to associate with other researchers while co-authoring a paper on the basis of socio academic characteristics such as nationality, gender or professional position. The results confirm our assumption. We found that, indeed, the researchers, in varying proportions (depending on the parameter) do show preferences in associating with other researchers based on these characteristics. However, several authors also commented that preferences had no place in academia.

Keywords: Homophily, Association, Preference, Socio Academic Characteristics.

1. Introduction

Homophily is the tendency to associate with someone due to some kind of similarity (McPherson, et al., 2001). A common phrase, “birds of same feather, flock together” sums up this idea. Homophily is a social characteristic that spans through various species on the Planet – from birds to human beings. We would have often noticed people talking in groups who may belong to the same nationality or speaking the same language. Researchers form a resilient community and it is generally anticipated that researchers would always give priority to other researchers’ academic talents and credentials over anything else when associating with them. Hence, we wanted to see if this interesting phenomenon also applies to the scientific community, i.e. whether scientists also associate (or more particularly co-author research articles) with one another based on some kind of similarity, i.e. nationality, ethnicity, professional position, etc.

Research Collaboration is a building block of research nowadays. Benefits of collaboration are numerous – from saving time in doing a research (as works gets divided among co-authors, thus reducing time in the completion of a project), to the utilization of expertise and resources at the disposal of co-authors (Beaver, 2001; Katz & Martin, 1997; Presser, 1980). Hence, it comes no surprise when authors look out for partners who they could work together on a research project. Unlike joint research done in large labs by teams (or ‘big science’), a significant proportion of collaboration happens in the social domain where author chooses who he or she would like to collaborate with. Expertise and talent remain the formal pre-requisites for such collaboration. What are the informal parameters that play a role in this decision? Whether researcher show preferences due to certain similarities such as nationality, gender or professional position? With this research objective in mind, we set out to carry out the present study.

More specifically, the research question for the study is:

Whether authors show preferences when choosing their co-authors due socio-academic factors?

The results from the study are significant as it would shed light on the hitherto (and probably taboo area) of informal choices being made by researchers when associating for serious roles such as doing research and writing a research paper together.

In the next section, we delineate the data collection technique and research method applied in carrying out the study. We then discuss the results and state our conclusions in the last section.

2. Material and Methods

Only 'Articles' artefacts were skimmed between 2003 and 2010 from the Web of Science that had 'Economics' as the subject category and had at least one south-east Asian nation in the address section of the bibliography. A total of 1024 email IDs are garnered. Using an online questionnaire individual emails were sent out to researchers requesting them to participate in the study. The email letter contained the link to the online questionnaire which they could click and participate in the survey straightaway. The Duration from 2003 was selected as e-mail ids in the bibliometric records were mentioned more consistently from this year onwards. Before 2003, the email IDs in the author information was only mentioned in a few records and that, too, sporadically.

Participation in the online survey was voluntary, and respondents could choose not to answer question(s) that he or she was not comfortable answering; they could also discontinue participation in the study at any time. The first part of the questionnaire asked for specific information about the researcher (such as age, professional position, etc.) and the other sections had questions including those about their preferences when associating with other researchers. A section was also provided for researchers to provide any other comment about collaboration or work preferences.

The first batch of e-mails were sent to 498 respondents and the subsets of first 50 completed questionnaire responses received an internal reliability rating of Cronbach's alpha of 0.74. Rest of the email requests were then sent out. One e-mail reminder was sent to those who had not yet responded. While 111 responses were received directly via an online google database, one respondent sent his response through e-mail. A total of 342 emails bounced due to email ID failure. Out of the 682 remaining e-mail IDS, we received 112 responses giving a response rate of 15.85%. This response rate seemingly appears to be low, however, a similar study related to international scientific collaboration among Turkish scientists (Aytac, 2010) had a response rate of 8.76%. Until recently, response rate was viewed as an indicator of survey quality. However, rigorous studies conducted in recent years have challenged the notion that a lower response rate means lower survey accuracy. A 1996 survey Visser, et al., (1996) found that a survey with lower response rates (near 20%) yielded more accurate results than did surveys with higher response rates (60%- 70%). Yet another study by Keeter, et al., (2006) found that 77 of the 84 comparisons between Pew Research Center's methodology (usual response rate 25%) and a more rigorous survey conducted over a much longer field period (achieving a response rate of 50%) yielded statistically indistinguishable results. Sheehan, (2001), in a review of e-mail surveys, has reported seven different studies with response rates between 6% and 20%. In survey research, response representativeness is more important than response rate (Cook, et al., 2000).

Descriptive analysis was used. In the tables, valid (n) refers to the number of respondents who chose to respond to this question option. These researchers ($n = 112$) came from 89 institutions representing 28 countries (see table 1). A large number were from south-east Asian countries (ASEAN), which was expected given the fact that the dataset is ASEAN-centric. Malaysia had the maximum number of respondents (15), followed by Thailand (12) and Singapore (9). Respondents were also from large countries like USA, Canada, Australia and also from lesser known countries like Nepal.

Australia	7	Germany	4	Malaysia	1	Norway	1	Sweden	1	USA	1
ANU	2	Heidelberg U	1	UKM	5	U Oslo	1	Linkoping U	1	Morgan State U	1
Monash U	1	Hohenheim U	1	Valuecap S/B	3	Philippines	7	Taiwan	1	Oklahoma State U	1
Murdoch U	1	U Kiel	2	Multimedia U	1	ADB	1	Taiwan Pharmacist Association	1	Pennsylvania State U	1
U Queensland	2	India	1	USM	2	De La Salle U	1	Thailand	1	U Chicago	1
U Western Australia	1	South Asia U	1	U Nottingham Malaysia Campus	2	IRRI	1	Assumption U	2	U Alabama	1
Austria	1	Indonesia	3	UPM	1	PIDS	1	Chulalongkorn U	2	U Michigan-Dearborn	1
U Natural Resources and Life Sciences	1	ERIA	1	UTAR (Perak)	2	U Santo Tomas	1	Kasetsart U	2	U Notre Dame	1
Bangladesh	1	Indo Inst of Sc	1	U Sultan Zainal Abidin	1	U Philippines Los Baños	2	Mahidol U	1	U Pittsburgh	1
BRAC	1	WHO	1	WorldFishCenter	1	Singapore	9	Naresuan U	2	U Texas at San Antonio	1
Canada	2	Israel	1	Mexico	1	ASEAN+3 Macroeconomic Research Office	1	Prince of Songkla U	1	U Wisconsin Green Bay	1
Simon Fraser U	1	Hebrew U	1	ITAM	1	NTU	3	Thammasat U	1	Texas A&M U	1
U Ottawa	1	Italy	1	Nepal	1	NUS	4	Chiang Mai U	1	U Wisconsin Madison	1
Czech	1	FAO	1	Tribhuvan U	1	SMU	1	U of the Thai Chamber of Commerce	1	Vietnam	3
Silesian U	1	Japan	5	Netherlands	4	Spain	2	UK	8	Hue U	1
France	3	GRIPS	1	Erasmus U	1	CEMFI	1	Imperial College London	1	Inst LaborSc Social Affairs	1
CIRAD	1	ICSEAD	1	U Groningen	3	U Pompeu Fabra	1	Plymouth U	1	CFVG - European Excellence in	1

									MgmtEdn
Toulouse Sch Eco	1	Intl U Japan	1	New Zealand	2	Sri Lanka	1	U Portsmouth	1
U Limoges	1	Osaka Prefecture U	1	U Auckland	1	U Kelaniya	1	U Aberdeen	1
		Ritsumeikan Asia Pacific U	1	U Canterbury	1			U Birmingham	1
								U Nottingham	2
								U Sussex	1



A large proportion of the respondents held a PhD degree and was aged over 35 years and were married (see Table 2). The male to female ratio was almost three to one. Those aged between 35 – 45 years of age formed the largest respondent group. About 35% of the respondents were already Professors. I believe this is one of the most diverse samples of respondents for a survey on co-authorship associations.

Table 2: Descriptive statistics of the sample (n= 112)

Descriptives	Valid (n)*	Frequency	Valid Percent	
Gender	111	Male	83	74.8
		Female	28	25.2
Age	112	less than 35	12	10.7
		35-45	43	38.4
		46-55	31	27.7
		56 and above	26	23.2
Marital Status	106	Single	16	15.1
		Married	86	81.1
		Divorced	1	0.9
		Widow/er	1	0.9
		Others	2	1.9
Type of Institution	110	University Educational Institution (non-university)	92	83.6
		Research Institute (non-university)	9	8.2
		Organization	2	1.8
			7	6.4
No. of years of service in current institution	112	less than 1 year	3	2.7
		1-5 years	28	25
		6 - 10 years	25	22.3
		more than 10 years	56	50
Highest qualification	112	Masters	11	9.8
		Doctorate (PhD)	93	83
		Doctorate (others)	2	1.8
		Doctorate with PostDoc	6	5.4
Professional Position	105	Lecturer	4	3.8
		Senior Lecturer	13	12.4
		Asstt. Professor	7	6.7
		Associate Professor	31	29.5
		Professor	36	34.3
		Researcher	10	9.5
		Economist	2	1.9
		Others	2	1.9

* number of respondents who choose to respond to this question option.

3. Results and Discussion

Assortativity in academia is a contentious issue. Most researchers would silently acknowledge the fact that it exists, but few are ready to formally admit it. While putting out this survey, the author was well aware of this fact. However, since the survey was anonymous, there was hope that some picture would be revealed.

60% of respondents in this study revealed that they had shown a preference based on nationality at least sometimes (includes ‘most of the time’ and ‘every time’). With gender and ethnicity, this is again a sizable 40%. In fact, about 20% of respondents revealed that they had shown nationality preference, either ‘every time’ or ‘most of the time’. This is, to a certain extent, a revelation.

However, it must be noted that certain preferences may be due to circumstances rather than choices. As a respondent noted:

“On the final question, this more reflects the fact that most senior researchers in my field are males of European ethnicity, rather than a preference on my part for working with others of the same gender and ethnicity”.

We asked researchers if they preferred to link with someone well known in the field and, if so, to what degree. The survey found that about 50% of authors preferred to link with someone, either ‘always’ or ‘most of the time’, who is well-known in their field. New Researchers, in any research field, preferred to attach themselves to well-known people in the field. In fact, the very basis for the growth of networks (here a community of researchers) is, in part, preferential attachment (Albert & Barabasi, 2002). There was more affinity towards preferring to do work within one’s discipline (intra) than to do multidisciplinary work (see Table 3).

Table 3: Preference to co-author with other researchers based on socio-academic parameters

Preference	Demographics			Field of research			Professional position		
	Nationality	Gender	Ethnicity	Well-known in my field	In my field of research	Not in my field of research	Equal academic rank	From university/institution department	juniors /students
Valid (n)*	104	103	101	106	107	106	104	102	106
Always	3	1	2	11	18	3	2	1	8
	2.88%	0.97%	1.98%	10.38%	16.82%	2.83%	1.92%	0.98	7.55
Most of the time	17	6	4	41	67	17	15	28	27
	16.35%	5.83%	3.96%	38.68%	62.62%	16.04%	14.42%	27.45%	25.47%
Some times	44	35	31	45	20	69	49	58	53
	42.31	33.9	30.69	42.4	18.69%	65.09%	47.12%	56.86%	50.00%

Never	%	8%	%	5%					
	40	61	64	9	2	17	38	15	18
	38.46	59.2	63.37	8.49	1.87%	16.04%	36.54%	14.71%	16.98%
	%	2%	%	%					

When asked about their preference for collaboration based on professional position, again, a good percentage showed this preference. Over 30% revealed that they preferred to work with their juniors/students, 'always' or 'most of the time'. These results reveal that authors do indeed have preferences, albeit with a smaller overall percentage ('always' and 'most of the time'), when co-authoring a paper.

The fact that homophily, due to nationality or gender, can happen due to personal preference, some authors contend that association for a certain choice, i.e. gender or nationality, could happen as it may be the only option or they bring more resources on the table. As a female researcher from an Austrian University comments:

I think there is a large difference between co-authoring a paper with a colleague (similar rank) and with a PhD student. When writing a paper with a colleague, usually one person (the first author) writes a full draft, the others comment on it and make a few changes. When writing a paper with a PhD student, the student writes the first draft, but it usually needs massive changes before submission. However, several authors did suggest that preferences had no place in academia.

Conclusion

The study used empirical data to gauge homophily among researchers due to socio-academic factors. Our findings show that indeed homophily does happen in academia. Homophily due to subject expertise is expected and in our study too this is demonstrated. However, the fact that homophily due to, i.e. nationality, is as high as 60% is significant enough to acknowledge the fact the demographic-based homophily does take place in academia.

Some recent studies have looked into the assortativity of scholars based on socio academic parameters using bibliometric data. However, the current study is perhaps one of the first studies that gauge the perception of assortativity through first-hand responses.

Acknowledgement

The study is supported by University of Malaya, project number: RPO20D-14AFR.

References

- i. Albert, R. & Barabasi, A., 2002. Statistical mechanics of complex networks. *Review of Modern Physics*, 74(1), pp. 47-97.
- ii. Aytac, S., 2010. International scholarly collaboration in science, technology and Medicine and social science of Turkish Scientists. *The International Information and Library Review*, 42(4), pp. 227-241.
- iii. Beaver, D., 2001. Reflections on scientific collaboration, (and its study): past, present, and future. *Scientometrics*, 52(3), pp. 365-277.
- iv. Cook, C., Heath, F. & Thompson, R., 2000. A meta-analysis of response rate in web-or-internet based surveys. *Educational and Psychological Measurement*, 60(6), pp. 821-836.
- v. Katz, J. & Martin, B., 1997. What is Reseach Collaboration?. *Research Policy*, 26(1), pp. 1-18.
- vi. Keeter, S. et al., 2006. . Gauging the impact of growing nonresponse on estimates from a national RDD telephone survey. *Public Opinion Quarterly*, 70(5), pp. 759-779.
- vii. McPherson, M., Smith-Lovin & Cook, J., 2001. Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1), pp. 415-444.
- viii. Pressor, S., 1980. Collaboration and the quality of research. *Social Studies of Science*, 10(1), pp. 95-101.
- ix. Sheehan, K., 2001. Email survey response rates: A review. *Journal of Computer-mediated Communication*, 6(2).
- x. Visser, P., Krosnick, J., Marquette, J. & Curtin, M., 1996. Mail surveys for election forecasting?. *Public Opinion Quarterly*, 60(2), pp. 181-227.

