

UNIVERSITY OF TARTU
Faculty of Social Sciences
School of Economics and Business Administration

Kristina Tsvetkova

DIVERSITY OF SCIENCE IN THE EU AND UK

Bachelor Thesis

Supervisor: Junior Lecturer in Accounting Mariia Chebotareva

Tartu 2023

I have written this Bachelor Thesis independently. Any ideas or data taken from other authors or other sources have been fully referenced.

Table of contents

Introduction.....	4
1. Theoretical background of measuring diversity	6
1.1. Defining diversity of science	6
1.2. Overview of different approaches for measuring diversity	8
2. Empirical analysis of diversity indexes for the EU countries and UK	15
2.1. Methodology of measuring diversity of science by the means of different indexes	15
2.2. Discussion of empirical results	25
Conclusion	30
References.....	32
APPENDICES	36
Appendix A.....	36
Appendix B	37
Appendix C.....	38
Resümee.....	39

Introduction

Science never stands still. Every day scientists around the world explore new areas of perception and life, invent new laws and disprove previous research based on newer and more reliable knowledge. Due to this constant process of development of the world and technology we are confronted with a “diversity of science” concept.

An interest to the diversity of science is growing, which is thought to be threatened by science policy's pursuit of excellence. The prevailing technique for achieving excellence is to focus funding on the top achievers. This policy indicates that fewer units will be supported in each field, which is likely to reduce diversity. (Adams, 2003)

The adaptive behaviour of scientists is a danger to variety. Researchers are likely to select initiatives that are supported by the scientific community and have a high likelihood to get funded. Such safe projects use techniques that are recognized to produce results and adhere to a field's mainstream (Schmidt et al., 2006). These arguments lack an empirical basis, despite being compelling. There hasn't been a strong indicator of research diversity yet. According to empirical research on how the British Research Assessment Exercise (RAE) affected the diversity of economic research, academics' perspectives are becoming more homogeneous (Harley, 1997).

There is also a strong connection between the diversity of science and the sustainability of science. The educational and scientific system of the country is more stable when it has more diverse research fields (Lacey et al., 2019). In other words, the more diverse system, the more sustainable it is. A diverse scientific community is more likely to generate new and innovative ideas, approaches, and perspectives that can help address complex scientific challenges, such as climate change, biodiversity loss, and public health crises. These challenges require multidisciplinary and diverse approaches that consider different cultural, social, and environmental factors. Stilgoe and Guston (2017) emphasize the importance of considering diverse perspectives and engaging stakeholders in scientific decision-making processes in order to promote sustainability of science. A lack of diversity in the scientific community can result in blind spots and biases that can hinder progress in addressing these challenges (Spangenberg, 2011).

The term “diversity” is actively used in different fields. In ecology “diversity” is defined as “the richness and variety of natural ecological communities” (Pielou, 1975), in biology as “the variety and abundance of species in a defined unit of study” (Magurran, 2004 & Schmidt et al, 2006). In social sciences, “diversity” can be generally understood as providing fair chances for employment, but in a broader sense, it entails embracing all the

unique aspects that set individuals apart from one another (Carrell et al., 2006). Only a few authors dealt with the “diversity of science” concept (e.g. Schmidt et al. 2006; Stirling, 2007). However, “diversity of science” is still not clearly defined and there is no uniformly applied metrics for assessing the diversity of scientific fields.

Therefore, the current thesis is driven by a lack of analytical approaches for determining the level of scientific diversity in the EU countries and United Kingdom (UK). All available methodologies taken independently are insufficient to measure sectoral diversity of science. As a result, the author intends to specify methodology for measuring diversity of science by combining various methods (i.e. literature analysis, document analysis and statistical data analysis).

The aim of the thesis is to verify applicability of diversity indexes for measuring diversity of scientific fields in the European countries.

In order to achieve the aim, the following research tasks have been set up:

- to define "diversity of science" for the purposes of science and research;
- to provide an overview of existing methodological approaches for evaluating diversity of scientific fields;
- to collect data on number of students enrolled and graduated by doctoral/equivalent level;
- to calculate the diversity of scientific fields based on selected indexes in the EU and UK;
- to interpret and discuss the results.

The topic is relevant and important since scientific system with more diverse structure is more resistant to the effects of various negative factors (Adams & Smith, 2003). It is also worth noting that the more diverse the science, the faster and more extensive it develops (Leydesdorff et al., 2019). Moreover, a diverse scientific community can help ensure that research questions are asked from a variety of perspectives and that data is analysed and interpreted in a way that considers different cultural and societal contexts. This can lead to more accurate and comprehensive scientific findings. (Swartz et al., 2019) Therefore, this thesis can be of interest for the academics at universities and other research organizations, policy makers in ministries of science and education as well as for the society as such.

The thesis consists of two chapters: theoretical background and empirical analysis. Subchapter 1.1 includes an overall explanation of the concept of diversity of science. In addition, a table of different definitions of diversity in various fields of science is made and diversity of science is defined in the framework of the current thesis. Subchapter 1.2 provides

an overview and comparative analysis of different indexes for measuring diversity as well as advantages and disadvantages of several commonly used indexes. Chapter 2 is an empirical chapter where the diversity of science is measured by the means of several diversity indexes in subchapter 2.1 and the results in the context of the EU and UK are discussed in subchapter 2.2.

Keywords: diversity, science, variety, metrics, indexes, sustainability

1. Theoretical background of measuring diversity

1.1. Defining diversity of science

The concept of "diversity" has been researched across a number of scientific areas and by different authors (Adams, 2003; Hurlbert 1971; Schmidt et al. 2006; Hill 1973; Stirling 2007; Sander 1968; Magurran 2004, etc.). The term is most commonly used in biological and ecological studies, where "diversity" refers to the variability and abundance of species (see Table 1).

In Schmidt et al. (2006), one can notice the author's reasoning about diversity as approximately a single term on the examples of several sciences, while Stirling (2007) devoted all his attention to the definition of diversity in many different sciences such as economics, ecology, agronomy, technology, engineering, etc. At the same time Kelly (2018) analysed the topic through the ethnic diversity of science. Nevertheless, there are still authors who define diversity through the prism of technology. They recognize that institutional diversity and technological diversity interact in different ways (Hughes, 1987; MacKenzie & Wajcman, 1985). As a result, these interactions constitute the breeding ground for technical diversity (Matthews & McGowan, 1992). Also, according to Matthews & McGowan (1992), diversity is widely regarded as a barrier to achieving the benefits of scale economies. This is because when technological efforts are divided into (potentially) diverse factions, the resources available for any one solution-avenue are diminished (Matthews & McGowan, 1992). Authors interpret "diversity" only in terms of product range, not variation in the processes that produce these items (Gatsios & Seabright, 1989).

Patil and Taillie (1977) define diversity as "the rarity of species" which they claim is an average property of a group. They go on to establish the stochastic ordering that is the same as the internal ordering of communities in terms of variety.

Schmidt, et al. (2006) mention such a term as biodiversity and define it as the number of species in a unit of study or "species richness". A framework for an interdisciplinary

investigation of diversity was presented in Stirling (2007). Stirling (2007) introduces three important properties of diversity, namely "variety", "balance" and "disparity" which together form a general concept of diversity relevant to any field of science (see Figure 1).

"Disparity" denotes the number of categories present in the system, while "balance" refers to the distribution of elements within each category. Lastly, "variety" describes the extent of variations between categories and how distinct they are from one another. When analysing the diversity of scientific fields, these elements of diversity are important to consider in order to gain a comprehensive understanding of the system and its constituent parts.

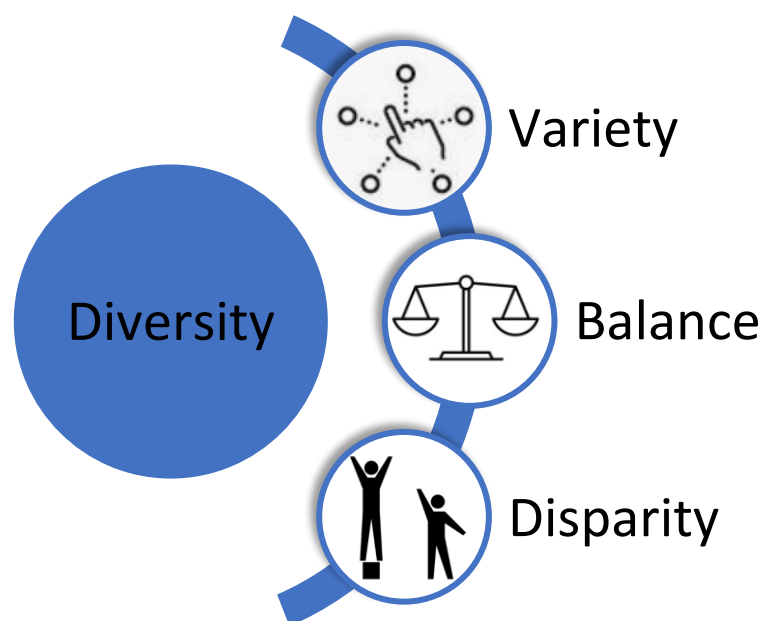


Figure 1. Definition of diversity by Stirling (2007)

Source: compiled by the author

Diversity offers a way to support pluralism, promote innovation, combat ignorance, and reduce insularity Stirling (2007). It presents a key tactic for acquiring the sustainability-essential traits of caution, fortitude, and resilience. This article proposes a common framework for comprehending diversity in a variety of various situations and specialties in order to accomplish these goals.

The study by Schmidt et al. (2006), however, focuses on the potential for using diversity metrics in a specific area and the identification of "species," or research fronts, to be used on both a national and worldwide scale while Kelly (2018) define diversity as "the effective number of ethnic types in a dataset". In other words, it is the number of ethnic

categories that must be evenly inhabited for the average fractional populations of the categories to match it in the actual dataset.

Table 1

Definitions of diversity in various sciences

Definition	Author; year	Field of research
Biological diversity is the variability among living organisms from all sources including, inter alia, terrestrial, marine and other aquatic systems and the ecological complexes of which they are part.	Magurran, 2004	Biology
Diversity is the variety and abundance of species in a defined unit of study.	Magurran, 2004 & Schmidt et al, 2006	Biology
Ecological diversity is the richness and variety of natural ecological communities.	Pielou, 1975	Ecology
Diversity is the number of species and their numerical composition.	Sander, 1968	Biology
Diversity within the population has meant sizeable affluent and impoverished segments, the "haves" and "have nots."	Matthews & Mcgowan, 1992	Economics
Diversity is a concept that is used in all sciences and contains 3 main components: "variety", "balance" and "disparity".	Stirling, 2007	Interdisciplinary

Source: compiled by the author

Since the thesis is focusing on comparing the diversity of science in different countries of the EU and UK, the definition by Stirling (2007) is the most appropriate for current work as it is universal and can be applied to different scientific fields.

1.2. Overview of different approaches for measuring diversity

There are several approaches how diversity can be measured. For instance, citation patterns and bibliometric approach can be used to assess the diversity of science. (Jarneving,

2005; Zitt, 2005; Nalimov & Mul'čenko, 1969; Hirv, 2018) This approach can help to identify biases and power dynamics within the scientific community (Swartz et al., 2019). The author concludes that research collaboration increases bibliometric visibility (Hirv, 2019).

Schmidt et al. (2006) created a system for evaluating scientific diversity applying a co-citation cluster analysis. The authors made the supposition that funding determines how diverse a field is. Resources get concentrated in the hands of specific groups of researchers as the most successful study is sponsored, which causes a decline in diversity. The authors used the concept of biodiversity to categorize authors of publications from the same subject by the same species in order to measure diversity in research. Thus, the number of species in a scientific unit (species' richness) and uniformity are two indicators of variety that are considered (variability in species and abundance).

Kelly (2018) focuses most of his research on the Shannon and Simpson indexes (see Table 2). The diversity index value is changed to the resultant expression. The same formula is used to determine actual diversity regardless of which diversity index is selected from the Shannon or Simpson family. Narrowing down the subject of research, Kelly (2018) focuses on ethnic diversity at Cambridge and Oxford universities over the previous ten years to address one of the most contentious issues in higher education: why ethnic diversity at Oxford and Cambridge universities has increased by more than 25% over the past ten years, but the issue of black students' under-enrolment in Britain persists.

Shannon's index, which is thought to stress the richness component of variety, and Simpson's index, which emphasizes the evenness component, are two of the most well-known measures used to quantify landscape composition (Nagendra, 2002). These indexes can vary significantly in response to variations in the richness and evenness of the environment (Nagendra, 2002).

The Herfindahl-Hirschman index is used to regulate market share in the US (US Department of Justice and the Federal Trade Commission, 1992). The higher the HI, the less competitive the market. It is computed as:

$$HI = \sum_{i=1}^N S_i^2 \quad (1)$$

where S is the sum of squares of market shares, N is the total number of market participants.

Probably the most frequently used index of diversity is the Shannon Index of diversity (SHDI) (Shannon & Weaver, 1949). It is defined as:

$$H = -\sum_{i=1}^N p_i \ln p_i \quad (2)$$

where N is the total number of types in the community and p_i is the proportion of N made up of the i^{th} type in the dataset, and $\ln p_i$ is the natural logarithm of p_i . The logarithms of the fractions (p) in natural form are evidently negative. Therefore, the 'minus' in the Shannon formula serves the purpose of adjusting the sum to yield a positive outcome. (Kelly, 2017)

It is also possible to calculate the Shannon Equitability Index. The closer the Shannon Equitability Index is to one, the more complete the equitability. It is computed as:

$$E_H = \frac{H}{\ln N} \quad (3)$$

The Simpson Index (SIDI) is also commonly used and defined as the likelihood that two randomly chosen, equally sized sub-units of the landscape correspond to different cover categories (Simpson, 1949; Nagendra, 2002). It is calculated as:

$$\lambda = \sum_{i=1}^N p_i^2 \quad (4)$$

This is the weighted arithmetic mean of the proportionate abundances p_i of the kinds, with the proportional abundances themselves serving as weights. Proportional abundances by definition have values between 0 and 1, but their weighted arithmetic mean, and hence λ , is obtained when all types are equally abundant. The greater the Simpson Index, the less diverse the population: $\lambda = 1$ represents infinite diversity and $\lambda = 0$ represents no diversity. (Kelly, 2017)

Simpson frequently employs transformations that result in an increase of the outcome as diversity increases. An example of such a transformation is the Inverse Simpson Index. (Kelly, 2017) It is defined as:

$$\lambda^{-1} = \frac{1}{\sum_{i=1}^N p_i^2} \quad (5)$$

where N is the total number of types and p_i is the proportion of N .

The Evenness (E) can be computed by dividing λ^{-1} by the maximum possible value λ^{-1} can have (λ^{-1}_{\max}) when items are uniformly distributed across all types, resulting in N . A value between 0 and 1 represents the degree of evenness or equitability, with 1 indicating complete evenness. (Kelly, 2017) It is calculated as:

$$E = \frac{\lambda^{-1}}{\lambda^{-1}_{\max}} = \left[\frac{1}{\sum_{i=1}^N p_i^2} \right] / N \quad (6)$$

The relative weights given to variety and balance can be affected by the logarithm base used in Shannon and the exponent value used in Simpson-Herfindahl (Hill, 1973). However, in actual practice, the implications for analysis are rarely examined. In other words, even widely used "non-parametric" indices like these are conceptually parameterized at a fundamental level even if it is implicitly. It is reasonable to change the values of these

fundamental parameters in order to provide different representations of variety. (Stirling, 2007)

Schmidt et al. (2006) applies the Simpson and Shannon indexes as well because they are the most appropriate to the field of the study. However, the authors claim that the Pearson coefficient, which is frequently utilized, has certain theoretical issues. It should only be used with data sets that are normally distributed, as opposed to bibliometric data, which are skewed, because it assesses the degree to which a linear function exists between two variables. Beyond this technical point, one may contend that normalizing the precise number of citations rather than measuring the existence of a linear relationship is a better way to gauge similarity than doing so directly.

With regard to the findings of the study, Stirling (2007) outlined a framework for an interdisciplinary investigation of diversity, listed several elements that have contributed to interest in diversity, and defined a broad framework for comprehending diversity in a variety of situations and specialties.

Diversity in education has increased over the past ten years as evidenced by the Simpson index's decline, which also represents a decline in diversity as measured by nationality (Kelly, 2018). When adopting the Shannon index, the comparable rise for Cambridge and Oxford is 22% and 30%, respectively. For the reasons already mentioned, this does not imply that Oxford is more diverse than Cambridge, but it does show that both institutions have significantly increased their diversity in the last ten years ago.

Schmidt et al. (2006), on the other hand, concluded that the single-link cluster method produces pretty disappointing outcomes. Additionally, there were issues with how applied diversity measures were interpreted. Disparities between two measurements at the site, the number of study fronts and the Shannon index, amply demonstrate the need for extreme caution while defining and quantifying variety. Although the number of study fronts is simpler to assess, it ignores how much research effort is allocated to each study front. The authors also concluded that the bibliometric approach to research diversity can possibly address these issues, making it a viable tool for scientific diversity research.

Table 2 presents a variety of indices for measuring diversity of science. Considering the relative population sizes of various types or species, a diversity index in general calculates the number of unique types or species present in a dataset (Colwell, 2009).

The number of types or the relative abundance of the various types both rise, which enhances diversity. A dataset's "richness" is measured by the variety of types it contains, and its "evenness" or "equitability" by the quantity of each type. Due to the fact that it also

considers the relative abundances of the various types, a diversity index offers more details on the composition of a community than richness alone. An ideal index should be relatively easy to compute, capable of discriminating effectively when diversity varies, and essentially unaffected by the size of the relevant dataset. (Kelly, 2017)

Table 2

Indexes for measuring diversity in many fields

Index	Field of research	Country	Object of measurement	Results
Shannon Index	Electrochemistry; ecology	France, Germany, Russia, UK and USA	Number of publications, variation in ethnicity, and one local education authority in London (Kelly 2017; Schmidt et al. 2006).	Shannon's equality index, which ranges from 0 to 1 and represents complete equality or disparity.
Renyi entropy			Shannon index generalization	
Simpson Index	Ecology; society	United Kingdom and USA	Concentration of types-classified specimens (Simpson 1949).	The diversity decreases as the index rises. D=0 denotes limitless diversity. D = 1 - No diversity
Gini-Simpson Index			Reformulation of the Simpson Index	
Hunter-Gaston Index			Form of the Simpson Index	
Herfindahl-Hirschman Index	Economy; education	United Kingdom	Evaluating the industry's organizations' level of competitiveness (Kelly 2017).	The market is "very competitive" if HI 0.01. The market is "unconcentrated" if HI = 0,01-0,15. The market is considered "moderately concentrated" if HI = 0,15-0,25. It is a "highly concentrated market" if HI > 0,25.
Gibbs-Martin Index	Psychology; sociology	United Kingdom, USA	The same as Gini-Simpson index.	$GM = 1 - \frac{\sum x^2}{(\sum x)^2}$ where X denotes the number of people employed in each industry.
Inverse Simpson Index	Social and natural sciences	United Kingdom, USA	Simpson's indexes are a way of quantifying the biodiversity of communities (Singh, 2022).	The scale is from 0 to 1, with 1 denoting complete equality.

Source: compiled by the author based on Eerma et al. (2022)

Indexes in Table 2 have certain advantages and disadvantages. Such as, for example, Simpson and Shannon Indexes are relatively easy to calculate, i.e. they do not require professional scientific knowledge and skills for their application, and they are widely used in many scientific fields. While the Herfindahl-Hirschman Index (HHI) is not as extensively used in ecology and biodiversity research as the Simpson or Shannon Diversity Indexes, it is a well-established tool for measuring market concentration and has recently been adapted to analyse biodiversity.

Another benefit of the Simpson Index is providing a complete measure of diversity. The index considers both the number of different species present in a community and the relative abundance of each species. In addition, because the index is highly sensitive to changes in the relative abundance of different species, it can be used to track changes in ecosystems through time. Overall, the Simpson Diversity Index is a great indicator for evaluating the diversity of different ecosystems and can be a useful tool for researchers, conservationists, and politicians alike. (Dong-June & Hong, 2011)

The Shannon Diversity Index also has a number of positive characteristics such as its suitability for samples with small sizes making it an excellent tool for studying rare or difficult-to-find species. The Shannon Diversity Index considers both the number of species present in a community and their relative abundance. This implies that it considers not just the number of species but also how they are distributed within the community. (Magurran, 2004)

These indexes are easy to calculate and the corresponding methodology for measuring the diversity of science does not require the resources of scientists' time, salary payments and any other costs. The simpler the calculation method, the easier and more profitable it is to work due to the fact that, for example, you do not need to hire a scientist with a doctorate who would conduct research and analyse the data obtained.

Nevertheless, above mentioned indexes also possess some disadvantages. All three indexes have one common important drawback, which is insensitivity to rare species. The Simpson, Shannon and Herfindahl-Hirschman Indexes are less sensitive to rare or uncommon species and more sensitive to changes in the abundance of dominating species. This implies that it might not fully represent the ecosystem's overall diversity. The Simpson and Shannon Indexes in turn have the disadvantage of being sensitive to sample size. The accuracy of the index can be affected by the sample size used to calculate it. Smaller sample sizes may result in less accurate estimates of diversity. (Barrantes & Sandoval, 2009)

In addition, The Shannon Index assumes that the community under study is a closed system with no species' immigration or emigration. This assumption may not apply to all ecosystems or communities. Overall, while the Shannon Index can be a useful tool for measuring diversity in certain contexts, it is important to consider its limitations and to use it in conjunction with other measures to gain a more comprehensive understanding of diversity.

The HHI, like the other diversity indexes, is limited to assessing diversity within a particular community or ecosystem and cannot be used to evaluate diversity across several communities or ecosystems. The HHI is less useful in communities with more than two dominating species. The requirement to include the market shares of all enterprises functioning in the market is the disadvantage of the HHI as well (Song et al., 2013).

In the following chapter, the author focuses on Shannon, Simpson and Herfindahl-Hirschman Indexes in order to evaluate diversity of science in different European countries based on statistical data from Eurostat and OECD.

2. Empirical analysis of diversity indexes for the EU countries and UK

2.1. Methodology of measuring diversity of science by the means of different indexes

The issue of science diversity is multifaceted and needs to be approached from various angles. This entails examining diversity in terms of the enrolment and graduation rates of doctoral students across various scientific fields, the volume of publications and citations, the presence of researchers, and the funding allocated to different fields. According to OECD, "researchers" can be defined as professionals who contribute to the development and generation of new knowledge, products, processes, methods, and systems, as well as the management of related projects (OECD, 2019). Annex A has been compiled to compare the share of researchers in the population of the European countries.

Based on Eurostat data, the proportion of researchers per capita has been calculated for the year 2019. Figure 2 illustrates variations in the number of researchers among 25 European countries using different colours. For the United Kingdom and France, data for 2019 is not available but the absence of information displayed on the map by the dash (-). In Figure 2 and Annex A, researchers are listed as a percentage of the total population of the country.

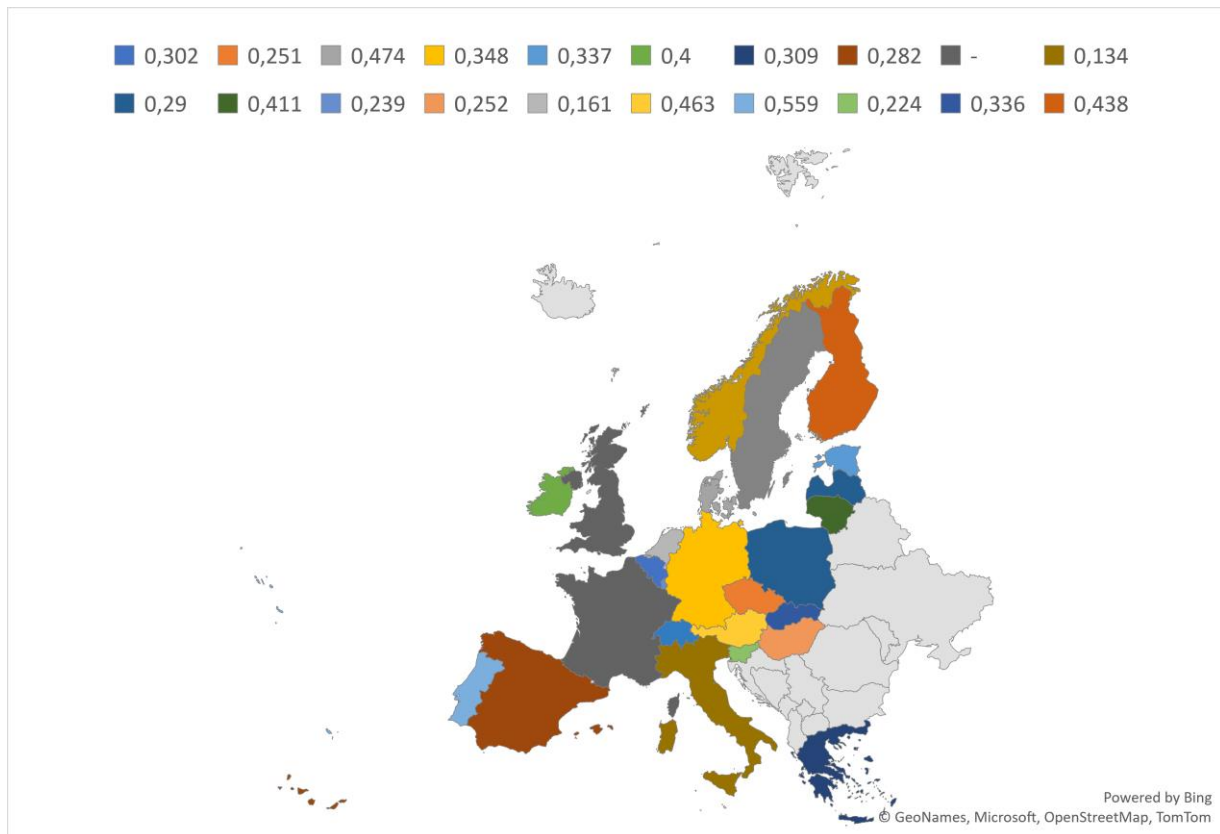


Figure 2. % of researchers in European countries

Source: Author's calculations based on Eurostat data (2019)

The highest the number of researchers per capita is in such countries as Switzerland (0.592%), Portugal (0.559%) and Norway (0.541%). It should be highlighted that although Portugal and Norway have similar percentages of researchers, Portugal has double the number of researchers and population size compared to Norway.

Among the countries with the smallest percentage of researchers are Italy (0.134% with a huge population of 59 816 673 people), Netherlands (0.161% with a population of 17 282 163 people) and Slovenia (0.224; population is 2 080 908 people). This indicates that these countries have relatively small number of researchers per capita. Thus, it is not enough to analyse only the number of researchers per capita. In order to have a deeper understanding of science diversity in the European countries, diversity metrics should be applied as well.

To measure diversity of science in different countries of Europe and UK, three main indexes highlighted in bold in Figure 3 are applied.

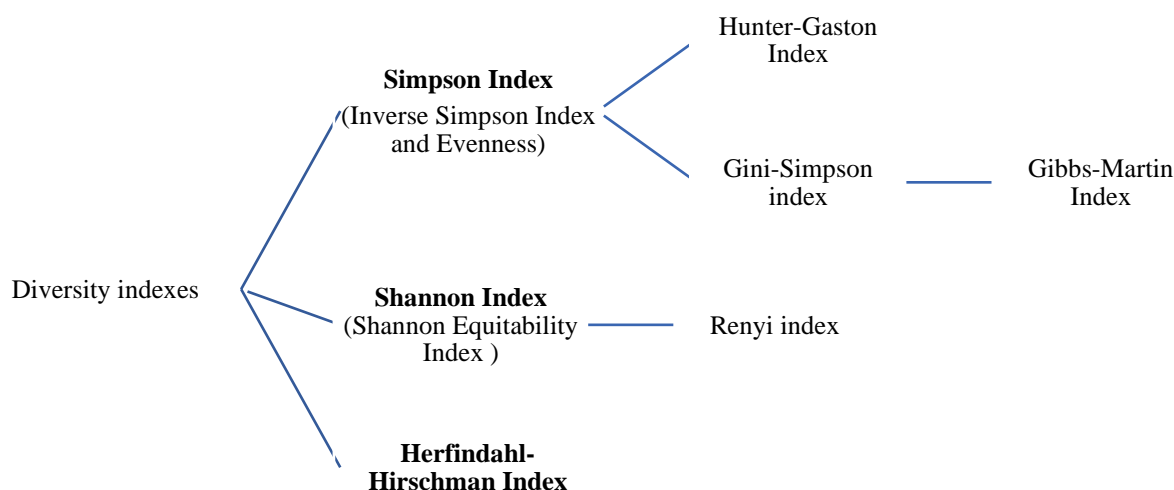


Figure 3. Classification of different diversity indexes

Source: compiled by the author

The Simpson index, Shannon index and Herfindahl-Hirschman indexes have been calculated applying data from Eurostat and OECD statistical databases. The Herfindahl index can be calculated as the sum of squares of market shares, assessing concentration and variety. The higher the HI, the less market competition there is.

The Shannon Diversity Index is calculated by taking the number of each species, the proportion each species is of the total number of individuals, and sums the proportion times the natural log of the proportion for each species (Callahan & Nolan, 2006). Also, indicating a more complete equivalence is the Shannon Equitability Index, which is closer to one (Colwell, 2009).

The Simpson Index can be used to evaluate biodiversity, as well as to assess population diversity disparities in schools, communities, and other settings. One of the index's more useful features is the ability to compare two sets of data to determine which is more diversified (Stephanie, 2021). The Simpson Index score ranges from 0 to 1. A high score implies a high level of diversity, whereas a low score suggests a low level of diversity. The community comprises only one species when the diversity index is 0 (i.e., no diversity). The diversity index rises and approaches one as the number of different species increases and species population distributions become more even. (Singh, 2022)

The study is based on classification which includes 11 scientific fields, which are presented in Table 3. The International Standard Classification of Education (ISCED) is used by OECD and Eurostat for the purposes of gathering statistical information.

Table 3

Fields of education and training according to the International Standard Classification of Education (ISCED)

The name of the scientific field
Generic programmes and qualifications
Education
Arts and humanities
Social sciences, journalism and information
Business, administration and law
Natural sciences, mathematics and statistics
Information and Communication Technologies
Engineering, manufacturing and construction
Agriculture, forestry, fisheries and veterinary
Health and welfare
Services

Source: Eurostat

Figure 4 shows the received data for 2019 on the Inverse Simpson Index and Evenness based on the number of doctoral students enrolled in 11 fields. Simpson (1949) defined the Simpson Index (SI) as the degree of grouping of individuals across species. Simpson frequently employs Inverse Simpson as a modification to raise the value of the index as variety increases (Kelly 2017).

The evenness determined from the Inverse Simpson Index ranges between zero and one, with one indicating total evenness. Portugal has the highest evenness among other countries (0.706). Denmark has the smallest evenness with the value (0.425) which means that in Denmark compared to Portugal there are some scientific fields that have more students doctoral students enrolled than in others.

The diversity indexes are computed using OECD statistics based on not only the number of doctoral students enrolled in doctoral studies, but also the number of doctoral graduates by year and field of research. The OECD statistics on graduates include the same fields as in Table 4.

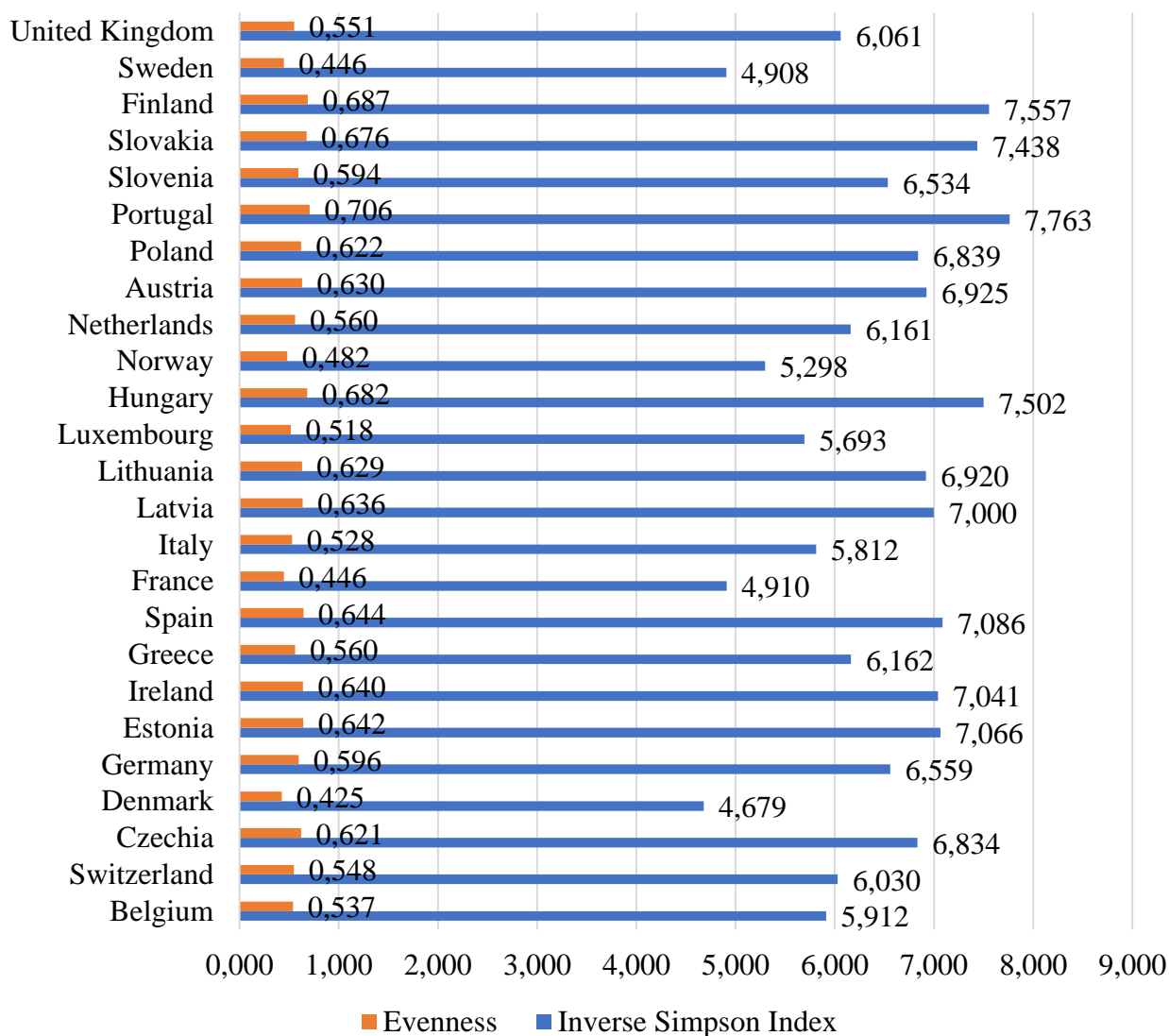


Figure 4. Inverse Simpson Index and Evenness (based on the number of doctoral students enrolled)

Source: compiled by the author based on Eurostat data 2019

In Figure 5 presents the results of the Inverse Simpson Index and Evenness based on the number of doctoral students graduated. Some changes in the evenness values can be reported. For example, such countries as Portugal (0.688), Slovakia (0.686), Poland (0.631) and Finland (0.621) have the greatest evenness while Denmark (0.405), Sweden (0.424) and Luxembourg (0.429) have the lowest values.

For example, the Netherlands and Norway have quite similar values of indexes for graduated students (0.449 and 0.438 respectively). When it comes to the students enrolled the numbers grow (0.560 and 0.482 respectively). This means that during the enrolment stage for doctoral students at universities, they are more evenly distributed across scientific fields compared to the end of their studies.

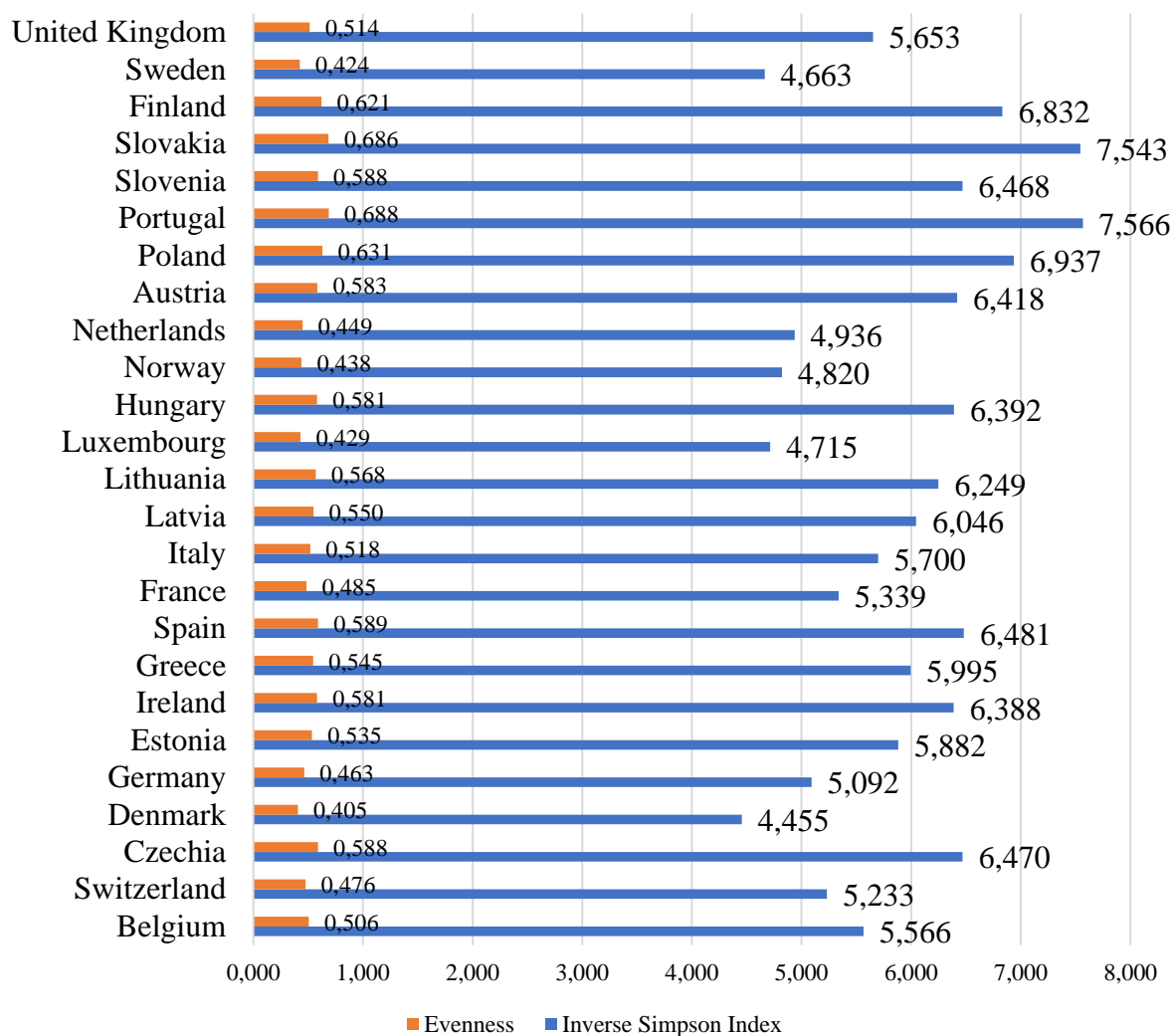


Figure 5. Inverse Simpson Index and Evenness (based on the number of doctoral students graduated)

Source: compiled by the author based on Eurostat data 2019

The Shannon Equitability Index, also known as the Evenness Index, is a metric used to quantify the degree of evenness or balance in the distribution of values across distinct categories or groups (Colwell, 2009). In the context of comparing student enrolment in various fields of science across countries, the Shannon Equitability Index can be used to assess the degree to which enrolment is uniformly distributed throughout distinct fields of science within each country. It is also possible to compute the Shannon Equitability Index.

The resulting Shannon Equitability Index ranges from 0 to 1, with 1 meaning perfect evenness or balance in the distribution of students across fields of science within each country. Interpreting the Shannon Equitability Index results for different countries might provide insights into the degree of evenness or balance in student enrolment in various fields of science within each country. A higher index value indicates that student enrolment is more

evenly spread among fields of science, whereas a lower value indicates that enrolment is more biased towards certain fields. By comparing the index values across nations, it is possible to establish patterns or trends in the distribution of students across scientific subjects and evaluate the effectiveness of fields aimed at fostering more balanced enrolment.

In Figure 6 the results of calculations of the Shannon Equitability Index based on the number of doctoral students enrolled in different scientific fields are shown. Denmark has the lowest value on the Shannon Equitability Index (0.692) which means that Denmark has a slightly weaker evenness distribution of doctoral students in scientific fields compared to other countries. The highest equitability is present in such countries as Hungary (0.890), Portugal (0.889), Slovakia (0.883), Finland (0.881) and Estonia (0.873).

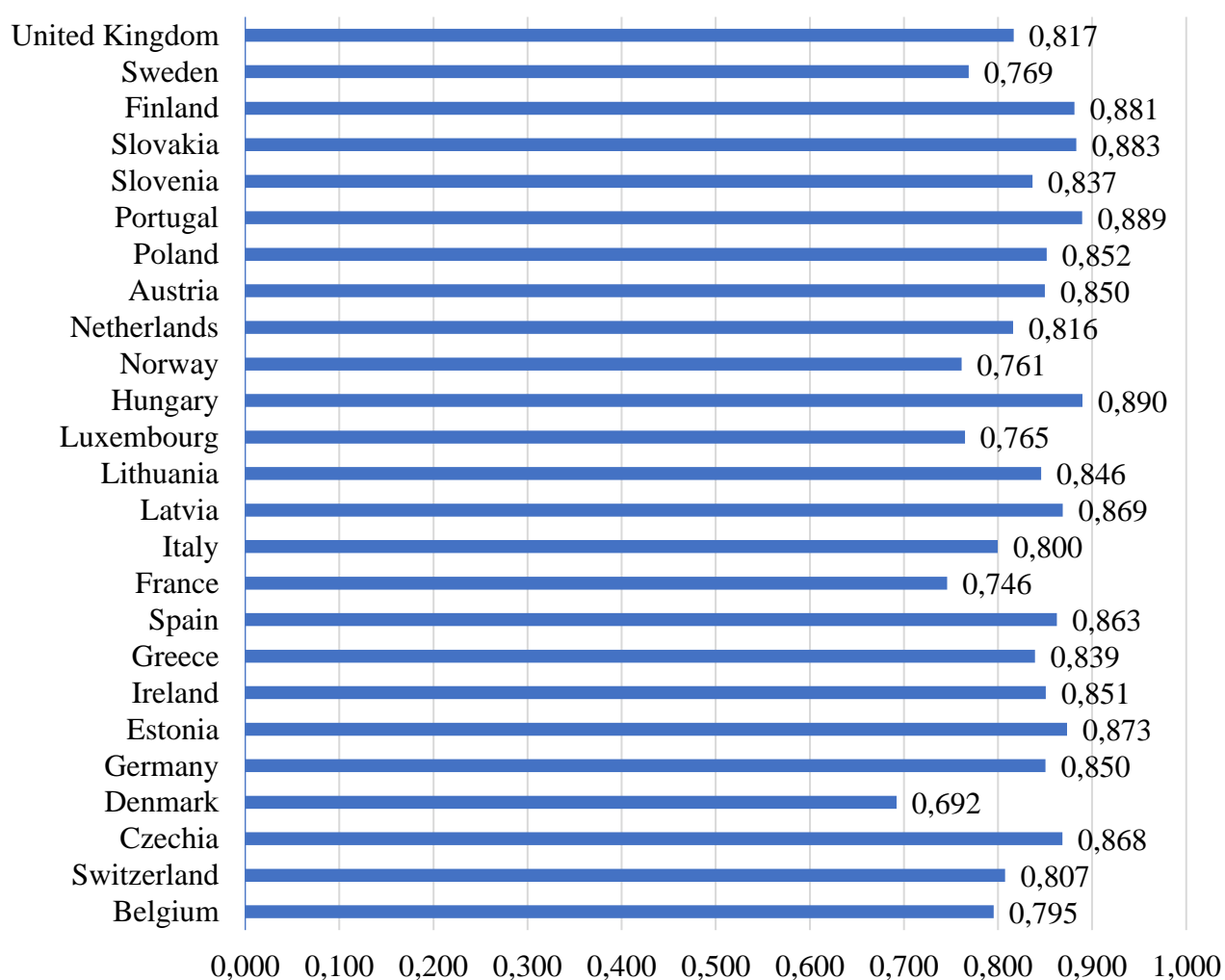


Figure 6. Shannon Equitability Index (based on the number of doctoral students enrolled)

Source: compiled by the author based on Eurostat data 2019

Figure 7 presents the outcomes of the computation of the Shannon Equitability Index utilizing the count of doctorate students who have completed doctoral (or equivalent)

programs in various fields. Denmark still has the lowest value on the Shannon Index (0.692). However, in France, for example, the diversity index increased from 0.746 to 0.792, which means that in France, diversity in scientific fields is greater among students who graduated from university than among those who enrolled in the program.

Conversely, in Germany, the Shannon Index reveals that there exists more diversity in the distribution of students across scientific fields at the stage of university admission (0,850), whereas this variability appears to decrease upon graduation from these programs (0,775).

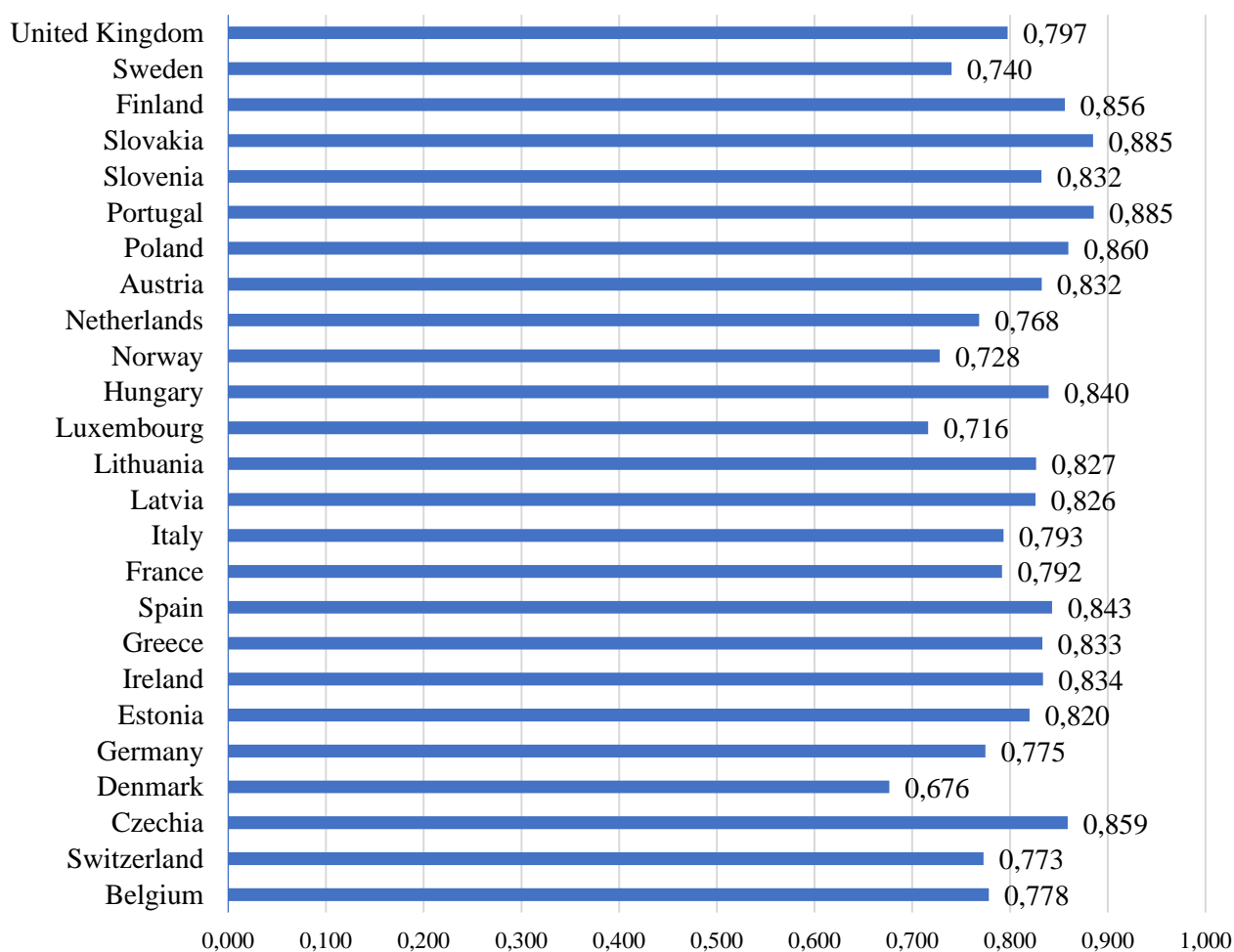


Figure 7. Shannon Equitability Index (based on the number of doctoral students graduated)

Source: compiled by the author based on Eurostat data 2019

Figure 8 depicts the findings of the computations of the Herfindahl-Hirschman Index, which is based on the number of doctoral students who are registered in distinct doctoral programs.

The Herfindahl-Hirschman Index (HHI) is a market concentration metric that is often used in economics to assess the level of rivalry within an industry. The HHI can be used to

examine the concentration of doctoral students enrolled in various scientific subjects in the context of countries and their research output.

A country with a high HHI value in a certain scientific subject has a relatively concentrated enrolment of doctoral students in specific subfields within that scientific field. This could imply that the country takes a more concentrated or specialized approach to doctoral education in that scientific field.

A low HHI value, on the other hand, would imply that the country has a more diverse enrolment of doctoral students in that scientific field across different subfields. This may imply that the country takes a more balanced approach to doctoral education in that scientific field.

The Herfindahl Index can be used to interpret the level of competition within a market according to the US Department of Justice's Antitrust Division's Merger Guidelines. The interpretation criteria are the following:

- A market with an HI value of less than 0.01 is considered to be highly competitive.
- An HI value between 0.01 and 0.15 indicates an unconcentrated market.
- An HI value between 0.15 and 0.25 indicates a moderately concentrated market.
- An HI value greater than 0.25 indicates a market with a high level of concentration.

Based on this classification, 25 countries can be divided into 2 groups: unconcentrated market and moderately concentrated market since the values less than 0.01 (meaning high competition) or more than 0.25 (mean a high level of market concentration) were not obtained.

The first category with values between 0.01 and 0.15 includes countries: Portugal (0.129), Finland (0.132), Hungary (0.133), Slovakia (0.134), Spain (0.141), Ireland and Estonia with the same value 0.142, Latvia (0.143), Austria (0.144), Lithuania (0.145), Poland and Czechia with the same value 0.146.

The second category with values between 0.15 and 0.25 includes countries: Germany (0.152), Slovenia (0.153), Netherlands and Greece with the same value 0.162, United Kingdom (0.165), Switzerland (0.166), Belgium (0.169), Italy (0.172), Luxembourg (0.176), Norway (0.189), Sweden and France with the value 0.204, Denmark (0.214).

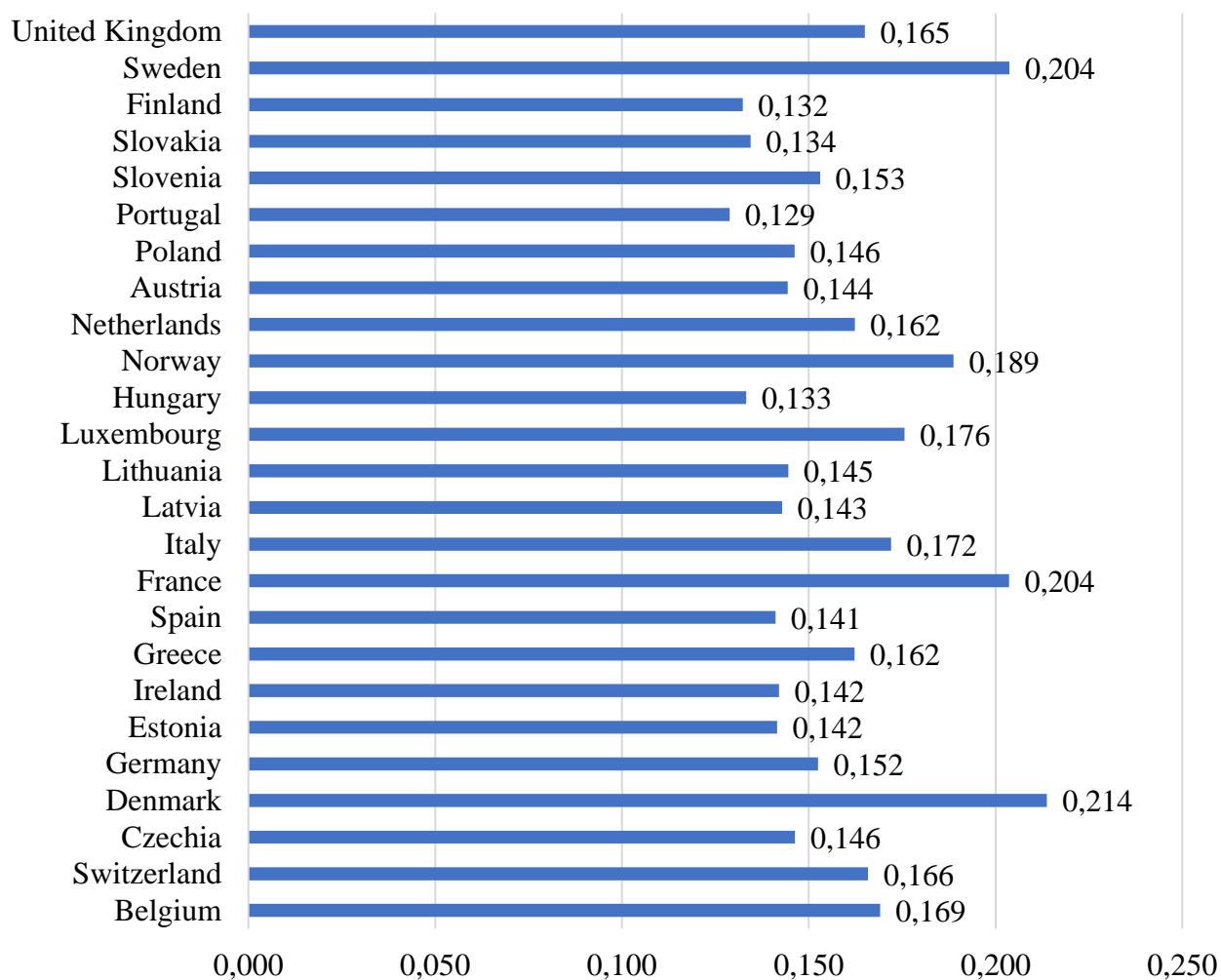


Figure 8. Herfindahl-Hirschman Index (based on the number of doctoral students enrolled)

Source: compiled by the author based on Eurostat data 2019

Figure 9 depicts the results of the computations of the Herfindahl-Hirschman Index, which is based on the count of doctoral students who have completed their doctoral programs. Countries such as Portugal (0.132), Slovakia (0.133), Poland (0.144) and Finland (0.146) fall in the category of unconcentrated market, which consists of values ranging between 0.01 and 0.15. The graduation of doctoral students is relatively evenly distributed across different scientific fields.

Countries such as Spain (0.154), Czechia and Slovenia (both at 0.155), Austria and Hungary (both at 0.156), Ireland (0.157), Lithuania (0.160), Latvia (0.165), Greece (0.167), Estonia (0.170), Italy (0.175), United Kingdom (0.177), Belgium (0.180), France (0.187), Switzerland (0.191), Germany (0.196), Netherlands (0.203), Norway (0.207), Luxembourg (0.212), Sweden (0.214) and Denmark (0.224) belong to the category of a moderately concentrated market, encompassing values ranging from 0.15 to 0.25. While there may be

some variation in student graduation among different fields, there is no dominant subfield or topic that stands out considerably.

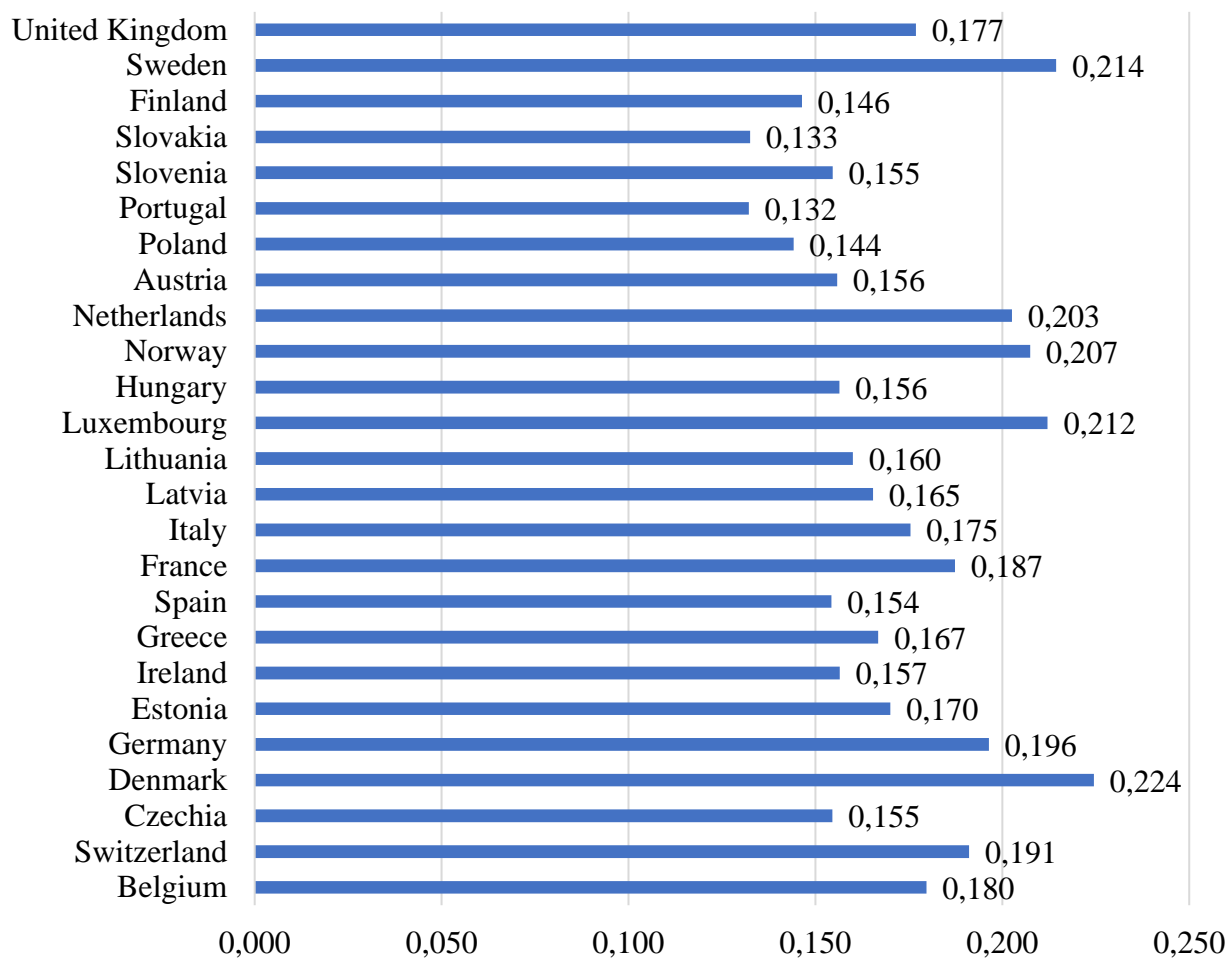


Figure 9. Herfindahl-Hirschman Index (based on the number of doctoral students graduated)

Source: compiled by the author based on Eurostat data 2019

The subsequent chapter highlights the author's interpretation of the results derived from Shannon, Simpson, and Herfindahl-Hirschman Indexes. The focus of this analysis is to deeply explore and explain the reasons for received values of indexes.

2.2. Discussion of empirical results

Table 4 provides a comprehensive overview of the diversity of science across various European countries, as measured by the Simpson, Shannon Equitability and Herfindahl-Hirschman Indexes. By examining the values of these indexes for each country, it becomes possible to identify overall trends and patterns in the diversity of scientific fields across Europe. This information can be particularly useful for policymakers and stakeholders who are interested in improving the diversity of scientific fields in their countries. Through a

careful analysis of these indexes, it may be possible to identify areas where additional support and resources are needed in order to promote greater diversity and inclusivity in the scientific community.

Table 4

Results of the diversity indexes

	Simpson Index (Evenness)		Shannon Equitability Index		Herfindahl-Hirschman Index	
	enrolled	graduated	enrolled	graduated	enrolled	graduated
Belgium	0,537	0,506	0,795	0,778	0,169	0,180
Switzerland	0,548	0,476	0,807	0,773	0,166	0,191
Czechia	0,621	0,588	0,868	0,859	0,146	0,155
Denmark	0,425	0,405	0,692	0,676	0,214	0,224
Germany	0,596	0,463	0,850	0,775	0,152	0,196
Estonia	0,642	0,535	0,873	0,820	0,142	0,170
Ireland	0,640	0,581	0,851	0,834	0,142	0,157
Greece	0,560	0,545	0,839	0,833	0,162	0,167
Spain	0,644	0,589	0,863	0,843	0,141	0,154
France	0,446	0,485	0,746	0,792	0,204	0,187
Italy	0,528	0,518	0,800	0,793	0,172	0,175
Latvia	0,636	0,550	0,869	0,826	0,143	0,165
Lithuania	0,629	0,568	0,846	0,827	0,145	0,160
Luxembourg	0,518	0,429	0,765	0,716	0,176	0,212
Hungary	0,682	0,581	0,890	0,840	0,133	0,156
Norway	0,482	0,438	0,761	0,728	0,189	0,207
Netherlands	0,560	0,449	0,816	0,768	0,162	0,203
Austria	0,630	0,583	0,850	0,832	0,144	0,156
Poland	0,622	0,631	0,852	0,860	0,146	0,144
Portugal	0,706	0,688	0,889	0,885	0,129	0,132
Slovenia	0,594	0,588	0,837	0,832	0,153	0,155
Slovakia	0,676	0,686	0,883	0,885	0,134	0,133
Finland	0,687	0,621	0,881	0,856	0,132	0,146
Sweden	0,446	0,424	0,769	0,740	0,204	0,214
United Kingdom	0,551	0,514	0,817	0,797	0,165	0,177

Source: Compiled by the author based on OECD and Eurostat data (2019)

In general, there appears to be a trend of declining values in the Simpson (Evenness) and Shannon Equitability Indexes from the time of admission to university to the point of graduation. To provide an example, we can examine a selection of countries such as Denmark, Portugal, Norway and Finland. The Shannon Equitability Index for Finland in 2019 revealed an enrolment value of 0.817, which declined to 0.797 at the time of graduation from university. Similarly, other selected countries also exhibited a decrease in diversity.

This trend suggests that students enter university with a wide range of scientific interests, but then either switch to other fields or drop out. Such processes can have a negative impact on the diversity of scientific fields, resulting in declining values for the indexes.

This explanation also applies in the opposite direction, where the diversity index may start out lower and increase by graduation, as observed in countries like France, Poland, and Slovakia. This trend may suggest that students initially have a strong interest in popular and in-demand fields, but then shift their focus to other programs during the course of their studies. As a result, the diversity index increases over time.

However, it is important to note that the Shannon Equitability Index only considers the distribution of students across scientific fields within each country and does not consider other factors that may affect the quality of education or students' ability to pursue careers in their chosen field. As a result, additional measures may be required to adequately assess the success of education policies and practices in fostering equitable access to education and job prospects in many sectors of science. (Rafols & Meyer, 2010)

In contrast to the Shannon and Simpson indexes, the Herfindahl-Hirschman Index has an opposite interpretation. A higher value of the index indicates a lower level of diversity. Specifically, as the index approaches 1, it signifies a greater concentration of doctoral students in fewer scientific fields. In other words, a high Herfindahl-Hirschman Index implies that there is a dominant field or a small group of fields with a larger share of doctoral students compared to other fields. (Carley & Porter, 2012)

It should be noted that the interpretation of the HHI values is also dependent on the context and aims of the investigation. For example, a country with a high HHI value in a given scientific subject may have purposefully focused its resources and efforts on specific subfields or issues within that field in order to become a leader in those areas. As a result, the HHI should be used in conjunction with other metrics and qualitative research to provide a more complete picture of a country's scientific output and doctorate education.

Furthermore, to support previous arguments, Figure 10, Figure 11, Appendix B and Appendix C have been created to demonstrate the distribution of students across different scientific fields. Figure 10 displays 4 countries (Denmark, France, Sweden and Norway) with the lowest Inverse Simpson Index and Evenness based on the number of students enrolled in fields. Scientific fields such as Natural sciences, Mathematics, statistics and Health, welfare are more popular among students and tend to have a higher concentration, while others have very few students. As a result, there is a low level of diversity in science in these four countries.

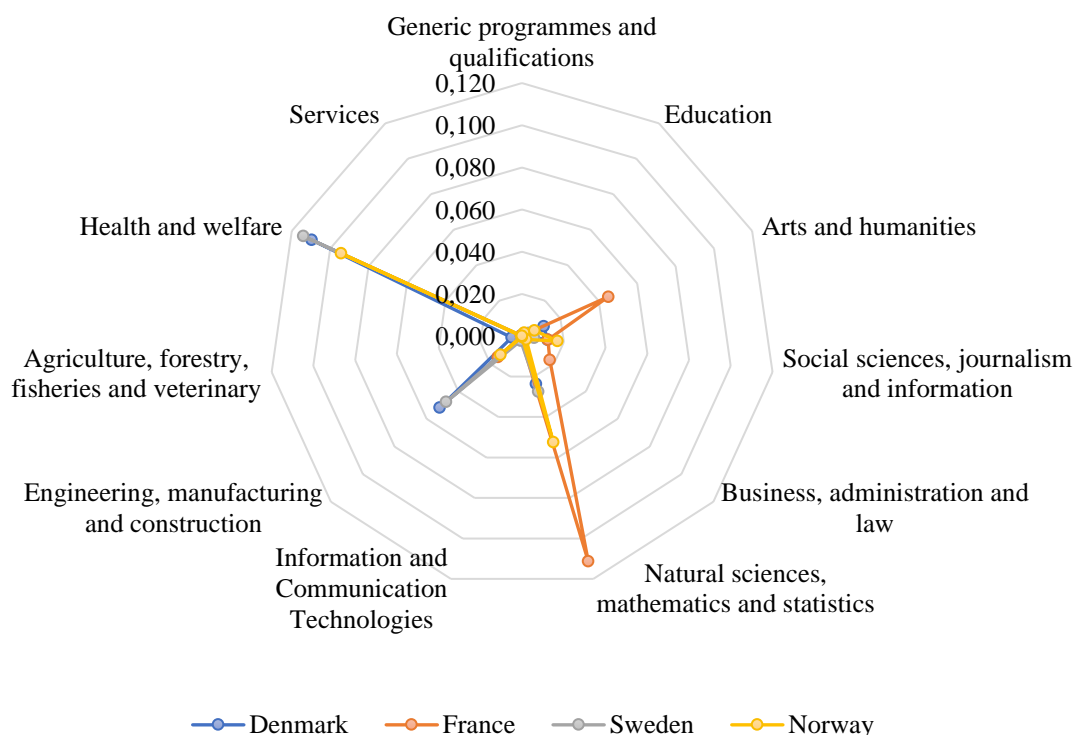


Figure 10. Results of the Simpson Index in each scientific field (based on doctoral students enrolled)

Source: Eurostat data, compiled by the author

The differences in the distribution of fields of science between countries highlighted in Figure 10 and Figure 11 can provide valuable insights into the unique characteristics of each country's scientific landscape. As illustrated in Figure 11, Poland appears to have a greater emphasis on Arts and Humanities, as well as Natural Sciences, Mathematics, and Statistics, with a slightly lower proportion of students graduating in Health and Welfare and Engineering, Manufacturing, and Construction fields.

Conversely, Finland seems to have a higher proportion of students graduating in Health and Welfare and Engineering, Manufacturing, and Construction fields, indicating a greater emphasis on these areas of study. These differences in emphasis may reflect the unique economic and social priorities of each country, as well as differences in academic and research strengths.

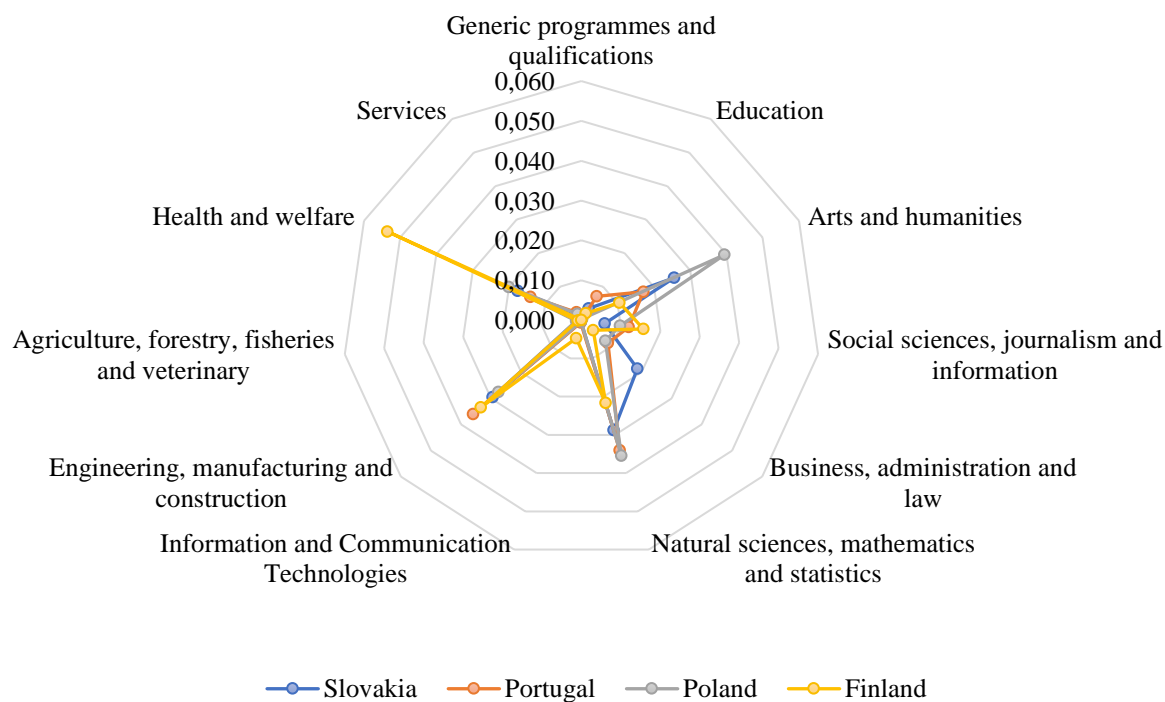


Figure 11. Results of the Simpson Index in each scientific field (based on doctoral students graduated)

Source: Eurostat data, compiled by the author

The provided data allows for monitoring the evolution of scientific diversity in various European countries and comprehending the factors contributing to the high or low diversity index in the selected countries.

The Amable classification system divides academic areas into five categories: physical sciences, biological sciences, engineering sciences, social sciences, and humanities. For example, if the Simpson Index for enrolled students is high in the physical sciences category, this indicates a lack of diversity and suggests that attempts to improve participation from other groups may be necessary. Similarly, a high Shannon Index for graduated students in the humanities indicates increased diversity and suggests that the discipline is attracting a larger spectrum of individuals. (Amable, 2003)

Overall, understanding data on diversity indices based on the Amable classification necessitates careful consideration of context and factors that may contribute to diversity discrepancies. Better knowledge of the diversity of academic subjects and populations by evaluating and comparing these indexes across categories, institutions, and locations should be acquired, measures to encourage greater diversity and inclusion can be suggested. (Amable, 2003)

Conclusion

Diversity is an important concept for developing the precautionary, resilient, and robust features that are essential for sustainability. To that purpose, the study identifies a broad framework for comprehending diversity. This entails viewing diversity as a function of three required properties: variety, balance, and disparity.

Throughout this bachelor thesis, the concept of diversity has been thoroughly explored and different methods for measuring diversity have been examined providing insights into the advantages and disadvantages of each metrics.

The current science policy's goal of excellence may jeopardize scientific diversity, leading to blind spots and prejudices that impede progress in addressing difficult scientific concerns. A more diverse scientific community is more likely to provide novel ideas, methodologies, and viewpoints that can aid in the resolution of these issues. As a result, developing analytical methodologies for measuring the extent of scientific variety in various countries is critical.

The thesis proposes a quantitative approach for measuring diversity of science by applying Simpson, Shannon and Herfindahl diversity indexes. The thesis calculates the diversity of scientific fields in the EU and UK using specified indexes based on the number of students enrolled and graduating at the doctoral/equivalent level.

The study's empirical findings indicate that the Simpson, Shannon, and Herfindahl indexes can be effective tools for measuring the diversity of scientific research in European countries and the UK. However, it's important to note that the reliability of the data collected is somewhat limited due to the study's failure to account for various factors that impact diversity measurement. The author acknowledge that the lack of available data on these factors represents a significant limitation, and caution that any conclusions drawn from the research must be considered with this in mind.

Once the calculations have been performed and the data has been compiled into various graphs and tables, an assessment of the overall diversity of scientific fields in different countries can be made. By analyzing the data, it becomes apparent that certain countries, such as Portugal, Finland, Slovakia, and Poland, exhibit high levels of diversity across all indices. Conversely, countries such as Denmark, France, Sweden, and Norway show the lowest levels of diversity in science.

Overall, the analysis did not reveal any countries with extremely low or high levels of diversity in science. However, it's important to note that this does not necessarily indicate that all countries have achieved optimal levels of diversity in their scientific fields. There may

still be room for improvement in increasing representation from underrepresented groups and promoting greater inclusivity in science.

The directions for future research are driven by the limitations of the thesis. First of all, there is lack of European level dataset on the number of researchers by fields (the one that is similar to the doctoral students and graduates by fields), thus, development of such a dataset would make the analysis of diversity of science more complete. Furthermore, the thesis analyses diversity of science based on the number of doctoral students (enrolled and graduated). However, some other crucial aspects such as the number of publications and citations by fields of research, funding by fields are not investigated in this thesis. Due to unavailability of data on international level, the latter aspect can be studied only on national level for each country separately. Then, comparability of results is questionable since countries might have different funding schemes. A unified approach to the data collection should be developed.

References

1. Adams, J. and D. Smith, (2003). Funding research diversity: The impact of further concentration on university research performance and regional research capacity, A report to Universities UK by Evidence Limited.
2. Amable, B. (2003). The Diversity of Modern Capitalism. In Oxford University Press eBooks. <https://doi.org/10.1093/019926113x.001.0001>
3. Barrantes, G., Sandoval, L. (2009) Conceptual and statistical problems associated with the use of diversity indices in ecology, *Revista de Biología Tropical*, 57(3): 451-60.
4. Blume, M. E., & Friend, I. (1975). The Asset Structure of Individual Portfolios and Some Implications for Utility Functions. *The Journal of Finance*, 30(2), 585–603.
5. Carley, S., Porter., A. L. (2012). A forward diversity index. *Scientometrics*, 90: 407-427
6. Carrell, M., Mann, E. and Sigler, T. (2006) Defining workforce diversity programs and practices in organizations: a longitudinal study. *Labor Law Journal*, 57(1): 5-12.
7. Colwell, R. K. (2009). III.1 Biodiversity: Concepts, Patterns, and Measurement. *The Princeton Guide to Ecology*, 257–263. <https://doi.org/10.1515/9781400833023.257>
8. Dong-June, Y., Ji Hong, K. (2011). Comparative evaluation of species diversity indices in the natural deciduous forest of Mt. Hood. *Jeombong. Forest Science and Technology*, 7(2): 68-74.
9. Eerma, D., Chebotareva, M., Espenberg, S., Hirv, T., Pavlov, D., Urmann, H., Remmik, M. (2022). Teaduse valdkondliku mitmekesisuse määratlemine, selle seire meetodika ja mõõdiku väljatöötamine. Tartu: Tartu Ülikool.
10. Gatsios, K. & Seabright, P. (1989). Regulation in the European community. *Oxford Rev. Econ. Policy* 5, 37–60. (doi:10. 1093/oxrep/5.2.37)
11. Graduates by field. (2019). OECD. Retrieved from https://stats.oecd.org/Index.aspx?DataSetCode=EDU_GRAD_FIELD
12. Harley, S. and F.S. Lee, (1997). Research Selectivity, Managerialism and the Academic Labor Process: The Future of Nonmainstream Economics in UK Universities. In *Human Relations*, 50, pages 1427-1460.
13. Hill, M. O. (1973). Diversity and Evenness: A Unifying Notation and Its Consequences. *Ecology*, 54(2): 427-432.
14. Hirv, T. (2018). Effects of European Union Funding and International Collaboration on Estonian Scientific Impact. *Journal of Scientometric Research*, 7(3): 181-188.
15. Hirv, T. (2019). Research consortia determine a significant part of the bibliometric visibility of Estonian science. *Trames*, 23: 287-308.

16. Hughes, Thomas P. (1987). The evolution of large technological systems. In *The social construction of technological systems. New directions in the sociology and history of technology*, edited by W. E. Bijker, T. P. Hughes and T. Pinch. Cambridge, Massachusetts & London, England: MIT Press, 51-82.
17. Hurlbert, S.H. (1971). The Nonconcept of Species Diversity: A Critique and Alternative Parameters. *Ecology*, 52(4): 577-586.
18. Jarneving, B. (2005). A Comparison of two Bibliometric Methods for Mapping of the Research Front. *Scientometrics*, 65(2): 245-263.
19. Kelly, A. (2017) *Developing Metrics for Equity, Diversity and Competition. New measures for schools and universities.* Routledge.
20. Kelly, A. (2018). A new composite measure of ethnic diversity: Investigating the controversy over minority ethnic recruitment at Oxford and Cambridge universities. *British Educational Research Journal*, 45(1), 41–82. <https://doi.org/10.1002/berj.3482>
21. Lacey, R. E., Dolan, E. L., & Kalinowski, S. T. (2019). Cultural diversity and sustainability science: A synthesis for action. *Sustainability Science*, 14(5), 1345-1357.
22. Leydesdorff, L., Wagner, C. S., Bornmann, L. (2019). Interdisciplinarity as diversity in citation patterns among journals: Rao-Stirling diversity, relative diversity, and the Gini coefficient. *Journal of Informetrics*, 13: 255-269.
23. MacKenzie, D., & Wajcman, J. (1985). *The Social Shaping of Technology* (0 ed.). Open University Press.
24. Magurran, A.E., (2004). *Measuring Biological Diversity.* Oxford, Blackwell.
25. Matthews, M., & McGowan, F. (1992). Reconciling diversity and scale: some questions of method in the assessment of the costs and benefits of European integration.
26. Nagendra, H. (2002). Opposite trends in response for the Shannon and Simpson indices of landscape diversity. *Applied Geography*, 22(2), 175–186. [https://doi.org/10.1016/s0143-6228\(02\)00002-4](https://doi.org/10.1016/s0143-6228(02)00002-4)
27. Nalimov, V. V., Mul'chenko, Z. M. (1969). *Naykometry. [Seismometry].* Moscow. (Nalimov, B. V., Mulchenko, Z. M. (1969). Moscow: Nauka).
28. Nolan, K.A. & Callahan J.E. (2006). Beachcomber biology: The Shannon-Weiner Species Diversity Index. Pages 334-338, in *Tested Studies for Laboratory Teaching, Volume 27* (M.A. O'Donnell, Editor). Proceedings of the 27th Workshop/Conference of the Association for Biology Laboratory Education (ABLE), 383 pages.
29. Patil, G. P., Taillie, C. (1977). Diversity as a Concept and its Measurement. *Journal of the*

- American Statistical Association*, 77(379): 548-561.
30. Pielou, E. C. (1975) Ecological diversity. New York: Wiley InterScience. *Revue D'économie Industrielle*, 59(1), 222–234. <https://doi.org/10.3406/rei.1992.1416>
 31. Rafols, I., Meyer, M. (2010) Diversity and network coherence as indicators of interdisciplinarity: case studies in bionanoscience. *Scientometrics*, 82: 263-287.
 32. Rosenberg, N. (1996) Uncertainty and technological change. In *The mosaic of economic growth* (eds R. Landau, T. Taylor & G. Wright). Stanford, CA: Stanford University Press.
 33. Sander, H. L. (1968). Marine Benthic Diversity: A Comparative Study. *The American Naturalist*, 102 (925): 243–82.
 34. Schmidt, M., Glaser, J., Havemann, F., Heinz (2006). A Methodological Study for Measuring the Diversity of Science. *Proceedings International Workshop on Webometrics, Informetrics and Scientometrics & Seventh COLLNET Meeting, Nancy (France)* 129-137.
 35. Shannon, C. E., & Weaver, W. (1949). *The Mathematical Theory of Communication*. Urbana, IL: The University of Illinois Press, 1-117.
 36. Simpson, E.H. (1949) Measurement of Diversity. *Nature*, 163, 688.
<http://dx.doi.org/10.1038/163688a0>
 37. Singh, P. (2022, November 1). *Simpson's Diversity Index Calculator*. Retrieved from <https://www.omnicalculator.com/statistics/simpsons-diversity-index>
 38. Song, Y., Merlin, L., Rodriguez, D. (2013) Comparing measures of urban land use mix. *Computers, Environment and Urban Systems*, 42: 1-13.
 39. Spangenberg, J. H. (2011). Sustainability science: a review, an analysis and some empirical lessons. *Environmental Conservation*, 38(3), 275–287.
<https://doi.org/10.1017/s0376892911000270>
 40. Stephanie, S. (2021, May 24). Simpson's Diversity Index: Definition, Formula, Calculation. *Statistics How To*. Retrieved from <https://www.statisticshowto.com/simpsons-diversity-index/>
 41. Stilgoe, J., & Guston, D. H. (2017). Responsible innovation and the governance of science: The need for policy coherence. *Sustainable Science*, 12(1), 149-161.
 42. Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society Interface*, 4(15), 707–719.
<https://doi.org/10.1098/rsif.2007.0213>

43. Swartz, T. H., Palermo, A. S., Masur, S. K., & Aberg, J. A. (2019). The Science and Value of Diversity: Closing the Gaps in Our Understanding of Inclusion and Diversity. *The Journal of Infectious Diseases*, 220(Supplement_2), S33–S41.
<https://doi.org/10.1093/infdis/jiz174>
44. Total researchers by sectors of performance - head count. (2019). Eurostat. Retrieved from <https://ec.europa.eu/eurostat/databrowser/view/tsc00003/default/table?lang=en>
45. US Department of Justice and the Federal Trade Commission. (1992) *Journal of the American Statistical Association*.
46. Zitt, M. (2005) Facing Diversity of Science: A Challenge for Bibliometric Indicators. *Measurement: Interdisciplinary Research and Perspectives*, 3 (1): 38-49

APPENDICES

Appendix A

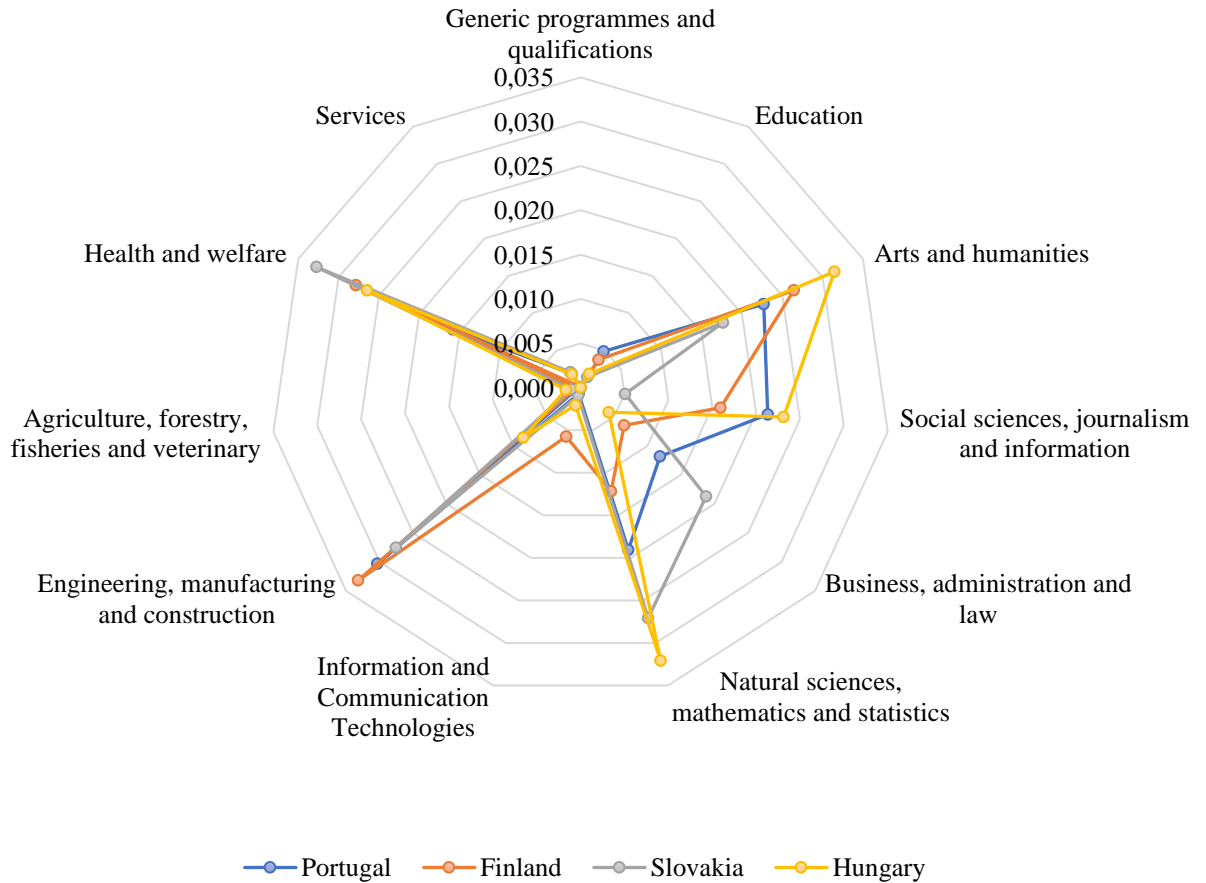
Number of researchers in European countries

Countries	Researchers	Population	% of researchers
Belgium	34 623	11 455 519	0,302
Czechia	26 766	10 649 800	0,251
Denmark	27 513	5 806 081	0,474
Germany	288 837	83 019 213	0,348
Estonia	4 461	1 324 820	0,337
Ireland	19 634	4 904 240	0,400
Greece	33 144	10 724 599	0,309
Spain	132 551	46 937 060	0,282
France	-	67 177 636	-
Italy	80 260	59 816 673	0,134
Latvia	5 564	1 919 968	0,290
Lithuania	11 495	2 794 184	0,411
Luxembourg	1 469	613 894	0,239
Hungary	24 612	9 772 756	0,252
Netherlands	27 766	17 282 163	0,161
Austria	41 021	8 858 775	0,463
Poland	110 004	37 972 812	0,290
Portugal	57 416	10 276 617	0,559
Slovenia	4 658	2 080 908	0,224
Slovakia	18 340	5 450 421	0,336
Finland	24 152	5 517 919	0,438
Sweden	36 183	10 230 185	0,354
Norway	28 821	5 328 212	0,541
Switzerland	50 583	8 544 527	0,592
United Kingdom	-	66 647 112	-

Source: Eurostat data (2019), author's calculations

Appendix B

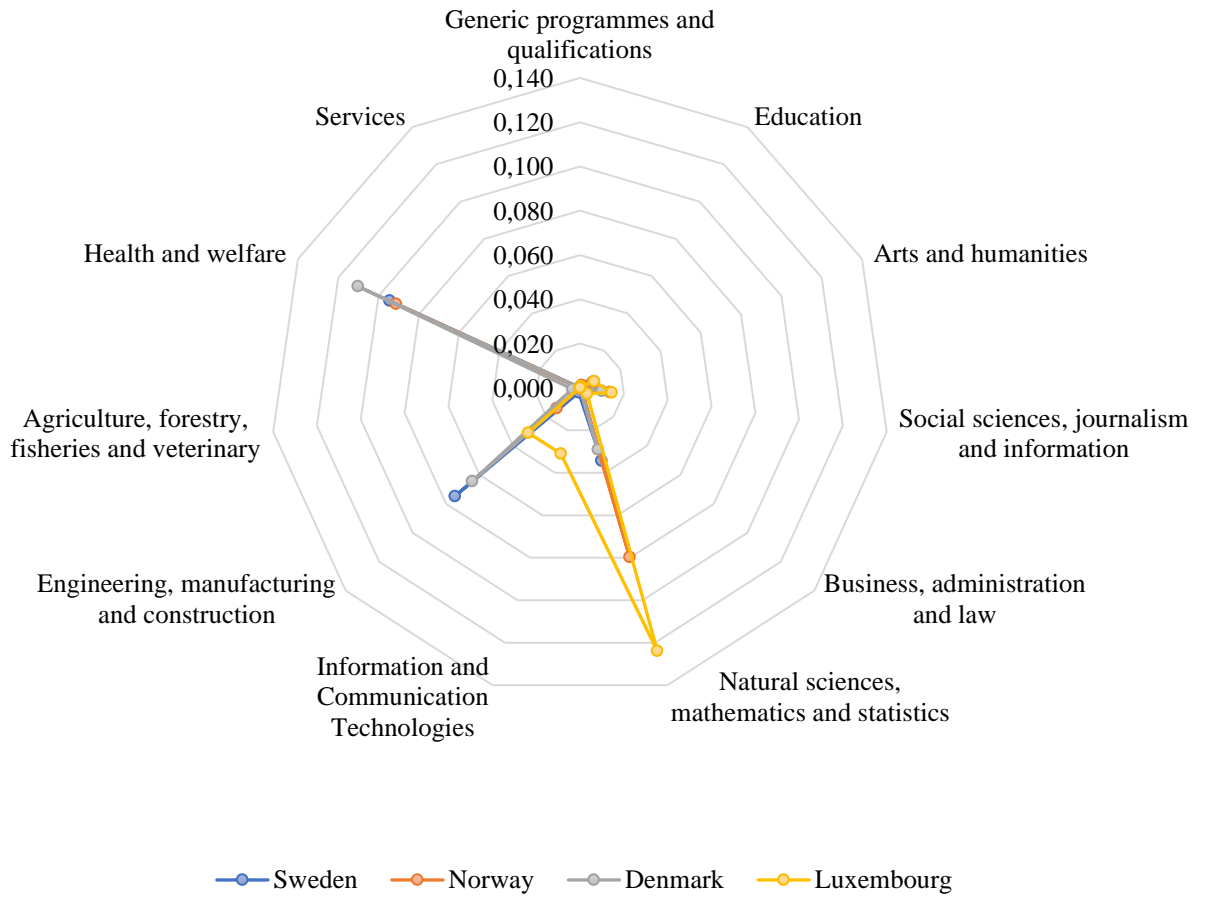
Results of the Simpson Index in each scientific field (based on doctoral students enrolled)



Source: Eurostat data, compiled by the author

Appendix C

Results of the Simpson Index in each scientific field (based on doctoral students graduated)



Source: Eurostat data, compiled by the author

Resümee

TEADUSE MITMEKESISUS ELIS JA ÜHENDKUNINGRIIGIS

Kristina Tsvetkova

Mitmekesisus on oluline kontseptsioon jätkusuutlikkuse jaoks oluliste ettevaatlike, vastupidavate ja kindlate omaduste väljatöötamiseks. Sel eesmärgil määratletakse uuringus lai raamistik mitmekesisuse mõistmiseks. See tähendab mitmekesisuse vaatamist kolme nõutava omaduse funktsioonina: mitmekesisus, tasakaal ja erinevus.

Samuti on tugev seos teaduse mitmekesisuse ja teaduse jätkusuutlikkuse vahel. Riigi haridus - ja teadussüsteem on stabiilsem, kui sellel on mitmekesisemad uurimisvaldkonnad (Lacey et al., 2019). Teisisõnu, mida mitmekesisem süsteem, seda jätkusuutlikum see on.

Lõputöö koosneb kahest peatükist: teoreetiline taust ja empiiriline analüüs. Alapeatükk 1.1 sisaldab teaduse mitmekesisuse mõiste üldist selgitust. Lisaks koostatakse tabel erinevate teadusvaldkondade mitmekesisuse erinevate määratluste kohta ja praeguse lõputöö raames määratletakse teaduse mitmekesisus. Alapeatükk 1.2 annab ülevaate ja võrdleva analüüsi erinevatest indeksitest mitmekesisuse mõõtmiseks ning mitme tavaliselt kasutatava indeksi eeliste ja puuduste kohta. 2. peatükk on empiiriline peatükk, kus teaduse mitmekesisust mõõdetakse mitme mitmekesisuse indeksi abil alapeatükis 2.1 ning tulemusi ELi ja Ühendkuningriigi kontekstis käsitletakse alapeatükis 2.2.

Lõputöös pakutakse välja kvantitatiivne lähenemisviis teaduse mitmekesisuse mõõtmiseks, rakendades Simpsoni, Shannoni ja Herfindahli mitmekesisuse indekseid. Lõputöö arvutab mitmekesisust teadusvaldkondades ELis ja Ühendkuningriigis, kasutades kindlaksmääratud indekse, mis põhinevad õpilaste arvul, kes õpivad ja lõpetavad doktoriõpet / samaväärset taset.

Uuringu empiirilised leiud näitavad, et Simsoni, Shannoni ja Herfindahli indeksid võivad olla tõhusad vahendid teadusuuringute mitmekesisuse mõõtmiseks Euroopa riikides ja Ühendkuningriigis. Siiski on oluline märkida, et kogutud andmete usaldusväärsus on mõnevõrra piiratud, kuna uuringus ei võetud arvesse erinevaid mitmekesisuse mõõtmist mõjutavaid tegureid. Autor tunnistab, et nende tegurite kohta kättesaadavate andmete puudumine kujutab endast olulist piirangut, ja olge ettevaatlikud, et kõiki uurimistööst tehtud järeldusi tuleb seda silmas pidades arvestada.

Tulevaste uuringute suunad on ajendatud lõputöö piirangutest. Esiteks puudub Euroopa tasandi andmekogum teadlaste arvu kohta valdkondade kaupa (See, mis sarnaneb doktorantide ja lõpetajate arvuga valdkondade kaupa), seega muudaks sellise andmekogumi väljatöötamine teaduse mitmekesisuse analüüsi terviklikumaks. Lisaks analüüsib väitekiri

doktorantide arvul (kes õpib ja lõpetas) põhinevat teaduse mitmekesisust. Kuid mõnesid muid olulisi aspekte nagu publikatsioonide arv ja tsitaadid uurimisvaldkondade kaupa, rahastamine valdkondade kaupa ei uurita selles väitekirjas. Kuna andmed ei ole rahvusvahelisel tasandil kättesaadavad, saab viimast aspekti uurida ainult riiklikul tasandil iga riigi kohta eraldi. Seega on tulemuste võrreldavus küsitav, kuna riikidel võivad olla erinevad rahastamisskeemid. Tuleks välja töötada ühtne lähenemisviis andmete kogumisele.

Non-exclusive licence to reproduce thesis and make thesis public

I, Kristina Tsvetkova,

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,

DIVERSITY OF SCIENCE IN THE EU AND UK

supervised by Mariia Chebotareva

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Kristina Tsvetkova

11.05.2023