CEB Working Paper

Centre Emile Bernheim
Research Institute in Management Sciences

Solvay Brussels School
Economics & Management

# Social Network Sustainability Metrics: A Study of Co-authoring Behaviors in the Social Sciences, Using 2008-2017 Scopus Data for Vietnam

**Tung Manh Ho, Hong Kong Nguyen-To, Thu-Trang Vuong and Quan-Hoang Vuong**

The study examines the co-authoring behaviors of 412 Vietnamese social scientists over the 2008-2017 period via a new method – social network analysis – to determine if these researchers have formed sustainable scientific communities, using Scopus data. The dataset provides an insightful look into the predominant form of collaboration, i.e., co-authorship, within the Vietnamese social science research communities. Through basic network metrics such as density and clustering coefficient, the study hypothesizes that the socially sustainable research communities are those with low clustering and high density. As any scholar's position in a network can be specified by three quantities: number of publications, connections, and years in research, the distance metrics from the most productive to the rest are computed and compared. The study hypothesizes that if the distance is too large; it reflects the socially unsustainable situation in the network. The results indicate that certain level of social unsustainability exists in social sciences groups in Vietnam. Though the results are only indicative, it has opened up a fertile space for future enquiry into this matter.

Keywords: publishing behavior, co-authoring behavior, sustainable scientific communities, social sustainability distance, social network analysis

JEL Classifications: D85, D91, Q01

CEB Working Paper N° 17/027
October 2017

*Université Libre de Bruxelles - Solvay Brussels School of Economics and Management*
*Centre Emile Bernheim*
*ULB CP114/03 50, avenue F.D. Roosevelt 1050 Brussels BELGIUM*
*e-mail: ceb@admin.ulb.ac.be Tel.: +32 (0)2/650.48.64 Fax: +32 (0)2/650.41.88*

ULB

# Social Network Sustainability Metrics: A Study of Co-authoring Behaviors in the Social Sciences, Using 2008-2017 Scopus Data for Vietnam

**Tung Manh Ho [1, 2], Hong Kong Nguyen-To [3], Thu-Trang Vuong [4], Quan-Hoang Vuong [2, 5, *]**

[1] Institute of Philosophy, Vietnam Academy of Social Sciences, Hanoi, 100000, Vietnam
[2] Centre for Interdisciplinary Social Research, Western University Hanoi (ĐH Thành Tây), Hanoi, Vietnam; tung.ho@wu.edu.vn (T.M.H), hoang.vuong@wu.edu.vn (Q.-H.V)
[3] Vietnam Panorama Media Monitoring, Hanoi 100000, Vietnam; hongkong.tnguyen@gmail.com
[4] Sciences Po Paris, Campus de Dijon, 21000 Dijon, France; thutrang.vuong@sciencespo.fr
[5] Université Libre de Bruxelles, Centre Emile Bernheim, Brussels 1050, Belgium; qvuong@ulb.ac.be
[*] *Corresponding author*.

**Abstract:** The study examines the co-authoring behaviors of 412 Vietnamese social scientists over the 2008-2017 period via a new method – social network analysis – to determine if these researchers have formed sustainable scientific communities, using Scopus data. The dataset provides an insightful look into the predominant form of collaboration, i.e., co-authorship, within the Vietnamese social science research communities. Through basic network metrics such as density and clustering coefficient, the study hypothesizes that the socially sustainable research communities are those with low clustering and high density. As any scholar's position in a network can be specified by three quantities: number of publications, connections, and years in research, the distance metrics from the most productive to the rest are computed and compared. The study hypothesizes that if the distance is too large; it reflects the socially unsustainable situation in the network. The results indicate that certain level of social unsustainability exists in social sciences groups in Vietnam. Though the results are only indicative, it has opened up a fertile space for future enquiry into this matter.

**Keywords:** publishing behavior, co-authoring behavior, sustainable scientific communities, social sustainability distance, social network analysis

**JEL Classification:** D85, D91, Q01,

---

*This is an unpublished working manuscript*

## 1. Introduction

In April 2017, Vietnam's Ministry of Education and Training (MOET) issued Circular No. 08/2017/TT-BGDDT [1] in which all doctoral students are required to have papers published in Scopus and Web of Science-indexed journals, and Ph.D. supervisors to have at least one international publication. Meanwhile, in the public sphere, everyone is paying special attention to the debate on the proposal to reform the standard of the highest academic titles in Vietnam: Professor and Associate Professor. In the debate, the central issue is about the role of international publications in judging whether a Vietnamese academic is worthy of these honorary titles [2, 3].

While this contentious issue is raging on, Vietnam National Foundation for Science and Technology's Deputy Director Dr. Pham Dinh Nguyen publicly restated that the foundation's policy is to focus their investment primarily for strong research groups so as to promote high quality research [4]; the details of the policy is listed in Degree No. 37/2014/TT-BKHCN [5] issued by the Ministry of Science and Technology. These efforts are the governmental response to the public outcry that Vietnam's scientific education and research is suffering in both quality and quantity. In a report released by the *Times Higher Education* earlier in 2017, there was no Vietnamese university in the top 300 in Asia; according to many experts, this is the direct result of a lack of focus on research capacity of Vietnamese universities [6-7]. Furthermore, several popular stories in the Vietnamese press have exposed a low quantity of international publications of Vietnamese researchers and research institutions, especially in social sciences and humanities. For example, many have questioned the Vietnam Academy of Social Sciences (VASS), which is touted as one of the top governmental research institutions, for spending over USD 90 million in five years until 2016 only to publish 22 Web-of-Science-indexed papers in total [8-9]. In this context, it is clear that change is inevitable to create more internationally integrated, productive and sustainable research communities in Vietnam. In an online periodical of Vietnam's Ministry of Science and Technology, the Director of Vietnam National University of Social Sciences and Humanities, Mr. Pham Quang Minh, claimed that social sciences and humanities research in Vietnam are out of touch with the reality of the world and called for a systematic reformation [10]. The end goal is clear, yet, how to embark is uncertain for both the public and policy-makers.

The main reason for such uncertainty is the lack of actionable hard data and insights on international publication of Vietnamese researchers in social sciences and humanities. There have been only a few quantitative studies on scientific publication in Vietnam in general while collection of data on the situation of international publications of Vietnamese social scientists is still in a nascent stage. One study shows that in South East Asia, total research output in Vietnam is relatively low, with rate only equivalent to 13% of Singapore and 29% of Thailand in the period of 1991-2010 [11-12]. It is notable that the share of international co-authorship took up 77% of the total publications [13-14] and most of the leading authors are not from Vietnam [13]. Regarding social sciences and humanities, a new study shows that there is a correlation with the author ages and number of articles in which they played a leading role, while there is no correlation with their gender in Vietnamese authors [15]. Regarding collaboration trends in Vietnamese social science, it is shown that the number of Vietnamese leaders in research groups is still very small and about 75% of the Vietnamese authors never attempted publishing solo [16]. Another study on co-authorship network among Vietnamese social scientists demonstrates how sparse the connections among these researchers are, though these connections depend heavily on a group of intellectual elite, which consist of well-connected, productive and socially important individuals [17]. The prevalence of international co-authorship and the scattered co-authorship connections among Vietnamese social scientists provide an interesting contrast. As pointed out by Vuong and Napier [18-19], Vuong [20], in a Confucian society such as Vietnam, criticisms of an innovative idea could be seen as a personal distrust in the person who proposes it. It could be the case that collaboration among Vietnamese scientists, who supposedly operate in the frontier of innovation and creativity, might suffer as a result of this cultural burden. In many ways, this background suggests a level of fragility or socially unsustainability in the collaboration among Vietnamese social scientists. Thus, for Vietnamese public and policy-makers to improve the situation of weak research capacity in social sciences, it is of vital importance to understand whether the existing social scientific communities are socially sustainable, and to identify which research groups are sustainable. By collecting and analyzing data using network statistical analysis on co-authorship networks among Vietnamese social scientists who have successfully published in the Scopus-indexed journals, this study aims to provide a novel way to conceptualize the problem of identifying sustainable scientific communities in Vietnam.

## 2. Literature Review

### *2.1. What does sustainability mean?*

Although sustainability and sustainable development are two very popular concepts in recent decades, it is very difficult to pinpoint clear-cut definitions to them. Indeed, for many, sustainability and sustainable development are comparable to democracy for its ambiguity and desirability [21]. The beginning of these concepts traced back to the *Brundtland Report* (*Our Common Future*), in which sustainable development is defined as the kind of development which ''meets the needs of current generations without compromising the ability of future generation to meet their own needs'' [22]. This definition has received much criticism for its vagueness [23-26]. Many believe that different conceptions of sustainable developments and sustainability are mere reflections of the philosophical and political positions of those advocating the definition; there is no unambiguous scientific way of defining them [26]. However, there are also arguments supporting certain level of ambiguity since it opens up a possibility for flexible negotiations among interested parties [23] as well as helps avoid unhelpful ideological battles related to these concepts [27].

It seems unlikely that sustainability and sustainable development will have a precise definition; nonetheless different dimensions of these concepts seem quite uncontroversial. One of them is the growing awareness of the global society on the interconnectedness between human society and nature: human is an inseparable part of the ecosystem and our well-being depends on a clean, hospitable natural environment. Thus, many argue that our current lifestyles and economic activities are unsustainable, and they would cause ecological catastrophes for the future generations. One famous example is in 1992, about 1,700 scientists of the Union of Concerned Scientists (UCS), including 102 Nobel Laureates, issued an appeal called *World's Scientists' Warning to Humanity*, which detailed the way our current practices, if left alone, could inflict great harms to the environment and human civilizations, and called for a fundamental change [28].

Along with this line of reasoning, there have been several alarming discussions on the untenable relationship between human population growth and resources management, which could lead to the eventual collapse of human civilization. Malthus in his famous *An Essay on Principle of Population* [29], Erhlich in *Population Bomb* [30], and Meadows in *Limits to Growth* [31] all laid out the scenario of economic collapse and mass starvation as overpopulation leads to resources depletion. Others have sought different ways human societies could collapse; for instance, the inability to manage the growing complexity of social organizations as explained by Joseph Tainter in his book: *The Collapse of Complex Societies* [32] or the involvement of harsh environmental factors such as deforestation, water scarcity, habitat destruction, and new species as explained by Jared Diamond in *Collapse: How Societies Choose to Fail or Succeed* [33].

Another dimension is that the case against severe economic inequality, which is referred to in the *Brundtland Report* [22] and Holden *et al.* [24] as promoting intra-generational equity. Though sustainability is a relatively new concept, past prominent thinkers such as Karl Marx have long identified severe economic inequality as a threat toward a stable and thriving society [34]. This debate has recently been revived by current renowned economists such as Stiglitz in *The Price of Inequality* [35] and Piketty in *Capital in the Twenty-first Century* [36]. Scholars across disciplines also pointed out that severe economic inequality is a source for unsustainable future [33, 37-38]. In recent years, with rapid progress in technology, how to manage technological risks has been added as a new dimension to the problem of how to sustain human society long into the future [39-42].

Indeed, however vague sustainability appears to be as a concept, it is truly useful in capturing a wide range of problems which are fundamental to the long-term development of human society. Next, the existing approaches to solve these sustainability problems will be reviewed.

### *2.2. Different approaches in solving sustainability problems*

Given that since its birth, sustainability has attracted much attention from all sectors: policy-makers, scientific communities, and activists, among others, there is no shortage of approach toward solving this problem. This paper groups the approaches into three broad categories: (i) institutional level, (ii) individual level, and (iii) technological fix. While there is considerable overlap among these categories, this classification offers a useful way to understand the landscape of different approaches toward the sustainability problem.

First, for the institutional level, the main focus is on changing conceptual frameworks, policies and behaviors of institutions to be more effective in striking a balance between the environment and securing long-term human survivals as well as satisfying the needs of human society. Elkington suggested the Triple Bottom Line model for businesses to assess and direct their sustainable practices, in which profit as a conventional bottom line is added together with the planet together with the well-being of people [43]. In the case of governmental policy, there have been arguments for reserving sustainability as a concept for only environmental factors, which relates to the future generations, while socio-economic well-being of the present generation can be reflected by the measurements of Gross Domestic Products (GDP) and Human Development Index (HDI) [44]. Others argue that it is possible to think of sustainability as an all-encompassing concept for both socio-economic and environmental dimensions by simply adding Ecological Footprint in the assessment [24]. There are also different suggestions to institutional changes. For example, Max Tegmark, in the subject of managing technological risks, argued for a switch in an institutional level from the learning-from-past-mistakes framework to a more proactive "security engineering" approach [45]; Bettencourt and West [46] made the case for adopting a new quantitative understanding of urban living in policy making as it could help human avoid the "planet of slums," and instead, arrive at "a sustainable, creative, prosperous, urbanized world expressing the best of human spirit."

Second, on the individual level, the focus lies in changing the behaviors and lifestyles of each individual in order to increase global sustainability. For example, the green living movement calls for each individual to be more aware of his or her relationship with the environment and thus adopt a more environmentally friendly lifestyle: using public transports rather than driving, classifying household trashes, boycotting eco-unfriendly products, curbing excessive consumptions, etc. This is widely regarded as a low-cost way for tackling the sustainability problem; however, studies show that policies that encourage eco-friendly lifestyle only succeed in a limited way [47]. Research in this area is still ambiguous in answering how to motivate people to adopt a more eco-friendly lifestyle; there seems to be no direct way to do so. It is also shown to be unnecessary for pro-environment behaviors to be linked to personal concerns or values regarding the environment [48]. In contrast, Whitemarsh and O'Neil in 2010 found out that self-identity can predict significantly pro-environment behaviors; people with green identity are more likely to take part in eco-friendly lifestyle [49]. It is also shown that motives to act pro-environmentally can be rooted on family's values rather than individual ideals [50-51]. In addition, psychological researches have shown when it comes to sustainable choices and behaviors, there is a well-documented gap between action and value [52-54]. Human psychology seems to be naturally prone to many significant mental barriers in adopting a more sustainable lifestyle: limited awareness about the problem, abstractness, sunk costs fallacy, experts' suspicion, fears of change, and denials of the problem, among others [55-57].

Finally, technical fix is a relatively less discussed approach. In this approach, science discovery and technological innovation are at the center of the bull's eye. Deutsch in 2011 discussed the concept of sustainability in the context of how to maintain an open-ended quest for knowledge creation, as he argued for the principle of optimism: "All evils are caused by insufficient knowledge." In this view, there is no way to foresee how new problems could arise from new scientific ideas and solutions to existing problems. Thus, the only sustainable way to organize a society is to nurture rapid progress in science and technology; in this process, a vibrant, creative and productive communities of scientists is necessary [58]. Geoffrey West arrived at the same conclusion

from a different line of reasoning: by applying network theory to mathematically study growth in biological systems, and human systems such as corporations and cities. He found out that it is possible to have an open-ended growth for social systems such as cities under one condition: continuous innovation [59-60]. West called for a grand unified theory of sustainability - a science that is quantitative and predictive - as he saw the same universal law of scaling can be applied for both the natural and human worlds [61].

### *2.3. Social sustainability, scientific communities and social network analysis*

As can be seen from the previous section, whether one chooses to address the problem in a personal or institutional level or by finding a technical fix, the role of creative problem solving and innovation is central to creating a more sustainable future. Hence, it is important to ask how communities of scientists operate and how to identify sustainable communities—the enduring and productive ones. As communities are essentially networks, here, social sustainability and social network analysis can be fruitful concepts to bring to this investigation.

Social sustainability, as a concept, has been applicable in many researches which belong to a wide range of subjects: sharing economy, supply chain, construction projects, community resilience, etc. [63-66]. However, similar to sustainability and sustainable development, social sustainability suffers from the problem of ambiguity con context-dependence, it is almost chaotic to survey the literature review related to this issue [67-68]. There have been many attempts to develop measures and indicators for social sustainability, however, as is the concept, the measures vary greatly and cover a broad range of aspects [69]. As Turcu pointed out the indicators are manifestations of the underlying local perspectives regarding sustainability [70]. In this situation, it is clear that there is an enormous difficulty in pin-pointing the exact definition of social sustainability and in finding the kind of measures that suit one's purpose. Nevertheless, the situation also implies there is enough space and flexibility in this line of research to apply the concept in the most constructive way.

Meanwhile, social network analysis, as a technique, has been employed widely to study the structures of scientific collaboration and dynamics. Newman showed scientific collaboration networks seem to have the "small world" properties when he applied this technique on the data of biomedical research, physics and computer science [71]. By studying citation networks, a study was able to uncover the landscape of sustainability science: there are 15 main research clusters such as rural sociology, tourism, forestry, ecological economics, urban planning, wildlife, etc. [72]. Similarly, Moody (2004) revealed that in sociology, researchers who do quantitative work tended to work with non-quantitative counterparts by studying 30 years' worth of data of sociology collaboration networks [73]. Application of social network analysis is also useful in predicting scientific performance. For example, a group of Taiwanese researchers found that position in a co-authorship networks can help predict citations of publications [74]. In China, a research team also found the same pattern when examining co-authorship network effects toward citation counts using library and information science data [75].

Indeed, there have been many attempts to utilize social sustainability and network analysis together. A study analyzed networks data and data on perceptions revealed that collaboration networks and sustainability perceptions are important in evaluating the implementation of climate change adaptation [76]. Another study applied network analysis in studying the interests of multiple stakeholders in construction projects in Saudi Arabia and concluded that as the needs of diverse stakeholders were satisfied, the situation of social sustainability improved for the construction projects [64]. However, there is no attempt yet at utilizing the two concepts in the context of scientific communities. Thus, this paper will explicate on how useful the technique of network analysis is in providing a quantitative understanding of social sustainability within the social scientific communities in Vietnam.

Standard network measure such as clustering coefficient (transitivity) has been shown to have straightforward mathematical relation to the spread of information in a network—the higher the cluster, the slower information spread [77]. And in sustainable scientific communities, one can reason that information, e.g. data, knowledge, expertise and experiences, should be communicated efficiently. Thus, clustering coefficient would offer an indirect way of assessing the status of social sustainability of research communities. Moreover, in a co-authorship network, a scholar's position can be defined by three types of quantity: the number of publications, connections, and years in research. As a result, it is possible to calculate the distance from the most productive member of a network to any other members of the network. This distance, referred to as "social sustainability distance" (*SSD*), might be meaningful in investigating social sustainability of the scientific network in question.

By collecting attribute data and relational data of 412 Vietnamese social scientists between 2008 and 2017, then constructing a co-authorship network of these scholars, this paper will attempt to answer three research questions:

**RQ1:** Applying social network analysis, what are the broad trends of co-authoring behavior in research communities that exist among 412 Vietnamese social scientists?

**RQ2:** Through clustering coefficient and density, what can be stated on the status of social sustainability in Vietnamese social scientists' communities?

**RQ3:** What does the SSD tell us about the social sustainability of Vietnamese social scientists' communities?

## 3. Materials and Methods

### *3.1. Conceptualization of research questions*

The concept of sustainability in research networks has never been explored, thus, it is important to clearly define what it means to say a research network is sustainable. Normatively speaking, a sustainable research network should: (i) be efficient in transferring scientific knowledge and expertise, (ii) remain well-connected even if a few members of the network are removed, and (iii) be comprised of productive members. Though it can be relatively intuitive and straight-forward to comprehend what would make a research network sustainable, it is hard to quantify the criteria. Here, the technique of social network analysis offers a novel way to solve this problem. Before getting to the criteria expressed in social network analysis language, it is useful to understand the basic terminologies of the technique.

A social network is defined as a collection of individuals, each of whom is related to others by one or more different kinds of relations such as friendship, kinship or co-authorship (Scott, 2017). And network analysis is an emerging field, based on the application of various mathematical tools and methods on the problems related to network. A graph or a network $G = (V, E)$ is a mathematical structure consisting of a set $V$ of vertices (or nodes) and a set $E$ of edges (or links); elements of $E$ are links between a pair of distinct vertices belongs set $V$. In this study, a vertex represents a Vietnamese social scientist. An edge represents a co-authorship connection between two distinct Vietnamese social scientists. The number of edges incident upon a vertex is called a vertex degree. Degree is an important concept in this study because it is used in a measure that helps identify sustainable research networks.

In this paper, we will use the words node, vertex, Vietnamese social scientist, researcher, scholar and scientist interchangeably. Similarly, edges, links, and connections are also equivalent in meaning. The words such as network, graph, research network and co-authorship network are also interchangeable. To see how these seemingly irrelevant concepts can be useful in identifying sustainable research networks, let's consider the way network analysis enables us to quantify the different aspects of a sustainable network.

First, to get a sense of how efficient information in a network is communicated, one can study how many connections there are in a network. This can be measured by the concept of density, the number of realized connections divided by the number of potential connections. One can assume that the higher the density, the more information can flow in a network. However, the measure of density might not reflect the full picture because the higher the number of members of a network, the more potential connections, thus, density can decrease as a result. Hence, it is important to view a network from another angle.

Another angle to get a sense how efficient information can be spread in a network is clustering coefficient, the likelihood that a connected triple will close to form a triangle. In network statistical analysis literature, it is a well-known fact that clustering coefficient has a straight-forward relationship with the speed of information spreading in a network: the higher the clustering, the slower information spread [77]. Thus, in the context of studying the sustainability of a research network, one can assume if a network has low clustering coefficient, it is likely that information can be communicated better in this network.

Through density and clustering coefficient, one can catch a glimpse into how effective scientific knowledge can spread in a network. However, ultimately, one must understand the end result of this process of knowledge dissemination, which is the gap between the most productive researcher or the most connected and others in the network. In a sustainable network, this gap should not be too big. To measure this gap, we develop a measurement for the distance between the most productive member of a research network and the rest in the network.

One can reason that the numbers of connections in a co-authorship network and the number of research years would have an effect on the productivity of a researcher. For example, the longer a person has been doing research, the more likely he or she can increase his or her productivity; or the more people a researcher co-authors with, the more likely he or she gain access to more ideas and knowledge, thus resulting in a gain in productivity. It can be seen that three quantities: number of publications, number of co-authorship links, and research years would constitute a vector that specifies any social scientist in a research network. Hence, one can compute the Euclidean distance from the vector that identifies the most productive scholar to other vectors that pick out any other nodes in a network. Whether this distance is large or small would be a telling sign if a research network is sustainable or not.

We believe these network metrics together will provide a quantitative method to identify sustainable research networks.

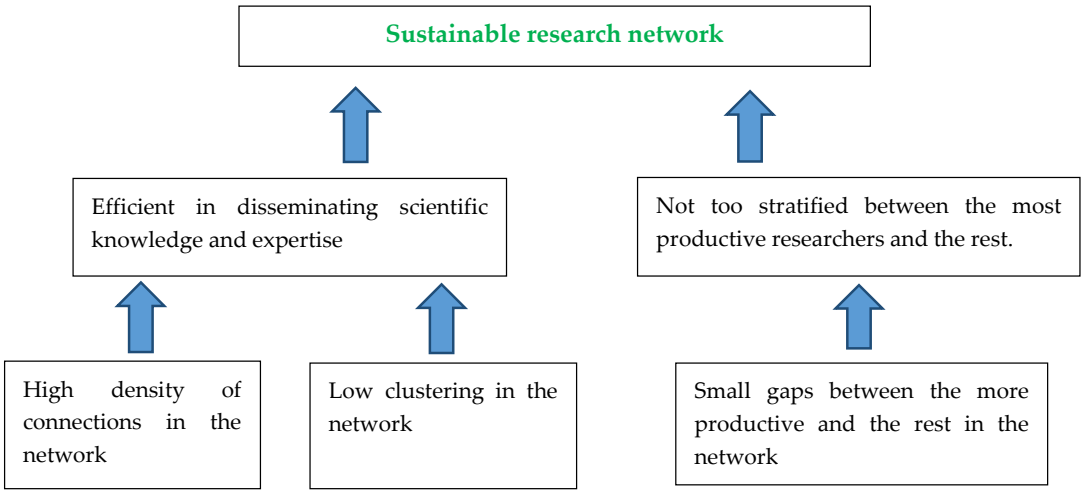



**Figure 1.** Conceptualization of the research questions

### *3.2. Materials: Attribute data and relational data*

The data for this study was derived from a dataset on the productivity of Vietnamese scientists in the field of social sciences and humanities collected by Vuong & Associates. The investigation, which took place within five months from February to July 2017, was conducted under the license V&A/03/2017, issued on 15 March, 2017.

#### 3.2.1. Attribute data

First, the criteria of the subject for data collection are determined. For a researcher to be a legitimate subject of this study, he or she must be:

a. A Vietnamese national
b. Satisfies either or both of these two conditions:
   - Have been affiliated with a Vietnamese institution
   - Have published at least one paper on a social scientific issue related to Vietnam

Being clear on the criteria, we then set out to collect attribute information of the researchers that satisfy the requirements. The result of the data collection process is a complete dataset of 412 scholars' details, consisting of: (i) age, sex, region; (ii) affiliations; (iii) fields of study; (iv) the number of publications in Scopus, (v) the number of research years since the Master graduation; (vi) the number of researchers they collaborated with; (vii) whether or not they have the title of "Professor/Assoc. Professor." To make sure the data is reliable, the research team collected data from various sources such as personal and institutional websites of scholars, journals' websites, Google scholars, and Scopus database. Then by comparing information in these online sources (Google scholar versus Scopus, personal sites versus organizational websites, etc.), we were able to eliminate potential errors. For instance, different versions of a Vietnamese name can result in two IDs for one person, or a person has Vietnamese name but is not a Vietnamese national. In the end, the clean set of data constitutes the "Nodes list" ("20170725_net412_ NODES.csv"), which contains the attribute information of each author.

#### 3.2.2. Relational data

From the data available after making the Nodes list, an "Edges list," ("20170729_net412_LINKS.csv") which contains relational data, is constructed. When two

researchers co-author a paper, they are considered to have a co-authorship link. Every time the same two authors appear together in a paper, it is counted toward the "weight" of the link. Figure 2 shows an example of how the edges list is derived from the original data. As an example, in the first row of the table on the left side, a published paper being co-authored by scientists ID s004, s076 and s079 is registered. Then, on the right side, the co-authorship links among these three scholars are documented; and the weight is the count of how many times each pair co-authors. The data was then processed and analyzed using statistical software R (v3.3.1).

| Title | year | Journal | ID |
|---|---|---|---|
| Does Economic Inequality Affect the Quality of Life of Older People in Rural Vietnam? | 2017 | Journal of Happiness Studies | s076 ;s004; s079 |
| The Effect of Having Children on Women's Marital Status: Evidence From Vietnam | 2017 | Journal of Development Studies | s004; s046 |
| Does firm privatisation benefit local households? The case of Vietnam | 2015 | Post-Communist Economies | s005; s004; s079; s080 |
| Firm agglomeration and local poverty reduction: evidence from an economy in transition | 2016 | Asian-Pacific Economic Literature | s005; s004; s076 |

(**a**)

| | | Weight |
|---|---|---|
| s004 | s076 | 2 |
| s046 | s004 | 1 |
| s079 | s004 | 2 |
| s004 | s080 | 1 |
| s005 | s004 | 2 |

(**b**)

**Figure 2.** An example of the process of deriving relational data from the original dataset. **(a)** The table shows an example of recording article titles and their relevant properties. **(b)** The table describes how relational data is created from table (a) (Adapted from [17])

The data for Net1 to Net20 were manually extracted from the full dataset. Nodes lists and edges list for these 20 networks were built by picking relevant edges and nodes from the original lists; all of which are available in the folder 20 Networks' Data. All the data sets are given in [81].

### *3.3. Method of analysis*

There are several reasons why we choose the method of statistical analysis of social network data for this study. First, we are naturally inclined to wonder what kind of interactions occurs among Vietnamese social scientists given the pervasiveness of co-authorship [16]. This leads us to expect that social network analysis would be a match with our interest to achieve a holistic and quantitative view of the characteristics of the co-authoring behavior of Vietnamese social scientists: how densely co-authoring connections occur, how clustering would affect the productivity of a co-authoring network, how socially sustainable it is. Another important aspect is the possibility of visualization of research networks. With the support of statistical software, graphic representations of research networks could be created; hence, not only we could learn from all the rigorous numerical analysis but we could also achieve a more intuitive understanding of the interactions among actors in a co-authoring network. It is indeed not difficult to see the advantages the method offers given our research questions.

As this study is strictly limited to the co-authoring behavior of Vietnamese scholars only, besides the advantages, there are a few caveats regarding the scope of analysis. First, it is likely that some features of research networks could be lost when the co-authorship connections with foreign scholars are not taken into account. To illustrate, a foreign researcher might co-author with a few Vietnamese scholars, but these Vietnamese might not publish together. This will result in some number of missing links. The aggregate effects of this phenomenon can make the network appear sparser and less clustering than in reality. Second, the history of network analysis began with problems in fields such as mathematics, chemistry, electrical circuits, operational research, and computer science before they are applied to study network of people [78]; consequently, the technique is not fully developed and matured. It is wise to keep in mind that there might well be

inherent shortcomings to the explanatory power of the technique. However, considering both *pros* and *cons*, we believe social network analysis's advantages out-weights its limitations, in conjunction with other types of techniques such as [77-79]. Next, we will turn to the formulas to calculating the network metrics that would help us to decipher the overall characteristics of co-authoring behavior among Vietnamese social scientists.

#### 3.3.1. Standard network measures

The formula to compute density of a network is:

$$density = 2l/[n(n\text{-}1)] \quad (1)$$

where $l$ is the number of edges/links exists in a network, and $n$ the number of nodes in the network. The formula to compute clustering coefficient is:

$$cl_{\tau}(G) = 3\tau\Delta(G)/\tau_3(G) \quad (2)$$

in which $\tau\Delta(G)$ is the number of triangles in the network $G$; and $\tau_3(G)$ the number of connected triples, the subgraphs consist of three vertices connected by two edges. As shown in the literature review and the conceptualization of the research question, low clustering coefficient would suggest knowledge in a research network is disseminated well.

#### 3.3.2. A proposed measure of social sustainability distance

As mentioned above, in a sustainable research network, the gap between the most productive researchers should not be too big. And we noticed that a researcher can be specified using three quantities: number of publications, number of connections in a network, and number of years in research, as provided by equations (1) and (2). Thus, to quantify the gap, we propose a formula to compute the Euclidean distance from a three-dimensional vector that specify the most productive scholar to other vectors that picks out any other nodes in a network:

$$SSD = \sqrt{(x_{\max} - x_0)^2 + (y_{\max} - y_0)^2 + (z_{\max} - z_0)^2}\ . \quad (3)$$

In equation (3), ($x_{max}$, $y_{max}$, $z_{max}$) is a vector represent three attributes of the most productive member in a network: $x_{max}$ stands for the total number of publications, $y_{max}$ stands for the total number of co-authorship connections that researcher possesses, $z_{max}$ stands for his or her research years. ($x_0$, $y_0$, $z_0$) is a vector representing the same three values: publications, connections, and research years that define any author in a research network.

In this study, for each co-authorship network, we compute the distance in five scenarios:

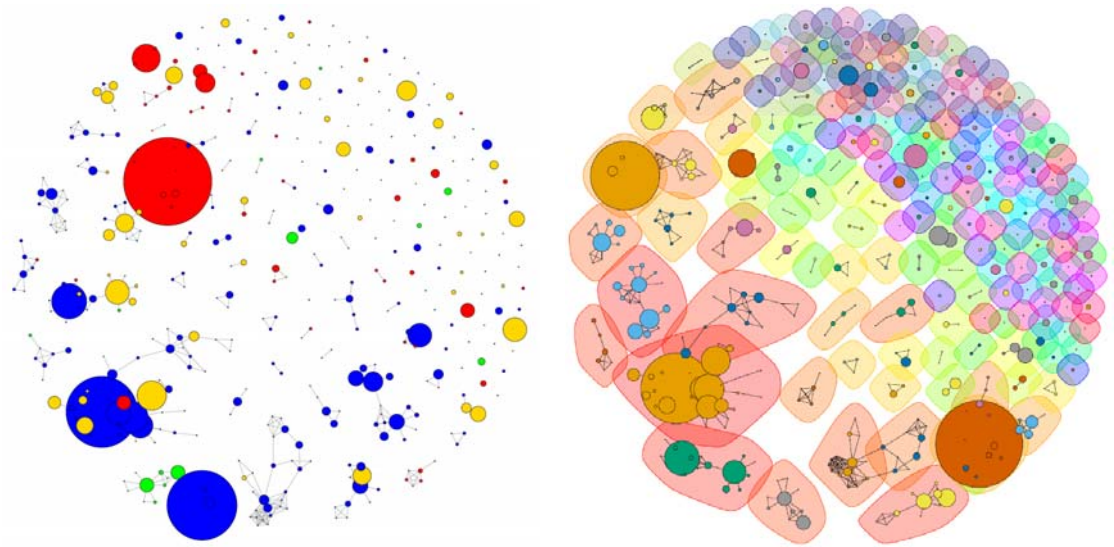
1. most productive researcher to the vector consists of the mean value of all three quantities
2. most productive researcher to the researcher who has the median value of the network in terms of publications
3. most productive researcher to the researcher who is in the min value in terms of publications
4. most productive researcher the researcher who has the second highest number of publications.
5. most productive researcher to the mean value of the low productivity group, those with equal or fewer than three publications

It is hoped that by calculating the *SSD* as a characteristic of a co-authorship network, then comparing this characteristic among different networks, interesting patterns regarding social sustainability of research networks will emerge.

## 4. Results

### *4.1. Standard network metrics and visualization*

First, it would be useful to visualize the co-authorship connections that exist among all 412 Vietnamese social scientists in our sample. All 412 nodes and their edges are plotted on a graph (called Net412) using "iGraph" package in R (v3.1.1), as shown in Figure 3. As node size represents the total articles a researcher has published, interesting patterns emerge. In most of the sub-networks or research communities, the most productive researchers seem to always be central to many different connections. Lying in the top corner of the circle in both figures are 125 researchers who either work alone or work with foreigners. The middle of the circle is occupied by small groups of researchers: size 2-9, which takes up about 40% of the total population. In the lower corner, there reside the large groups, comprising at least 10 members per group. These groups account for about 30% of the population.



(**a**) (**b**)

**Figure 3.** Net412 - A visualization of all 412 Vietnamese social scientists in two modes: (**a**) Node colors based on regions, in which blue is for north, red is for south, green is for center, gold is for overseas; (**b**) Node colors automatically generated by community detection algorithm, which yields 181 communities. There are 125 communities of only one member, the remaining is communities with more than two members. (No. nodes = 412; No. links = 401; Density = 0.0047; Clustering Coefficient = 0.59; Mean degree = 1.95; Mean publications = 3.56; Median publications = 2).

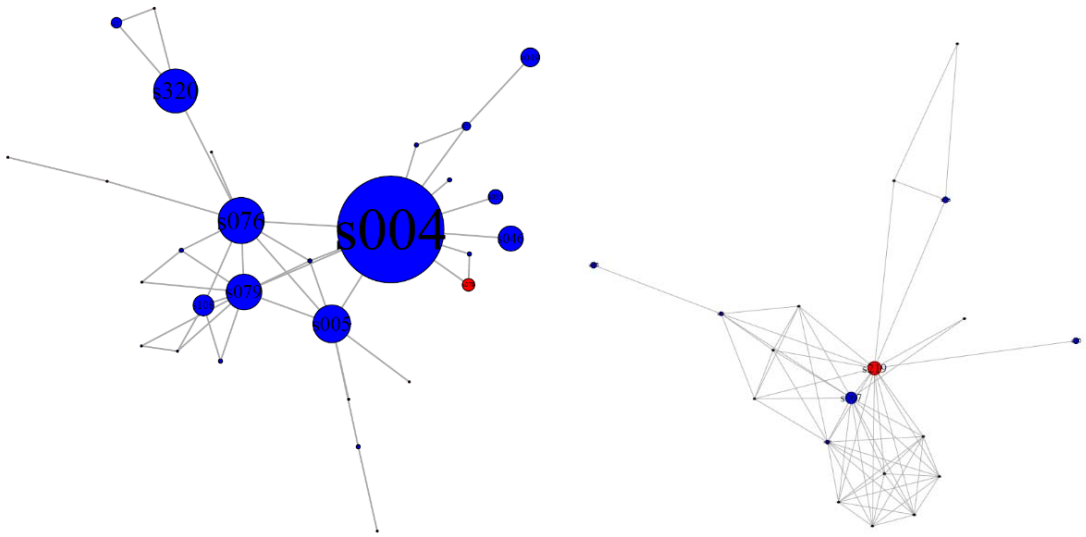
Having glimpsed at the structure of the network of 412 Vietnamese social scientists, one can then examine the question of social sustainability in more fine-grained details by looking at the properties of each scholarly community. Next, 20 communities with at least 5 members each are manually extracted from the total dataset (the data files for these communities can be found in the folder "20 Networks' Data" and "Rcommands and figures for all nets"). Each of these subgroups is considered a network in and of itself. In table 1, the standard network metrics of these networks are summarized.

**Table 1.** 20 (sub)networks of Vietnamese social scientists and their network metrics

| Network | Density | Clustering Coefficient | Mean Degree | Mean Publication | Median Publication | No. Nodes | No. Links |
|---|---|---|---|---|---|---|---|
| Net1 | 0.29 | 0.51 | 2.91 | 3.91 | 2 | 11 | 18 |
| Net2 | 0.31 | 0.44 | 2.80 | 3.90 | 1 | 10 | 14 |
| Net3 | 0.33 | 0.00 | 2.00 | 4.86 | 2 | 7 | 7 |
| Net4 | 0.36 | 0.51 | 3.20 | 3.20 | 1.5 | 10 | 16 |
| Net5 | 0.70 | 0.80 | 2.80 | 4.80 | 6 | 5 | 7 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Net6 | 0.36 | 0.71 | 6.53 | 2.00 | 1 | 19 | 62 |
| Net7 | 0.33 | 0.67 | 3.23 | 2.64 | 1 | 11 | 18 |
| Net8 | 0.11 | 0.30 | 2.97 | 7.10 | 2 | 29 | 43 |
| Net9 | 0.21 | 0.41 | 2.93 | 2.60 | 1 | 15 | 22 |
| Net10 | 0.48 | 0.63 | 2.86 | 2.43 | 2 | 7 | 10 |
| Net11 | 0.53 | 0.75 | 2.67 | 2.33 | 2.5 | 6 | 8 |
| Net12 | 0.24 | 0.52 | 3.33 | 4.60 | 3 | 15 | 25 |
| Net13 | 0.73 | 0.88 | 3.67 | 1.67 | 1.5 | 6 | 11 |
| Net14 | 0.43 | 0.57 | 2.57 | 2.14 | 2 | 7 | 9 |
| Net15 | 0.60 | 0.60 | 2.40 | 15.00 | 4 | 5 | 6 |
| Net16 | 0.33 | 0.38 | 2.67 | 3.78 | 2 | 9 | 12 |
| Net17 | 0.4 | 0.43 | 2 | 3.83 | 2.5 | 6 | 6 |
| Net18 | 0.5 | 0.5 | 2 | 2.8 | 1 | 5 | 5 |
| Net19 | 0.4 | 0 | 1.6 | 2.4 | 2 | 5 | 4 |
| Net20 | 0.87 | 0.87 | 4.33 | 19.83 | 15.5 | 6 | 13 |

In Table 1, each network of Vietnamese social scientists is represented by a series of different numbers. To aid our understanding, four most noticeable communities are chosen and plotted in the next two figures: one with highest number of nodes and one with highest number of links in figure 4; one with highest number of mean publications and one with lowest number of global clustering coefficient in Figure 5 (The R commands for figures 4 and 5 and the figures for all the nets that don't appear in the final paper can be seen in Dataset 6, 7 and 9). In Figure 4 and Figure 5, node size represents the number of publications; color is based on gender, blue for female and red for females; edges are links connect the nodes.



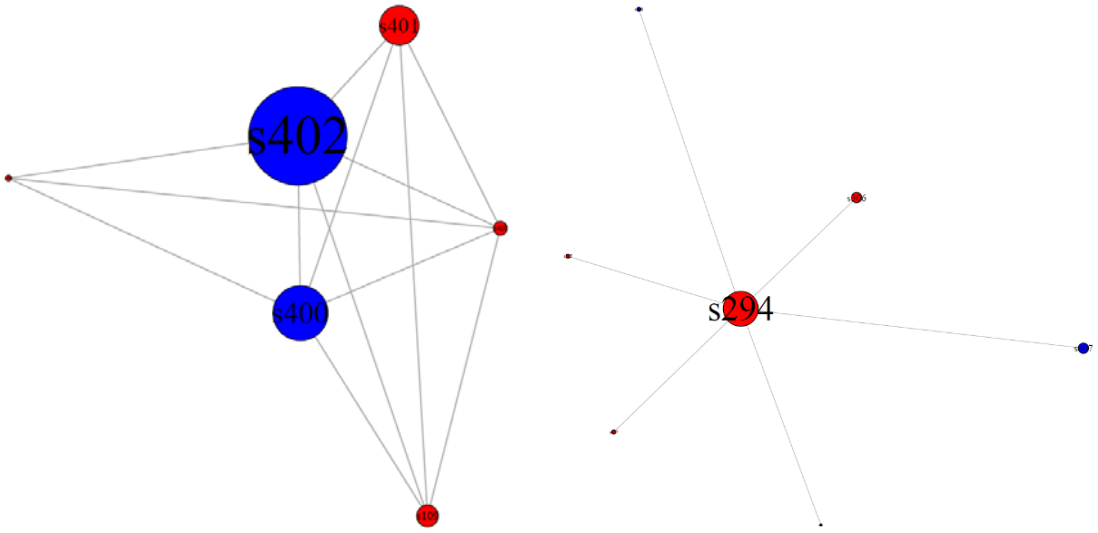

**(a)** Net8 is a community with the most researchers.

**(b)** Net6 is a community with the most links.

**Figure 4.** Two research communities among 412 Vietnamese social scientists standing out for the highest number of members and links. **(a)** Net8: Nodes=29; Links= 43; Density=0.11; Clustering=0.33; Mean degree=2.97; Mean publication=7.1; Median publication=2; **(b)** Net6: Nodes=19; Links=62; Density=0.36; Clustering=0.71; Mean degree=6.53; Mean publications=2; Median publication=1

In Figure 4, Net8 and Net20 are compared, as they are quite comparable in terms of nodes and links. In terms of clustering and density of connection, Net6 is higher than Net8 in both aspects; 0.36 vs. 0.11 in density, 0.71 vs. 0.33 in clustering coefficient. One can recognize the contrast among the

size of the nodes of Net8 (which has the most nodes) compared with Net6 (which has the most links). Net8 has more productive researchers in it; the fact is Net8 has 29 members but mean number of publication is 7.1, while Net6 has 19 members but the mean of publication is 2, which is equivalent to one third of the former.




**(a)** Net20 is a community with highest density of connection.

**(b)** Net3 is a community with no clustering coefficient.

**Figure 5.** Two research communities among 412 Vietnamese social scientists standing out for the highest density of connections. **(a)** a) Net20: Nodes=6; Links=13; Density=0.87; Clustering=0.87; Mean degree=4.33; Mean publication=19.83; Median publication=15.5; **(b)** Net3: Nodes=6; Links=7; Density=0.33; Clustering=0; Mean degree=2; Mean publication=4.86; Median publication=2

In Figure 5, Net20 and Net3 are compared as they are also comparable in terms of nodes and links. The most visible pattern is the contrast in clustering. Net20's clustering coefficient is 0.87 while Net3's is zero. Net20 consists of more productive members, and everyone collaborates with each other. Net3 is centered around the most productive member of the group. In terms of mean of publication, Net20 is about five times as much as that of Net3, 19.83 vs. 4.86.

Although there is no clear pattern in the data, our assumption is in a sustainable research community, information (knowledge, expertise, and experience) should be communicated efficiently; and for that to happen, the working hypothesis is the community should be low in clustering coefficient and high in density, in network metrics. That means, if we plotted a graph, in which the vertical axis is clustering coefficient, horizontal axis is density and each network is represented by a circle whose size is determined by its mean number of publications, we would be able to identify the bigger circle on the lower right quadrant of the chart to be our candidates for sustainable research networks. In Figure 6, density, clustering coefficient and mean number of publications are plotted to investigate the hypothesis.

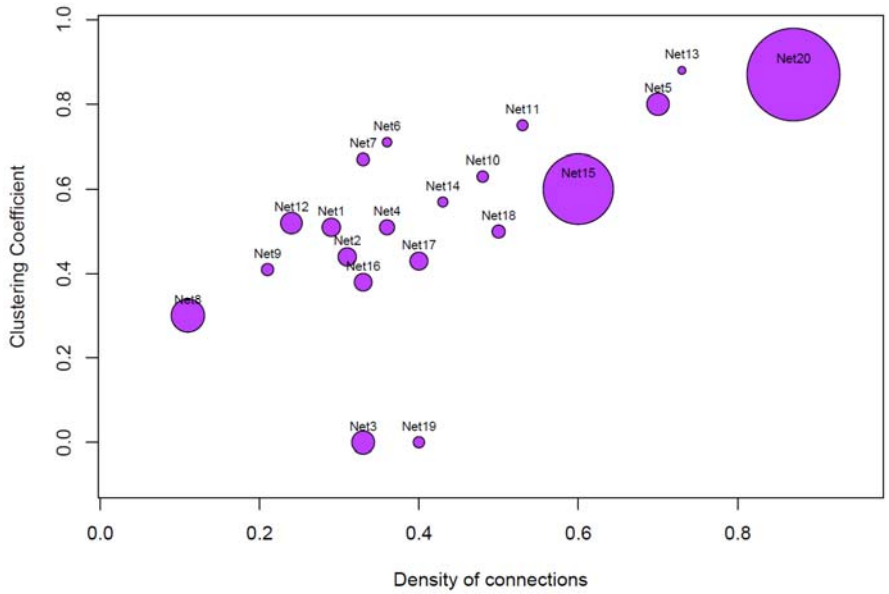



**Figure 6.** Relationship between network density, clustering coefficient and mean publications of each network. The vertical axis shows clustering coefficient, whose formula is $cl_T(G) = 3\tau\Delta(G)/\tau_3(G)$. The horizontal axis shows density, whose formula is $density = 2l/[n(n\text{-}1)]$. The size of the circle is equivalent to the mean publications of the co-authorship networks.

It is clear from the chart that there is no bigger circle on the lower right quadrant. There are some possible ways one can interpret this observation. First, it is possible that clustering and density might not have a strong enough connection with how productive a co-authorship network could be; thus, this approach to identify sustainable research networks might not be helpful. Second, it could also be the case that collecting data of 412 scientists in 10 years is not enough to investigate the matter; observing the movement of the network metrics over the years would provide a much clearer picture with regards to social sustainability of the research groups. The third possibility is the approach could be right and one can infer that no candidate can be qualified to be regarded as a sustainable research network in 412 social sciences in Vietnam. This is probable given how poorly social sciences in Vietnam have performed up until now [8-10]. In the next section, using the measure of social sustainability distance proposed in the previous section, it is hoped that insights into the status of social sustainability in Vietnamese social sciences communities could be unveiled.

### *4.2. Social sustainability distance*

Recall that, in a co-authorship network, any author's position could be defined using three quantities: number of publications, connections in the network and years in research. The assumption is, in a sustainable network of co-authors, the distance among the most productive ones and the rest should not be too great. In this paper, five measurements for this distance are calculated: distance from the most productive researchers to the mean, the median, the min, the second most productive and the mean of junior group (consists of people with equal or less than 3 publications). Table 2 summarizes the results.

**Table 2.** Social sustainability distance (*SSD*) measures and other attributes for 20 research networks of Vietnamese social scientists.

| Net | Max - Mean | Max - Median | Max - Min | Max -2nd | Max-Junior | No. Nodes | Mean publications |
|---|---|---|---|---|---|---|---|
| **Net1** | 9.38 | 11.05 | 12.41 | 6.71 | 12.21 | 11 | 3.91 |
| **Net2** | 22.99 | 30.41 | 30.41 | 28.86 | 25.55 | 10 | 3.9 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Net3** | 13.32 | 16.58 | 22.49 | 17.69 | 16.54 | 7 | 4.86 |
| **Net4** | 9.19 | 11.36 | 14.76 | 8.60 | 11.29 | 10 | 3.2 |
| **Net5** | 10.37 | 10.49 | 21.75 | 10.49 | 16.89 | 5 | 4.8 |
| **Net6** | 13.69 | 26.93 | 26.93 | 25.20 | 16.53 | 19 | 2 |
| **Net7** | 14.01 | 12.53 | 12.53 | 32.02 | 13.67 | 11 | 2.64 |
| **Net8** | 44.63 | 57.88 | 50.72 | 29.07 | 50.44 | 29 | 7.1 |
| **Net9** | 16.02 | 16.16 | 16.16 | 11.22 | 17.41 | 15 | 2.6 |
| **Net10** | 9.73 | 18.36 | 7.07 | 16.16 | 11.30 | 7 | 2.43 |
| **Net11** | 3.82 | 7.14 | 19.26 | 7.14 | 4.58 | 6 | 2.33 |
| **Net12** | 10.97 | 11.40 | 17.69 | 5.39 | 14.47 | 15 | 4.6 |
| **Net13** | 2.13 | 1.41 | 6.40 | 1.41 | 2.13 | 6 | 1.67 |
| **Net14** | 2.45 | 6.16 | 12.73 | 14.35 | 3.53 | 7 | 2.14 |
| **Net15** | 48.10 | 59.21 | 62.03 | 58.86 | 61.52 | 5 | 15 |
| **Net16** | 12.23 | 12.88 | 15.26 | 15.59 | 14.16 | 9 | 3.78 |
| **Net17** | 8.18 | 9.22 | 13.04 | 9.22 | 10.01 | 6 | 3.83 |
| **Net18** | 4.63 | 14.07 | 14.07 | 3.32 | 6.75 | 5 | 2.8 |
| **Net19** | 3.97 | 10.49 | 11.87 | 10.49 | 4.60 | 5 | 2.4 |
| **Net20** | 30.54 | 31.64 | 49.38 | 25.06 | 49.38 | 6 | 19.83 |

In Table 2, if we look at top 3 research networks in terms of the mean of publications—Net8, Net15, Net20—based on *SSD*, one can tell that Net15 is less sustainable than the other two. Net15 has only 5 members but *SSD* Max-2nd is about two times as much as the other two: 58.86 compared with 29.07 of Net8 and 25.06 of Net20. This measure shows such vast distance from the most productive member of Net15 to even the number two in terms of publications; for a group to maintain this huge gap among its members over 10 years, it shows a sign of unsustainable co-authoring behavior.

In the case of groups with small *SSD*, Net13 stands out as with lowest measures in all aspects: 2.13 for Max-Mean or 1.41 for Max-2nd for example. However, this is also a group with lowest mean number of publications, 1.67. This suggests when the gap is too small, the research group might be sustaining but sluggish.

To keep on investigating, what *SSD* measures could tell us about social sustainability of research networks, next, we plot all *SSD* measures of 20 networks on a radar chart in Figure 7.

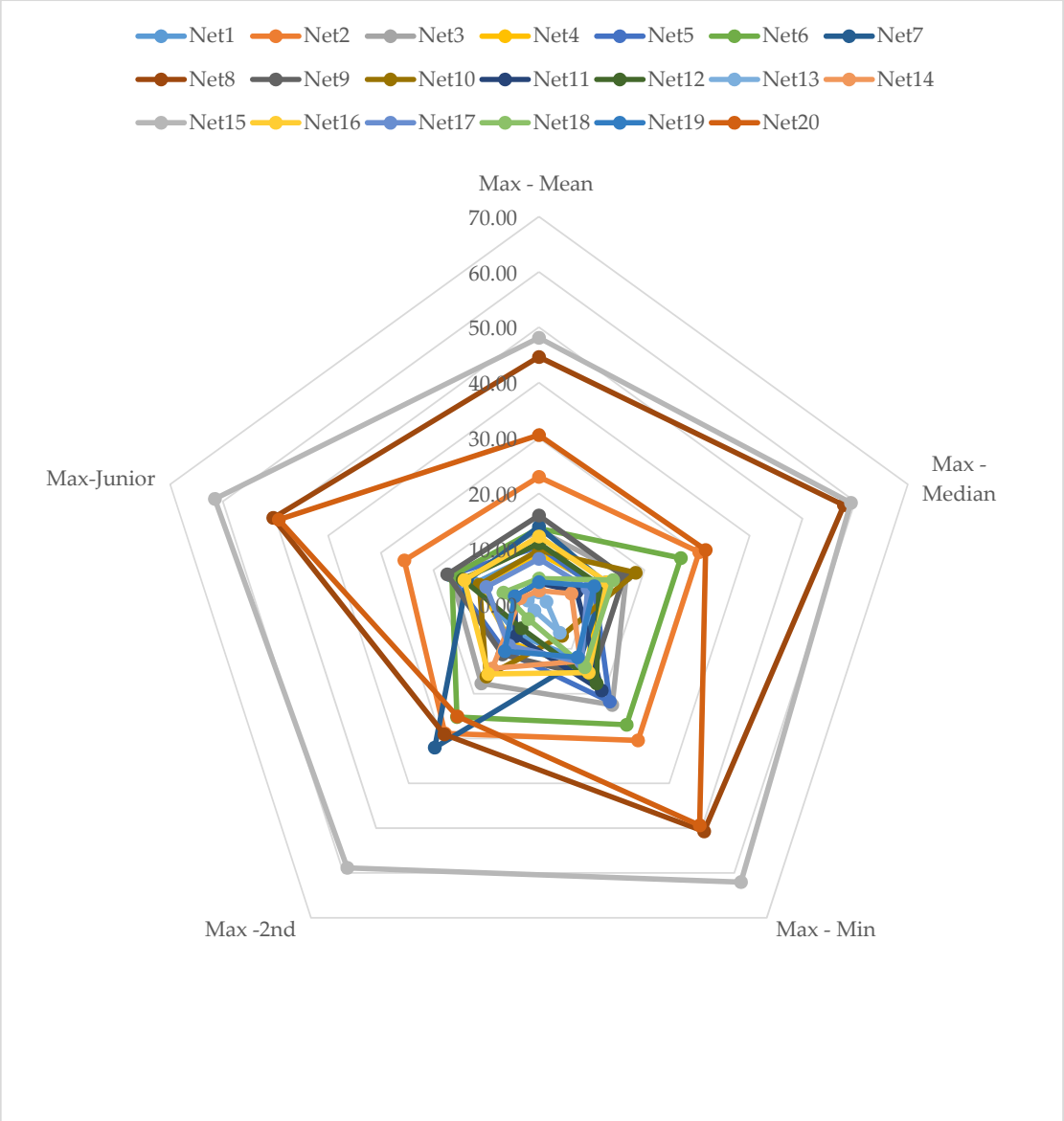



**Figure 7.** A radar chart plotting the different measures of social sustainability distance of 20 networks among Vietnamese social scientists

From Figure 7, it is quite obvious that across all networks, if the social sustainability distance is high in one kind of measurement, it is likely to be high in all others. This pattern holds even when the networks can differ wildly in various properties (see Table 2). One then can infer that over the period of 10 years, the most productive researchers tend to accelerate away from everyone else in the network and the gap would remain. From this data, one can make an educated guess that it is not easy to close the gap even between the most productive researcher and the second most productive one. This might be caused by the inefficient knowledge transfer among the researchers in co-authorship networks in social sciences in Vietnam, which implies a quite socially unsustainable situation.

## 5. Discussion

### *5.1. Policy implications for Vietnam*

As we have seen in the introduction, recently, the Vietnamese government has pursued science policies that incentivize international publications and strong research groups [7]. Hence, the ability to identify and create sustainable research groups will become central for policy-making in this area to be effective. Although the results presented in this study are still preliminary, the network standard measures and the proposed measures have offered a novel approach to address this issue. Density of connections, clustering coefficient and social sustainability distance allow us to glimpse and compare the structure and dynamics of the groups in a rather holistic way. This has allowed for the identification of two ways in which a research group could be unsustainable: the gap among its members in terms of productivity, connections and research years are either too large or too small.

Another interesting problem related to science funding in Vietnam is how to create more enduring scientific projects, in which the knowledge of one generation could be inherited by the next [4-5]. A useful approach to solving this problem is to apply social network analysis and observe the dynamics and evolutions in each research network by measuring how *SSD*, clustering and density change over the years in each network.

Similar to the world, Vietnamese society is taken by surprise with the speed of technological progress, which has created new and unforeseen social problems. In this context, the ability of Vietnamese social scientists to address these new problems will be essential for the country to continue making social progress in the coming years. Consequentially, for Vietnam, solving the problem of science funding and policy is not only the matter of improving the quality of academic research but also a matter of improving the odds that the country will thrive when facing the challenges of a new era.

### *5.2. On applicability of SSD metrics in an ever-technological advanced world*

In life, whenever there is a gap, there is a potential place for problems to arise, yet, a certain level of gap is acceptable or even desirable. The question of social sustainability is really about finding how big a gap is sustainable for a complex social system. This is indeed a difficult and also most central problem to the time we are living in. The world is witnessing such acceleration of technological progress that most humans find it hard to keep up. There is, without a doubt, a societal problem with the gap of understanding technology: the gap between the most scientific and technological adept and the laypeople. An increasing number of people whose lives are governed by technologies don't understand them well, while the elites who understand technologies could exploit the gap to their benefits. Though the study does not provide a direct diagnosis for this gap of understand technology, the measurement of social sustainability distance of is a suggestion for how the gap could be measured.

## 6. Limitations and future research directions

### *6.1. Refining and extending measurements of social sustainability distance*

For any new way of measuring, especially of such an elusive thing as social sustainability, it is certainly not without risks of being wrong and misguided. Therefore, it is imperative to cover as much empirical ground as possible by refining the measure of *SSD* as well as extending it. In terms of refining the measure, one can renormalize the measure by dividing the deduction of two quantities by the larger quantity. The new formula of a refined *SSD* metric can have the form as displayed by equation (4):

$$SSD_{\text{renormalized}} = \sqrt{\left(\frac{x_{\max} - x_0}{x_{\max}}\right)^2 + \left(\frac{y_{\max} - y_0}{y_{\max}}\right)^2 + \left(\frac{z_{\max} - z_0}{z_{\max}}\right)^2}\,. \tag{4}$$

As we have seen in Figure 3, the most productive members seem to always be central to many connections within a co-authorship network; in this study, we reasoned that the distance from the most productive member to the rest should be the most telling in terms of social sustainability and focused exclusive on this gap. However, to check against this assumption, one can explore options other than those appeared in this study, for example, mean-min, median-min, 2nd highest to mean, etc. Carrying out empirical investigations in this way would help us test the proposed measure in this study against all available options, thus finding out which one is the most reliable in helping us understand social sustainability of research groups.

### *6.2. More advanced application of social network analysis*

This study cannot claim to have fully utilized the power of the social network analysis. The study is limited to whether or not there exists a co-authorship connection and has not taken into account the weight of each co-authorship connection, which is defined as the number of times two authors published together. In social network analysis literature, techniques to study weighted connections have been well-developed [78]. One could well see this dimension of connection intensity among scholars could yield new insights into the social sustainability of the research communities.

Related to the question of social sustainability, another aspect of co-authorship network that is worth exploring is direction of the co-authorship connections. In this study, all connections among authors are treated as equal, i.e., there is no need to specify the order of two nodes that define an edge. However, in the real world, it is common that the first author or the corresponding author of a paper has a leadership role. One can argue that a sustainable research network would have a relatively high level of reciprocity among co-authors rather than a few authors dominantly play the leadership role. Hence, when this dimension of connection directionality is added into the analysis, network metric such as co-authoring reciprocity will help us achieve deeper understanding of social sustainability of a network.

### *6.3. Further development of empirical strategies*

It might be the case that the data collected for 412 Vietnamese social scientists is not enough to allow a truly thorough empirical investigation of the concepts and the measurements explored in this study. Consequently, the results of the study are limited to be only indicative of the situation of social sustainability of the research communities in social sciences in Vietnam. This problem can be solved by developing more advanced empirical strategies. For example, instead of investigating one block of time from 2008-2017, one can divide the block into smaller chunks of 3 years each. In this way, one can investigate how *SSD* metrics and clustering coefficient change over the years; as such, giving a better sense of how the research networks evolve or deteriorate. One can also study the problems raised in the study longitudinally by collecting the data of 30 or even 50 years. By following the research networks over a long period of time, the theoretical assumptions raised in this study can then by fully tested against reality.

## **7. Conclusion**

The study has developed a novel way to understand co-authorship behavior in Vietnam's social scientific communities as well as to evaluate their social sustainability. Using social network analysis, the study is able to answer the three research questions put forth in its premise. First, the technique helps generate a holistic picture of co-authorship among the 412 Vietnamese social scientists. Given the interconnected links among individual researchers, one can conclude that co-authorship is widespread and serves as the main collaborative form among these academicians. Through standard network metrics and visualization, the study shows that in the majority of Vietnamese research communities, the most productive researchers appear central to many different connections. The trends of co-authoring, then, are captured in three dominant groups: (i) researchers who work alone or with foreigners—the factor that has not yet been taken into account in this study,

(ii) researchers who work in a small group of 2 to 9 members, and (iii) researchers who work in a large group of at least 10 members.

Second, the study uses the basic network metrics, namely clustering coefficient and density of connection, and extracts 20 communities with at least 5 members each from the total dataset to analyze the extent of social sustainability in Vietnamese social scientists' communities. The resulting data support the argument that high density and low clustering indicate a socially sustainable research network. In other words, high density implies a condition where there is a large volume of information (knowledge and expertise) being exchanged while low clustering refers to a state where such information is being exchanged in an efficient manner.

Third, based on the ground that any researcher's position in a network can be defined by three quantities: (i) number of publications, (ii) connections, and (iii) and years in research, the study calculates and compares the distance between the most productive researchers and the rest of the communities. If the distance between the most productive researcher and his or her remaining group members is too large or too small, this research network might be socially unsustainable.

While the metrics used in this study suggest some level of unsustainability within Vietnam's social science communities, they nonetheless offer for the first time a quantitative way to identify both the productivity and sustainability of the local research networks. In the future, proper development of the empirical strategies could yield a more complete picture of this matter, perhaps in conjunction with related statistical investigations into transformed data types such as categorical data for acquiring conditional probabilities [79].

**Supplementary Materials:** Data availability is detailed as follows:
Dataset 1: "Folder Net412's Data". This folder contains two files: "20170725_net412_NODES.csv" and "20170729_net412_LINKS.csv". The former lists all 412 individuals in the study and their attributes, each individual is considered a node (vertex) in the network. The later lists the number of co-written articles between all 412 authors of the network, where relevant; each collaboration is counted as a link (edge) in the network.
Dataset 2: "Folder 20 Networks' Data". This folder includes all nodes lists and edges lists of 20 small networks extracted from the original.
Dataset 3: "Metrics all nets 20171003.csv" This dataset contains the summary of all metrics that represent 20 research networks in the study.
Dataset 4: "Folder SSD Calculation". This folder contains four files: "Research network details 20171012.xlsx" contains the nodes lists for 20 research networks in the study as well as the calculations for the junior members of each network; "SSD Calculation details 20171012.xlsx" contains the details of SSD calculations for all networks, "SSD Summary 20171012.xlsx" is the summary of final results of calculations for all SSD measurements of 20 networks in the study.
Dataset 5: "Folder Rcommand and figure3" contains R commands and the PDF file of Figure 3.
Dataset 6: "Folder Rcommand and figure4" contains R commands and the PDF file of Figure 4.
Dataset 7: "Folder Rcommand and figure5" contains R commands and the PDF file of Figure 5.
Dataset 8: "Folder Rcommand and figure6" contains R commands and the PDF file of Figure 6.
Dataset 9: "Folder Rcommand and figure for all nets" contains the R commands files and the PDF files for all nets that did not appear in the final paper.

**Acknowledgments:** We would like to thank Vuong & Associates for their research initiative *The Network of Vietnamese Social Scientists* ("NVSS"), which enabled the research process and provided the raw data for the study. We particularly thank Dam Thu Ha and Nghiem Phu Kien Cuong for their excellent research assistance. We specially thank Nguyen To Viet Ha, Pham Hung Hiep, Nancy K. Napier and Dam Quang Minh for their valuable comments. We are also grateful for useful comments from participants—who came from different academic disciplines—in the seminar "Beauty and the Light of Hope" (reflecting the spirit of [80)] co-organized by Western University Hanoi and Vuong & Associates in Hanoi, on October 3, 2017, particularly Nguyen Tu Cuong (mathematics), Nguyen Ngoc Chau (biology), Pham Duc Chinh (mechanics), Khuat Thu Hong (sociology), Le Van Canh (linguistics), Nguyen Viet Cuong and Tran Quang Tuyen (economics), Hoang Anh Tuan Kiet (physics), Tran Kien (law), Bui Quang Khiem (fine arts), Nguyen Pham Muoi (journalism/media), Bach Ngoc Chien (management), to name just a few.

**Author Contributions:** Q.-H.V. conceived and designed the study; all contributors (T.M.H., N.H.K, Q.-H.V.) contributed equally for the remaining tasks of the research and manuscript preparation, i.e. data set preparation, analysis, interpreting results, and writing and checking the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Grant Information:** The authors declared that no grants were involved in supporting this work.

## References

1. Ministry of Education and Training. Circular Number 08/2017/TT-BGDĐT, issued on April 4, **2017** by Minister of Education and Training. Available online: https://moet.gov.vn/van-ban/vanban/Pages/chi-tiet-van-ban.aspx?ItemID=1249 (accessed on 10 October 2017).
2. Ha, N. Reforming the criteria for the professorship titles: time to change. *Tuoi Tre Online* **2017** (March 27). Available: http://tuoitre.vn/doi-moi-tieu-chuan-chuc-danh-giao-su-den-luc-phai-thay-doi-1287110.htm (accessed on 18 September 2017; article in Vietnamese).
3. Chinh, P.D. Reforming by raising the academic standards right from inside of the State councils on professorship titles. *Vietnamnet* **2017** (February 8). Available online: http://vietnamnet.vn/vn/ giao-duc/khoa-hoc/du-thao-moi-ve-quy-dinh-tieu-chuan-giao-su-pho-giao-su-cai-cach-nang-chuan-tu-chinh-cac-hoi-dong-chuc-danh-giao-su-354890.html (accessed on 17 September 2017; article in Vietnamese).
4. Nhan, T. NAFOSTED invests in strong research teams. *Tia Sang* **2017** (August 24). Available online: http://tiasang.com.vn/-khoa-hoc-cong-nghe/NAFOSTED-dau-tu-cho-nhom-nghien-cuu-manh--10876 (accessed on 10 October 2017).
5. NAFOSTED. Circular Number 37/2014/TT-BKHCN, issued on December 12, **2014** by Minister of Science and Technology. Available online: http://www.nafosted.gov.vn/vi/archives/view/37242014TT-BKCN-26/ (accessed on 10 October 2017).
6. Hanh, H; Thu, L. Top 300 Asian Ranking: Why there have been no Vietnamese universities? *Dan Tri* **2017** (March 29). Available online: http://dantri.com.vn/su-kien/xep-hang-300-truong-dh-hang-dau-chau-a-vi-sao-dai-hoc-viet-nam-chua-ghi-ten-20170329162534079.htm (accessed on 18 September 2017; article in Vietnamese).
7. Huong, L. Nowhere in the ranking, what happened with Vietnamese higher education? *Giao Duc Viet Nam* **2017** (April 16). Available online: http://giaoduc.net.vn/Giao-duc-24h/Bang-xep-hang-khong-co-ten-dai-hoc-Viet-Nam-sao-the-nhi-post175835.gd (accessed on 18 September 2017; article in Vietnamese).
8. Diep, B. Debate around the 90 billion dong funding one internationally published article. *Dan Tri* **2016** (April 27). Available online: http://dantri.com.vn/kinh-doanh/tranh-cai-quanh-90-ty-dong-cho-1-bai-bao-khoa-hoc-quoc-te-2016042707233457.htm (accessed on 18 September 2017; article in Vietnamese).
9. Thuy, H. Why Vietnam Academy of Social Sciences had too few international publications? *Vnexpress* **2016** (April 23). Available online: https://vnexpress.net/tin-tuc/giao-duc/vi-sao-vien-han-lam-khoa-hoc-xa-hoi-it-co-nghien-cuu-cong-bo-quoc-te-3391989.html (accessed on 18 September 2017; article in Vietnamese).
10. Pham, Q.M. It is time to reform research in social sciences and humanities. *Tia Sang* **2017** (January 24). Available online: http://tiasang.com.vn/-quan-ly-khoa-hoc/Da-den-luc-phai-doi-moi-trong-nghien-cuu-KHXHNV-10389 (accessed on 18 September 2017; article in Vietnamese).
11. Nguyen, T. V.; Pham, L. T. Scientific output and its relationship to knowledge economy: an analysis of ASEAN countries. *Scientometrics* **2011,** *89*, 107-117, DOI: 10.1007/s11192-011-0446-2.
12. Yi, Y.; Qi, W.; Wu, D. Are CIVETS the next BRICs? A comparative analysis from scientometrics perspective. *Scientometrics* **2013**, *94*, 615-628, DOI: 10.1007/s11192-012-0791-9.
13. Manh, H. D. Scientific publications in Vietnam as seen from Scopus during 1996–2013. *Scientometrics* **2015**, *105*, 83-95, DOI: 10.1007/s11192-015-1655-x.
14. Nguyen, T. V.; Ho-Le, T. P.; Le, U. V. International collaboration in scientific research in Vietnam: an analysis of patterns and impact. *Scientometrics* **2017**, *110*, 1035-1051, DOI: 10.1007/s11192-016-2201-1.
15. Vuong, Q.H.; Ho, T.M.; Vuong, T.T.; Napier, N.K.; Pham, H.H.; Nguyen, H.V. 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*, **2017**, *43(3)*, 51-55, DOI: 10.20316/ESE.2017.43.006.

16. Vuong, Q.-H.; Ho, T.M.; Vuong, T.-T.; Nguyen, H.V.; Napier, N.K.; Pham, H.-H. *Nemo Solus Satis Sapit*: Trends of research collaborations in the Vietnamese social sciences, observing 2008–2017 Scopus data. *Publications* **2017**, *5*, 24, DOI: 10.3390/publications5040024.
17. Ho, T. M.; Nguyen, H. V.; Vuong, T. T.; Dam, Q. M.; Pham, H. H.; Vuong, Q. H. Exploring Vietnamese co-authorship patterns in social sciences with basic network measures of 2008-2017 Scopus data. *F1000Research* **2017**, *6*, 1559, DOI: 10.12688/f1000research.12404.1.
18. Vuong, Q.H.; Napier, N.K. Making creativity: the value of multiple filters in the innovation process. *Int. J. Transitions and Innovation Systems* **2014**, *3*(4), 294-327.
19. Vuong, Q.H.; Napier, N.K. Acculturation and global mindsponge: an emerging market perspective. *Int. J. Intercultural Relations* **2015**, *49*, 354-367.
20. Vuong, Q. H. Global mindset as the integration of emerging socio-cultural values through mindsponge processes: A transition economy perspective. In J. Kuada (Ed.) *Global Mindset: Exploration and Perspectives*, Routledge: New York, USA, **2008**; pp.109-126.
21. Lafferty, W.M. (Ed.) *Governance for Sustainable Development. The Challenge of Adapting Form to Function*. Elgar, Cheltenham, **2004**.
22. Brundtland, G.H. *Report of the World Commission on Environment and Development: Our Common Future*. Oxford University Press: Oxford, UK, **1987** (ISBN: 019282080X)
23. Robinson, J. Squaring the circle? Some thoughts on the idea of sustainable development. *Ecological Economics* **2004**, *48*(4), 369-384
24. Holden, E., Linnerud, K., & Banister, D. Sustainable development: our common future revisited. *Global Environmental Change* **2014**, *26*, 130-139.
25. Walker, P. (2017). Sustainability: a discipline and a political agenda?. *Environ. Hazards* **2017**, *16*(2), 93-98, DOI: 10.1080/17477891.2017.1296810.
26. Mebratu, D. Sustainability and sustainable development: historical and conceptual review. *Environmental Impact Assessment Review* **1998**, *18*, 493-520.
27. Sneddon, C.; Howarth, R.B.; Norgaard, R.B. Sustainable development in a post-Brundtland world. *Ecological Economics* **2006**, *57*(2), 253-268.
28. Union of Concerned Scientists. Available online: http://www.ucsusa.org/ (accessed on 20 September 2017)
29. Malthus, R.T. *An Essay on the Principle of Population.* J. Johnson: London, UK, **1798**.
30. Ehrlich, P.R. *The Population Bomb*. Ballantine Books: New York, USA, **1968**.
31. Meadows, D.H.; Meadows, D.L.; Randers, J.; Behrens, W.W. *The Limits to Growth*. The Universe Books: New York, USA, **1972**.
32. Tainter, J. *The Collapse of Complex Societies*. Cambridge University Press: Cambridge, UK, **1988**.
33. Diamond, J. *Collapse: How Societies Choose to Fail or Succeed*. Penguin: New York, USA, **2005**.
34. Marx, K.; Engel, F. Alienation and Social Classes (from the Holy Family). In *The Marx-Engels Reader*, 2nd ed.; Tucker, R.C.; W.W. Norton & Company: New York, USA, **1972**.
35. Stiglitz, J.E. *The Price of Inequality: How Today's Divided Society Endangers Our Future*; W.W. Norton & Company: New York, United States, 2012; ISBN: 978-0-393-34506-3.
36. Piketty, P. *Capital in the 21st Century*; Harvard University Press: Massachusetts, USA, **2013**.
37. Wilkinson, R.G.; Pickett, K. *The Spirit Level: Why More Equal Societies Almost Always Do Better*; Allen Lane: London, UK, **2009**.
38. Freeland, C. *Plutocrats: The Rise of the New Global Super-Rich and the Fall of Everyone Else*; Doubleday Canada: Toronto, Canada, **2012**.
39. Bostrom, N. Existential risk prevention as global priority. *Glob. Policy* **2013,** *1*, 15-31.
40. Posner, R.A. *Catastrophe: Risk and Response*; Oxford University Press: Oxford, UK, **2004**.
41. Rees, M.J. *Our Final Hour: A Scientist's Warning—How Terror, Error, and Environmental Disaster Threaten Humankind's Future in This Century—On Earth and Beyond*; Perseus Books Group: New York, USA, **2004**.
42. Bostrom, N.; Cirkovic, M. *Global Catastrophic Risks*. Oxford University Press: New York, USA, **2008**.
43. Elkington, J. Towards the sustainable corporation: Win-win-win business strategies for sustainable development. *Calif. Manag. Rev* **1994**, *36*, 90-100.
44. Kuhlman, T.; Farrington, J. What is sustainability?. *Sustainability* **2010**, *2*, 3436-3448, DOI: 10.3390/su2113436.
45. Tegmark, M. *Life 3.0: Being Human in the Age of Artificial Intelligence*. K. Doubleday: New York, USA, **2017**.
46. Bettencourt, L.; West, G. A unified theory of urban living. *Nature* **2010**, *467*, 912-913, DOI: 10.1038/467912a.

47. Whitemarsh, L. Behavioural responses to climate change: asymmetry of intentions and impacts. *J. Environ. Psychol.* **2009**, 29, 13–23, DOI: 10.1016/j.jenvp.2008.05.003.
48. Stern, P.C. New environmental theories: Toward a coherent theory of environmentally significant behavior. *J. Soc. Issues* **2000**, *56*, 407-424, DOI: 10.1111/0022-4537.00175.
49. Whitmarsh, L.; O'Neill, S. Green identity, green living? The role of pro-environmental self-identity in determining consistency across diverse pro-environmental behaviours. *J. Environ. Psychol.* **2010**, *30*, 305 - 314, DOI: 10.1016/j.jenvp.2010.01.003.
50. Barreto, M. L.; Szóstek, A.; Karapanos, E.; Nunes, N. J.; Pereira, L.; Quintal, F. Understanding families' motivations for sustainable behaviors. *Comput. Hum. Behav* **2014**, *40*, 6-15, DOI: 10.1016/j.chb.2014.07.042.
51. Grønhøj, A.; Thøgersen, J. Why young people do things for the environment: The role of parenting for adolescents' motivation to engage in pro-environmental behaviour. *J. Environ. Psychol* **2017,** *54*, 11-19, DOI: 10.1016/j.jenvp.2017.09.005.
52. Blake, J. Overcoming the 'value-action gap' in environmental policy: Tensions between national policy and local experience. *Local Environ* **1999**, *4*, 257-278, DOI: 10.1080/13549839908725599.
53. Kollmuss, A.; Agyeman, J. Mind the gap: why do people act environmentally and what are the barriers to pro-environmental behavior? *Envi. Edu. Res.* **2002**, *8*, 239-260.
54. Newton, P.; Meyer, D. Exploring the attitudes-action gap in household resource consumption: does "Environmental Lifestyle" segmentation align with consumer behaviour? *Sustainability* **2013**, *5*, 1211-1233, DOI: 10.3390/su5031211.
55. Gifford, R. The dragons of inaction: psychological barriers that limit climate change mitigation and adaptation. *Am. Psychol.* **2011**, *66*, 290, DOI: 10.1037/a0023566.
56. Rees, W. Whats blocking sustainability? Human nature, cognition, and denial. *Sustainability: Science, Practice, & Policy* **2010**, *6*(2).
57. Markowitz, E. M.; Shariff, A. F. Climate change and moral judgement. *Nat. Clim. Change* **2012**, *2*, 243-247, DOI: doi:10.1038/nclimate1378.
58. Deutsch, D.E. *The Beginning of Infinity: Explanations That Transform the World*. Viking: New York, USA, **2011**.
59. Edge. Why cities keep growing, corporations and people always die, and life gets faster. *Edge Conversation* **2011** (May 23). Available online: https://www.edge.org/conversation/geoffrey_west-why-cities-keep-growing-corporations-and-people-always-die-and-life-get (accessed on 04 September 2017).
60. West, G. Universal scaling laws from cells to cities: A physicist's search for quantitative, unified theories of biological and social structure and dynamics. *Bull. Am. Phys. Soc.* **2013**, *58*. Available online: http://absimage.aps.org/image/APR13/MWS_APR13-2013-020103.pdf
61. Ball, P. Complexity: Decoding deep similarities. *Nature* **2017**, *545*, 154-155, DOI: 10.1038/545154a.
62. Nica, E.; Potcovaru, A.M. The social sustainability of the sharing economy. *Econ. Manag. Fin. Markets* **2015**, *10*(4), 69-75.
63. Hollander, R.; Amekudzi-Kennedy, A.; Bell, S.; Benya, F.; Davidson, C.; Farkos, C.; ... Quigley, D. Network priorities for social sustainability research and education: Memorandum of the Integrated Network on Social Sustainability Research Group. *Sustainability: Science, Practice, & Policy* **2016**, *11*(2). Available online: http://www.tandfonline.com/doi/pdf/10.1080/15487733.2016.11908150.
64. Almahmoud, E.; Doloi, H.K. Assessment of social sustainability in construction projects using social network analysis. *Facilities* **2015**, *33*(3/4), 152-176, DOI: 10.1108/F-05-2013-0042.
65. Hutchins, M.J.; Sutherland, J.W. An exploration of measures of social sustainability and their application to supply chain decisions. *J Clean Prod* **2008**, *16*(15), 1688-1698, DOI: 10.1016/j.jclepro.2008.06.001.
66. Magis, K. Community resilience: An indicator of social sustainability. *Soc. and Nat. Resources* **2010**, *23*(5), 401-416, DOI: 10.1080/08941920903305674
67. Vallance, S.; Perkins, H. C.; Dixon, J. E. What is social sustainability? A clarification of concepts. *Geoforum* **2011**, *42*(3), 342-348, DOI: 10.1016/j.geoforum.2011.01.002.
68. Weingaertner, C.; Moberg, Å. Exploring social sustainability: learning from perspectives on urban development and companies and products. *Sus Dev* **2014**, *22*, 122-133, DOI: 10.1002/sd.536.
69. Karol, E.; Brunner, J. Tools for measuring progress towards sustainable neighborhood environments. *Sustainability* **2009**, *1*, 612-627, DOI: 10.3390/su1030612.
70. Turcu, C. Re-thinking sustainability indicators: local perspectives of urban sustainability. *J Environ Plann Man* **2013**, *56*, 695-719, DOI: 10.1080/09640568.2012.698984

71. Newman M.E. Scientific collaboration networks. I. Network construction and fundamental results. *Phys Rev E Stat Nonlin Soft Matter Phys* **2001**, 64(1 Pt 2): 016131, DOI: 10.1103/PhysRevE.64.016131
72. Kajikawa, Y.; Ohno, J.; Takeda, Y.; Matsushima, K.; Komiyama, H. Creating an academic landscape of sustainability science: an analysis of the citation network. *Sus. Sci.* **2007**, *2*, 221.
73. Moody, J. The structure of a social science collaboration network: Disciplinary cohesion from 1963 to 1999. *Am Sociol Rev* **2004**; 69, 213–238, DOI: 10.1177/000312240406900204.
74. Li, E.Y.; Liao, C.H.; Yen, H.R. Co-authorship networks and research impact: A social capital perspective. *Res. Pol.* **2013**, *42*(9), 1515-1530, DOI: 10.1016/j.respol.2013.06.012.
75. Ding, Y. Scientific collaboration and endorsement: Network analysis of coauthorship and citation networks. *J. Informetr.* **2011**; *5*(1), 187–203, DOI: 10.1016/j.joi.2010.10.008.
76. Ingold, K.; Balsiger, *J. Reg. Environ. Change* **2015**, *15*, 529, DOI:10.1007/s10113-013-0575-7.
77. Wu, X.; Liu, Z. How community structure influences epidemic spread in social networks. *Physica A: Statistical Mechanics and its Applications* **2008**; *387*(2), 623–630, DOI: 10.1016/j.physa.2007.09.039.
78. Kolaczyk, E.D.; Csárdi, G. *Statistical analysis of network data with R*. Springer: New York, USA, **2014.**
79. Vuong, Q.H. Survey data on Vietnamese propensity to attend periodic general health examinations. *Sci. Data* **2017**, *4*, 170142, DOI:10.1038/sdata.2017.142. Available from: www.nature.com/articles/sdata2017142.
80. Bach, N.C; Vuong, Q.H. *Bằng chứng cuộc sống: suy ngẫm về phát triển bền vững Việt Nam*. National Political Publishing House: Hanoi, Vietnam..


**Data Citations**:


81. Vuong, Q.H. NVSS Data Oct-2017 (N=412; 20 Groups) for New Metrics. *Mendeley Data*, v1 (**2017**); DOI: http://dx.doi.org/10.17632/bwgd7cdmbc.1 ;