PROMOTING SCIENTODIVERSITY THROUGH RESEARCH GRANTS

A Dissertation

Submitted to the National Graduate Institute for Policy Studies (GRIPS)

in Partial Fulfillment of the Requirements for the Degree of

Ph. D. in Public Policy

by

Yoshiaki Shimada

June, 2018

Abstract

The promoting diversity of science, or *scientodiversity* (variety, balance, and disparity in research subjects), is a prominent issue in science and technology policy because of their importance to research responsive to a wide range of socio-economic demands. However, resource allocation on science has been carried out on a publication performance-based without considering scientodiversity, given the lack of precise formulations and understandings such as those found in biodiversity studies. As a result, the decline in scientodiversity has emerged as a policy concern in Japan.

This problem on resource allocation is threefold; amount, distribution, and types. How much should we invest in science research as a country? What distribution of resource to research bodies is the optimum as a whole? What type of investment is the most appropriate to the promotion of scientodiversity? To answer these questions, this dissertation investigates the impact of resource allocation on the pattern and process of scientodiversity in three scales, such as country, university, and team, respectively.

First, I investigate the distribution of research subjects in the country-scale to develop a framework analogous to that of biodiversity. The result suggests that scientodiversity has similar statistical characteristics as biodiversity. Second, I evaluated the efficiency of universities in terms of the quantity and diversity of their publication. The results indicate the importance of the external research grant in university's research expenditure in terms of both publication and diversification. Third, I examined the impact of a mission-oriented grant on scientodiversity in the team-scale. The results show that the research subjects are better conserved under the mission-oriented program than the curiosity-driven one, a finding contrary to the conventional expectation.

These results may not only validate the adoption of sophisticated concepts and techniques from biodiversity studies in scientodiversity ones but also imply the possible "diversity-aware" design of science and technology policy.

Acknowledgements

I would like to express my gratitude to all people who have inspired and supported me during my study and research. This dissertation would not have been possible without those supports from many people. First, my deepest appreciation goes to my supervisors, Prof. Jun Suzuki, Prof. Tatsuo Oyama, and Dr. Naotoshi Tsukada for their dedicated advice and support. Without their guidance and persistent help, this dissertation would not have been possible. I stimulated and learned a lot from discussions with the supervisors. I have greatly benefited from the other committee members of my dissertation, Prof. Nagayasu Toyoda and Prof. Hitoshi Ieda, whose constructive comments helped me to improve the quality of my dissertation.

My sincere thanks also go to my colleagues at Japan Science and Technology Agency (JST), for their support and encouragement. More specifically, I would like to thank Tsutomu Kurosawa, Katsuji Matsumura, and Dr. Yasuhiro Yamashita, for their generous support of J-Global database. I would also like to express my gratitude to JST for the financial support. I would also like to express my gratitude to the anonymous referees in *Scientometrics* for their constructive comments that helped to improve my submitted papers. The results and conclusions in those accepted papers are the pillars of this dissertation.

Finally, I would like to express my deepest gratitude to my family whose encouragements have been the most powerful support.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Diversity in science

Diversity is a prominent concept widely used in wide variety of disciplines such as ecology, biology, physics, statistics, information science, sociology, economics, and policy studies (Grabher and Stark (1997); Hill (1973); Jost (2006); Limpert et al. (2001); May (1975); Newman (2005); Page (2010); Stirling (2007)). It has also been recognized as important in the science of science (Gibbons et al. (1994); Kuhn (1970); Merton (1973); Rosenberg (1996); Stirling (2007)). The impact of diversity in science on knowledge productivity has been studied in several aspects such as in terms of new combinations of existing knowledge (Rafols and Meyer (2010); Uzzi et al. (2013)), balance of expertise in interdisciplinary research teams (Aydinoglu et al. (2016); Barjak (2006); Lee et al. (2015)), efficiency in grant distribution (Hicks and Katz (2011); Shibayama (2011)), and geographical and gender balance (Williams and O'Reilly (1998)).

In particular, the promotion and maintenance of the diversity of research subjects, or *scientodiversity*, are prominent issues in science and technology policy because of their importance to research and innovation responsive to a wide range of societal demands (Gibbons (1999); Lund Declaration (2009)). Despite the known tradeoffs of diversity against transaction cost (Williamson (1993)), the economics of scale (Matthews and Mcgowan (1992)), and standardization (Cowan (1991)), scientodiversity has attracted great attention in policy discussions as a key driver of innovation.

However, only a few aspects have been investigated in terms of the impact of science policy on scientodiversity, and thus understanding of the key mechanisms of scientodiversity has been limited. In many studies, it has been shown what type of research theme is popular and/or growing (Börner (2010); Leydesdorff et al. (2013); Van Noorden (2015)), but merely mentioned how those distributions are maintained by means of science policy. This gap between observed scientodiversity and understanding of the key mechanism of it may hinder the evidence-based science policy. For example, statistical studies of the distribution of research topics, the universality of their distribution over many scales and time periods, and underlying mathematical models that generate and explain observed distributions may contribute to an understanding of the science of scientodiversity.

Despite the weak empirical evidence, a naïve conceptual analogy between scientodiversity and biodiversity raises primitive unanswered questions, notably the following. How are scientific research subjects distributed over disciplines in a given country, university, or research teams? Does that distribution differ substantially from country to country, and if so how? Are those distributions similar across countries? How did research portfolio of a university change with changes in policy and in time? What factors determine that distribution and how might that distribution change over time? What kind of policy tool is suitable for the effective and efficient maintenance of scientodiversity?

To answer these open questions, a promising way this thesis proposes is using an approach similar to that used in biodiversity studies. The ecosystem of sciences, like organisms' ecosystem, should be understood as an interactive system of science subjects and the environment. Local-scale patterns and dynamics of research subjects may be determined by local interaction among disciplines, like the inter-species interaction in bioecosystem, but are also interrelated with the external environment. The amount and structure of popular research topics in the local groups of different scales (*i.e.* countries, universities, research teams, etc.) are strongly influenced by their environment, in particular, by the resource allocation (research facilities, research expenses, human resources, etc.). Conversely, such environmental condition may be influenced by the composition of research subjects and the manner of interaction among them in given local groups through a discovery in local fields and/or the new combination of existing disciplines.

Therefore, an accurate grasp of the present situation and future prediction of scientodiversity are difficult without understanding the structure of the group of research topics and the interaction among constituent disciplines. In order to conduct evidence-based policies on science and technology, it is necessary to first understand the mechanism of science ecosystem as an interactive system including the environment, and then consider policy options for maintenance and/or promotion of scientodiversity based on these understandings.

1.2 Whys is scientodiversity a matter for concern?

In the philosophy and sociology of science, the scientodiversity has long been considered important to stimulating creative imagination (Kuhn (1970)) and improving rigor (Merton (1973)) of sciences, and to the creation of a portfolio of flexible strategic responses to uncertain futures (Rosenberg (1996)). From the viewpoint of innovation studies, promotion of learning (David and Rothwell (1996)), knowledge spillovers (Feldman and Kogler (2010)) and cross-fertilization (Braczyk et al. (2004); Levén et al. (2014)) are also important rationale for scientodiversity.

Recently, diversification of science has been regarded as a reflection of the complexity of social needs (Gibbons (1999)). Science in the modern society is expected not only to expand our intellectual frontier by discovery (Bush (1945)) but also to be responsive to a wide variety of socioeconomic demands, facilitating work in different academic domains and disciplines (Lund Declaration (2009); Schot and Steinmueller (2016)). Such a societal requirement of science has played a key role in opening the eyes of universities and public research institutes, as well as research funding agencies, to a vast unexplored territory of research subjects. In many countries, scientists have taken on such mission-oriented research, and funding agencies have consequently shifted their strategies towards socioeconomic interest and away from academic curiosity.

This new objective-driven research, which has been well investigated as mode 2 science (Gibbons et al. (1994)) or the Pasteur's Quadrant (Stokes (1997)), often lies outside traditional disciplinary boundaries of both basic or applied research. Complex and uncertain socioeconomic needs have surely shifted a certain amount of resource allocation across the spectrum of specific scientific subjects, towards mission-oriented research grant and specific research projects. The excessive expectation for practical application and solution to real problems may impair scientodiversity although scientodiversity is needed for solving those socioeconomic problems. Nevertheless, the impact of the mission-oriented grant has never been demonstrated empirically. The effect of resource allocation on the diversity of science is still an open question in the policy sphere, even though its understanding is crucial to decision-making regarding both concentration and diversification strategies. It is necessary to understand the mechanism of interaction between scientodiversity and science policy, and then to develop a new design of diversity-aware policy based on the understanding.

In the past, the importance of diversity as important as (1) the importance for the development of science itself, (2) the portfolio that flexibly responds to the demands of society has been debated. These are derived from the sociology of science, the economics of science, respectively. In this paper, we propose "view of science ecosystem" as the third reason that diversity is important. In scientific research, though the importance of ecosystems has been pointed out, there is no research on how scientific diversity is important to the ecosystem. According to ecological findings, the pattern of species distribution closely relates to ecological stability. It is known that the system stabilizes if there are a large number of species in special conditions where there is interaction like the problem raised for a long time. In recent years, it has also been pointed out that the diversity of interactions is important for the stability of the system. Even in scientific research, the relationship between scientific diversity and system stability is not expected to be simple. By understanding the role of science diversity in ecosystems, it is necessary to design policy considering the ecosystem.

1.3 Declining diversity in Japanese science

The decline in the diversity of science has emerged as a policy concern in Japan (National Institute of Science and Technology Policy (2015)). The results of a comprehensive questionnaire survey of Japanese researchers indicate that recent changes in the research environment, such as the growth of social expectations regarding rapid commercialization of research results, the creation of bureaucratic management systems to prevent research misconduct, and excessively performance-based evaluation of research, may serve to decrease the diversity of science, although there is as yet no concrete evidence of such a decrease. Quantitative analysis of clusters of highly cited articles has revealed a decrease in the coverage of clusters, *i.e.* of the percentage of research areas to which Japanese researchers contribute (from 41% in Science Map 2008 to 33% in Science Map 2012), (Igami and Saka (2016)).

Japan's stagnation of scientific activity has also shown in low publication performance as compared with other countries such as South Korea, China, and United Kingdom (Fuyuno (2017)). The total number of articles published by Japanese authors indeed slightly increased between 2005 and 2015 in Scopus database, but has not kept pace with the rapid growth of the world's average. Such decrease of Japan's presence in the research community is also evident in the decline of the ranking of Japanese research universities (Quacquarelli Symonds Limited (2017)Times Higher Education (2017)). A comprehensive survey on Japanese national universities in terms of both R&D budget and publication suggested that the total amount and distribution of budgets are associated to the observed decline of publication performance (Toyoda (2015)).

However, the relationship between the decrease in the number of papers and the decline of scientodiversity may not be simple. Indeed, the decrease ratio greatly varies from field to field observed 11 out of 14 research fields in Web of Science database (Fuyuno (2017)). Changes in Japan's total number of papers are dominantly influenced by changes in the number of papers in popular fields such as medicine, physics, chemistry, material science, engineering, biochemistry and molecular biology. In contrast, from a diversity perspective, relatively small changes in unpopular research areas often imply a crucial change in the whole ecosystem of scientific research activity. For example, mathematics is relatively small in terms of the number of publication but has a large impact on many disciplines. This concept, *i.e.* relatively rare elements having a wide influence on the entire system, has been well investigated as keystone species in ecology. Several diversity indices are known to be sensitive to the tail of a distribution, *i.e.* relatively rare elements, and widely used in ecology studies(Paine (1995); Power et al. (1996)). Moreover, the observed decrease in coverage of research areas in Science Map (Igami and Saka (2016)) itself should not be hastily taken as a decline in scientodiversity. Rather, it may well indicate a slow diversification in Japanese science, *i.e.* the percentage of research areas in which Japanese researchers participate cannot keep pace with the increase in the total number of research areas around the world.

Therefore, policies to improve average paper productivity do not always promote research diversity. Rather, those may impair scientodiversity, *i.e.* those will reduce the number of research topics and bring skew distribution. For example, it is expected that the competitive research grants stimulate researchers to produce more articles with better quality, but such grants sometimes motivate researchers to choose their research themes that researchers can easily obtain results. In particular, the mission-oriented grant is widely used and believed to limit researchers' freedom to choose their research themes, thus it should decrease scientodiversity. Promoting the mode 2 science (Gibbons et al. (1994)) or the Pasteur's Quadrant research (Stokes (1997)) is important in terms of response to the social demands and may also contribute to improving research productivity. However, excessive concentration of investment on a narrow area in the wide spectrum of possible research topics may prevent blooming of diverse emerging research.

1.4 Research object and challenges

The objective of this research is to understand fundamental mechanism of diversity and diversification of science and to apply it to design of new funding programs and comprehensive science, technology, and innovation (STI) policy.

The stagnation of government spending on science may have great impact on research performance in Japan as mentioned in previous studies (Fuyuno (2017); National Institute of Science and Technology Policy (2015); Toyoda (2015)). But the impact of such a stagnation of government expenditure on scientodiversity may not be simply estimated by the observed decline of the global share of the number of articles. The impediment to the survey on the impact of R&D funding to scientodiversity in country-scale is the lack of empirical and quantitative research because of the lack of standard framework for quantitatively evaluating scientodiversity. Thus, the establishment of such methodology will be the first challenge that this thesis has to tackle.

Another major challenge of this thesis is to show that the quantitative framework proposed by this thesis is effective in clarifying the relationship between scientodiversity and research fundings. This thesis will try to demonstrate the effectiveness of the framework on three different scales. First of all, I would like to clarify the relationship between R&D expenditure and scientodiversity in country-scale as a whole picture. Because the growth of Japanese R&D expenditure is about 20% in this decade, it is difficult to examine how scientodiversity will change when changing R&D funding on a logarithmic scale by analyzing Japanese data only. Therefore, an inter-country comparison is necessary. Previous reports on the decline of the number of research areas in comparison with other benchmarking countries also suggest the importance of country-scale comparison (Igami and Saka (2016)). The country-scale analysis of the relationship between scientodiversity and R&D expenditure will have some predictions and implications to help Japan's funding decision.

Second, I would like to grasp the relationship between scientific diversity and research funds in more detail than the national level. From the viewpoint of the flow of R & D funds, the proportion of government spending is only 15.4% of research and development funds nationwide, and more than 75% by private enterprises. Approximately half of this 15% of the funds will be invested in public research institutes and the other half in the university. On the other hand, more than 70% of all the papers published by Japanese institutions are published from universities with only 12% share as expenditure. Since these ratios differ from country to country, the result of the country-scale analysis will be only an understanding of the macroscopic behavior of scientodiversity. Therefore, in order to maintain consistency between input and output, it is necessary to analyze only for universities. The relationship between R&D funding and publication at Japanese national universities is well studied, and it has been reported that the decrease in research time and in the number of researchers in FTE count have a negative impact on the number of research articles (Toyoda (2015)). In addition, the concentration of grant to few universities has also been pointed out as a key cause of the decline of publication performance in terms of the effect of decreasing return to the scales (Shibayama (2011)). However, the effect of the latter on the paper productivity is reported to be smaller than the former, *i.e.* decrease in research time (Aoki and Kimura (2014)). Therefore, the analysis of scientodiversity at the university level is important in terms of the consistency of input-output, the richness of past research (*i.e.* it implies the easy availability of data), and also the high share of the research articles in Japan. Thus, universities should be investigated as the essential performing sector not only for publication but also for scientodiversity.

Third, I would like to empirically analyze the realistic case as a policy option. As already mentioned, it will be unrealistic for the Japanese government to increase the R&D investment on a logarithmic scale (for example double). If the block grant to the national university will not increase and competitive grants by the government and the grant from private companies will continue to increase, a possible policy option to promote and/or manage scientodiversity is devising a way of allocating competitive funds. Competitive funds are often allocated to researcher individuals or research teams. Therefore, the team-scale analysis is important to reveal the impact of the ways to allocate competitive grants (*i.e.* amount, distribution and type) on scientodiversity.

In this thesis, by overcoming the two key kinds of challenges (lack of framework and evidence) through three empirical studies, I will answer to the research question "how does the total amount, distribution and types of funding affect scientodiversity?".

1.5 Outline of thesis

The outline of this thesis is as followings (see also Figure 1.1). Chapter 2 draws a quantitative methodology to handle scientodiversity. First, we introduce (general) diversity indices developed and widely used in the field of ecology. Second, we compare several literature databases because the resolution of classification code is crucial for this research. We use J-Global database which has the high-resolution classification code scheme. Third, we introduce a lognormal distribution and a random multiplicative generative model to explain the local pattern of scientodiversity.

Chapter 3 shows that the statistical behavior of scientodiversity at the country scale can be understood by the analogy of ecology. In particular, we indicate that the characteristic of scientodiversity is determined by its scales. For example, it can be expected that richness of research subjects, which is one of the basic attributes of science ecosystem, depends on scales of R&D investment with assuming that the number of research topics is proportional to the number of articles (which is known to depend on R&D investment). Indeed, scale dependence called species-area relationship is widely known in the study of ecology (Hubbell (2001); May (1975)). By examining its analogy, we will indicate the importance of scale in scientodiversity. This is also a justification of the argument in this section.

Chapter 4 analyze the relationship between scientodiversity and budget at university scale. In particular, we investigate the efficiency of publication and diversification of Japanese national universities by data envelopment analysis when scientodiversity is regarded as one of the outputs of the research university. Research university plays a key role in scientific research in many countries and also well investigated from the viewpoint of public policy. However, the correspondence relationship between resource allocation among universities and the distribution of research subjects as a characteristic of each university is unobvious. The determining mechanism of university's scientodiversity may be different from the country's one even the size of the budget is in the same order. The

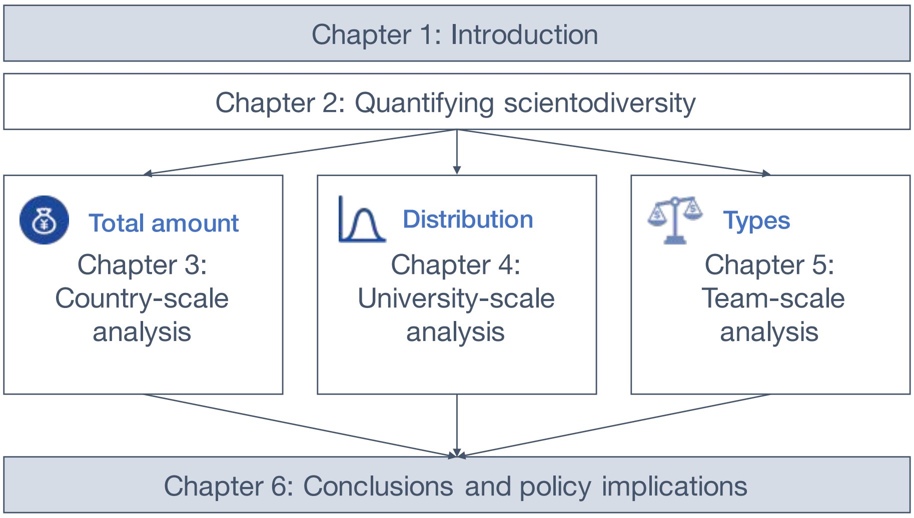


Figure 1.1: Schematic diagram of outline of the thesis.

chapter concludes the importance of external resources such as research grants from government and private companies.

Chapter 5 investigates the impact of the mission-oriented research grants on publication performance and scientodiversity at the team-scale. The relationship between interdisciplinarity in a research team and publication performance has been studied and the result justifies the importance of interdisciplinary research. However, what the policy-makers and funding agencies face in terms management and promotion of scientific research is a problem in much large-scale compared with the scales of these empirical studies. For example, even the importance is evident, policy design to incentivize researchers to build such an interdisciplinary team still needs concrete evidence. For this reason, how to predict macroscopic results from findings obtained on micro-scales, or vice versa, how to predict micro-scale changes from macroscopic environmental changes is required. This recognition is an underlying concept of this chapter.

Chapter 6 summarizes the results obtained in chapters 3 to 5 and policy implications and adds a comprehensive discussion. In particular, program design and recommendation to improve richness and balance will be explored with a focus on comprehensive analysis over three different scales.

Chapter 2

Quantifying scientodiversity

2.1 Diversity indices

The methodologies for formalization and quantification of diversity are well established in the fields of ecology and economics (Stirling (2007)). Several types of diversity index have been proposed for the measurement of three heuristic features of diversity (variety, balance, and disparity), and have been recently applied to scientometric studies (Mitesser et al. (2008); Pan et al. (2012); Rafols and Meyer (2010); Wagner (2010)). Many investigations of the structure of science have provided visualizations of the diversity and mutual relationship among elements of scientific knowledge as maps of science (Börner (2010); Leydesdorff et al. (2013)).

In this study, we use the total number of subject, the richness R, as a measure of variety because this is the simplest and the most popular way to characterize a biological community in a single parameter (May (1975)). Another popular single number for characterization is the total number of individuals, N. In our case, this is equal to the total number of papers in a group (*i.e.* a specified year and country). Notice that the research subjects and biological species are quite different in their definitions. While the biological species can be defined by objective methodologies (to some extent) such as morphological species, genetic species, and ecological species, the classification of the research theme in a bibliographic database is always more subjective (see also the limitation section in Chapter 6). One may ask how the individuals (*i.e.* papers) are distributed among species (research subjects). Usual models in ecology are defined as a class of functions for $S(n)$, the number of species with population n. The total number of species, or the richness R, and the total number of individuals N are thus related to $S(n)$ as

$$R = \int_0^\infty S(n)dn \tag{2.1}$$

$$N = \int_0^\infty nS(n)dn \tag{2.2}$$

We use Gini-Simpson index

$$1 - \lambda = 1 - \sum_{i=1}^{R} p_i^2 \tag{2.3}$$

for a measure of evenness.

Gini-Simpson index corresponds to the probability that two individual organisms you happen to see are of the same species when you are walking through a field. As same as Gini-Simpson index, Shannon-Wiener index $H' = -\sum_{i=1}^{R} p_i \ln p_i$ also evaluate rarity of an event that two of your sample are of same species. However, Shannon-Wiener index puts more weight on rare species than on common ones, while Gini-Simpson index puts same importance on them. This idea is derived from Shannon entropy in information theory in which the rarer news is considered to contain the richer information. Because of this formulation, Shannon-Wiener index is more sensitive to rare species while Gini-Simpson index is dominantly influenced by common species. Thus, Shannon-Wiener index is less robust than Gini-Simpson one against numerical fluctuation of counting for rare species. For example, in the beginning, you have $N = 1000$ samples and estimate that they can be divided into three species with relative abundance $p_1 = 998/1000$ and $p_2 = p_3 = 1/1000$. Then, you can calculate Gini-Simpson index $1 - \lambda = 0.996$ and Shannon-Wiener index $H' = 0.0158$. However, if you accidentally notice an error in your classification and finally find that the individual in the species 2 should be counted as the species 1, *i.e.* $p_1 = 999/1000$ and $p_3 = 1/1000$, both indices must be corrected as $1 - \lambda = 0.998$ and $H^{'} = 0.00791$, respectively. For the larger number of sample, this rare-species-sensitive behavior of Shannon-Wiener index makes the noisier result. In this study, we choose Gini-Simpson index for a measure of evenness since it is robust on a counting error.

Gini-Simpson index, by definition, depends on the value of R. It takes the maximum value $1 - 1/R$ when all classification codes have the equal population ($p_i = 1/R$ for any category i). In order to measure balance independently from variety (or richness), we also use an index called evenness defined as

$$E = 1/R \sum_{i=1}^{R} p_i^2 \tag{2.4}$$

corresponding to the Gini-Simpson index normalized by $1/R$.

In this study, we assume that all classification codes in J-Global database (see the next section) are equally different from each other, *i.e.* a disparity between subject i and j is same as the disparity between subject i and k for any i, j, and k ($i \neq j \neq k$). This setting

Table 2.1: Stirling's general diversity index.

	$\alpha = 0$	$\alpha = 1$
$\beta = 0$	Variety $\Delta_{00} = \sum_{ij(i \neq j)} (d_{ij})^0 = \frac{R(R-1)}{2}$	Disparity-weighted variety $\Delta_{10} = \sum_{ij(i \neq j)} d_{ij}$
$\beta = 1$	Balance-weighted variety $\Delta_{01} = \sum_{ij(i \neq j)} p_i p_j = \frac{\lambda}{2}$	Balance/disparity-weighted variety $\Delta_{11} = \sum_{ij(i \neq j)} d_{ij} p_i p_j$

is equivalent to Stirling's general diversity index (Stirling (2007)) in the case of $\alpha = 0$. The relation between Stirling's index $\Delta_{\alpha\beta}$ and richness R and Gini-Simpson index $1 - \lambda$ are listed in Table 2.1. Disparity d_{ij} between subject i and j can be defined by similarity in terms of co-occurrence pattern. Possible disparity-weighted calibration of diversity indices used in this study is discussed in Chapter 6.

2.2 Bibliometric databases and classification codes

The diversity of research topics is often measured in terms of diversity of classification codes attached to journals and/or papers, prompted by the availability of bibliographic datasets. The ISI subject category is the most popular for systematical classification of research subjects (Rafols and Meyer (2010)). These category codes are assigned to each journal but not to papers, so this coarse classification cannot be sufficient for analysis of the detailed structure of research topics. Descriptors of each paper, such as title, co-author, and reference, are often incorporated to improve the resolution of the classification. Several fine classification schemes are available, but only for specific fields and specific research subjects. For example, the JEL code and the PACS code are often used for classification in the field of economics and physics, respectively.

Clustering of research papers are also commonly used to quantify the diversity of research topics by evaluating the variety of forward citation (*i.e.* the number of other papers citing the paper) and/or reverse citation (number of distinct references in the paper) (Carley and Porter (2012); Van Noorden (2015)). This citation-based approach affords a powerful description of relations among scientific subjects and a bird's-eye view of whole network structures (Leydesdorff and Rafols (2011); Leydesdorff et al. (2013); Trajtenberg et al. (1997)). The clustering of papers by means of citation networks can generate a sufficiently granular classification of research topics without an a priori classification system (Mitesser et al. (2008); Schmidt et al. (2006)), although identification and consistency of clusters over time remain challenging research targets in bibliometrics (Igami and Saka

(2016)). Clustering has also made it possible to compute a number of diversity indices, such as Gini coefficient, Shannon entropy, Rao-Stirling index, as well as network parameters such as degree, betweenness centrality, and cluster coefficient (Leydesdorff and Rafols (2011); Wagner (2010)). Co-authoring network is one of the most useful foundation tools for measuring diversity within a research team or research community (Abbasi et al. (2011); Lee et al. (2015); Voutilainen and Kangasniemi (2015)).

Those clustering techniques are popular and mature for generation of the fine structure of research topics without an *a priori* classification scheme (Mitesser et al. (2008); Schmidt et al. (2006)). However, the consistency of clusters over time can hardly be guaranteed because of the arbitrariness underlying the identification of each cluster. Thus, detailed analysis of the diversity of subjects, which depends on the fine distribution of papers among clusters, needs to pay close attention to the definition of clustering parameters. Text-mining is one promising alternative technique for extracting detailed information for each paper and the structure of each research subject, although it requires considerable computational resources (Kostoff (2012)). Thus, the measure of diversity of research subjects with large coverage, fine granularity, and sure consistency is necessary for detailed analysis of the diversity of science.

In this study, we used the J-Global database in order to mitigate the shortage of resolution of the classification scheme. The characteristics of each publication database are shown in Table 2.4. The collection of the J-Global database is skewed to Japanese journals, but major international journals for the area we survey here are covered. The number of journal collected in Scopus, Web of Science, and J-Global are summarized in Figure 2.1. The classification scheme in Scopus, which is often used for bibliometric studies, is too coarse to analyze the diversity of research subjects at the level we need since the category code is assigned on a journal-by-journal basis. For example, articles on superconductivity and supernovae in Physical Review Letters cannot be distinguished in terms of Scopus classification because all articles in that journal are assigned to the one category 'physics and astronomy'. The resolution of the category code system is also insufficient in the ISI Web of Science database. The number of categories at the finest level in Scopus is 313 over all disciplines, which is slightly larger than that for the Web of Science database, 175. In contrast to those two major bibliographic databases, the J-Global database contains JST classification code scheme that hierarchically distinguish research subjects in science and technology field by 3,367 categories attached to each articles (Kitai (2008)), although the collection of journal in the J-Global database is slightly different from the Scopus and the Web of Science.

The granularity of the classification code influences the diversity index. Table 2.2 shows samples of distribution of the number of articles with each classification codes. There are two classification codes, namely "fine" and "coarse", and the structure of subdivision is different depending on the classification code, *e.g.* category d includes only 2 subcategories d1 and d2 while category e have 6 subcategories. As summarized in Table 2.3, we compute the diversity indices by using "fine" or "coarse" classification scheme for these 5 cases.

Case A is a flat distribution when viewed in the fine classification, but distribution with a somewhat biased distribution in category e is seen in the coarse classification. Although it is perfectly balanced distribution when viewed in the fine classification, the Gini-Simpson coefficient does not become 1 as described in the previous section. However, Evenness is equal to 1 because it is normalized by the value of R. Evenness computed by the coarse classification is 0.909, which indicates slightly skewed distribution. Conversely, Case B has a flat distribution ($E = 1$) in the coarse classification, but it looks skewed in the fine classification ($E = 0.882$). Case C is a flat distribution in the fine classification as same as Case A, but Gini-Simpson index for Case C is smaller than that for Case A because of its small R as discussed in the previous section. Case D and Case E have the same distribution shape in the fine classification and the diversity indices in the fine classification are exactly the same in those two cases. However, in the coarse classification, the distribution is not necessarily the same, reflecting the structure of the classification code, and then the distribution of Case E looks more skewed than that of Case D. That is, there are cases in which the skewness of distribution to be noticed is hidden by subdividing the classification code, but conversely, there may be a skewness that cannot be seen unless classification codes are enough granular. In parallel with the discussion of what to be the fair balance, it will be necessary to consider what to be the appropriate granularity of classification.

In the classification system, there are five hierarchical ranks expressed with a classification code $A_1A_2N_1N_2N_3N_4N_5A_3$ where A_i and N_i are alphabet character and number, respectively. Five ranks are described by A_1, A_1A_2, $A_1A_2N_1N_2$, $A_1A_2N_1N_2N_3N_4$, and $A_1A_2N_1N_2N_3N_4N_5$ in the order of fineness, and A_3 is a check digit. Number and examples of the classification code in each rank are listed in Table 2.5. For example, in the JST classification scheme, the finest-level code BM03043X represents Electric conduction in crystalline semiconductors. This code reflects the structure of code system as the followings; the first character B represents Physics and BM represents Electronic structure, electrical, magnetic and optical properties as a subcategory of B: Physics, the following two digits 03 represents Electrical properties: electronic conduction as a subcategory of BM, and the following 04 represents Electric conduction in semiconductors and insulators.

This classification scheme on scientific paper resembles to Linnaean taxonomy in terms of its hierarchical rank-based structure which opposed to cladistics approach. The classification code was developed based on the Universal Decimal Classification through

Table 2.2: Sample distributions for the comparison of granularity.

classification code		Case A		Case B		Case C		Case D		Case E	
a	a1	2,000	500	2,000	500	4,000	1,000	3,600	975	400	25
	a2		500		500		1,000		925		75
	a3		500		500		1,000		875		125
	a4		500		500		1,000		825		175
b	b1	2,000	500	2,000	500	4,000	1,000	2,800	775	1,200	225
	b2		500		500		1,000		725		275
	b3		500		500		1,000		675		325
	b4		500		500		1,000		625		375
c	c1	2,000	500	2,000	500	2,000	1,000	2,000	575	2,000	425
	c2		500		500		1,000		525		475
	c3		500		500		0		475		525
	c4		500		500		0		425		575
d	d1	1,000	500	2,000	1,000	0	0	700	375	1,300	625
	d2		500		1,000		0		325		675
e	e1	3,000	500	2,000	333	0	0	900	275	5,100	725
	e2		500		333		0		225		775
	e3		500		333		0		175		825
	e4		500		333		0		125		875
	e5		500		334		0		75		925
	e6		500		334		0		25		975
Total		10,000			10,000		10,000		10,000		10,000

Table 2.3: Comparison of diversity indices between the fine and coarse classification.

	Case A	Case B	Case C	Case D	Case E
Fine					
Richness	20	20	10	20	20
Gini-Simpson	0.950	0.943	0.9	0.933	0.933
Evenness	1	0.882	1	0.750	0.750
Coarse					
Richness	5	5	3	5	5
Gini-Simpson	0.780	0.800	0.640	0.739	0.667
Evenness	0.909	1	0.926	0.766	0.601

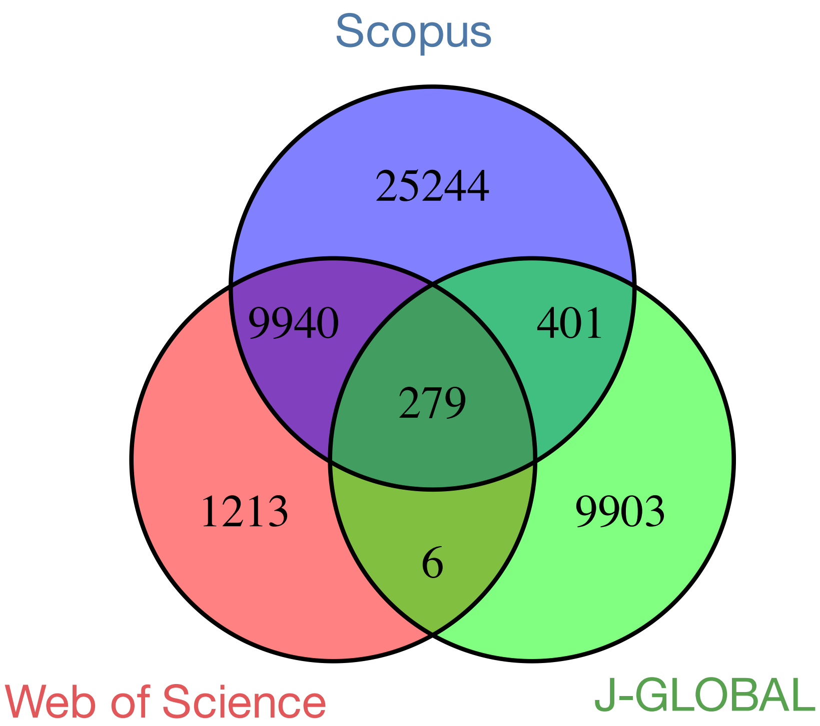


Figure 2.1: Comparison of bibliographic databases.
The numbers included in the Venn diagram represent the number of journals recorded in each database (as of December 2016). For example, 401 journals are registered in both J-Global and Scopus, while 9903 journals (including many Japanese journals) are recorded only in J-Global. Identification and matching of the journal were done by ISSN. The list of journals contained in the Web of Science and Scopus were downloaded from each web site.

Table 2.4: Comparison of bibliographic databases.

	Web of Science	Scopus	J-GLOBAL
Collection (international journal)	Approx. 12,000	19,829	5,011
Collection (Japanese journal)	373	417	9,514
Number of records	Approx. 55M	Approx. 30M	Approx. 35M
Number of category (in the coarsest level)	22	27	24
Number of category (in the finest level)	175	313	3,367
Unit of classification	Journal	Journal	Paper

As of April 2015

expert knowledge in 1975 and revised in 1981 with reference to the Broad System of Ordering developed by UNESCO (Sakagami (1989)). In this study, we use the latest version of the classification system, revised in 1993. Up to three classification codes can be attached to one article in J-GLOBAL database. The average number of attached classification code per paper is 1.31. Due to this property that multiple classification codes can be attached to a single article, it is possible to classify the paper of a new interdisciplinary research topic without any drastic revises of the classification code system. More than 90% of classification codes are attached to at least 1000 papers in J-Global database as shown in Figure 2.2.

2.3 Datasets

In this research, we used data from several sources. The number of papers (including original article, review, proceedings) is computed from the papers in the journals recorded in both J-Global and Scopus commonly. The raw data (tsv format) of J-Global database and the conversion table of the document ID between the databases was provided by the department of information planning, Japan Science and Technology Agency. Following with the lists of the article retrieved from Scopus on the web, SciVal calculates the number of highly-cited articles. J-Global gives the count of classification codes for a certain group of articles, and we can easily compute diversity indices from it. Scopus also gives the distribution of subject in its classification scheme (*i.e.* it is much coarser than that of J-Global) for a certain group of articles.

The expenditure on research and development (ERD) used in Chapter 3 has been retrieved and calculated from World Development Indicators (The World Bank (2017)).

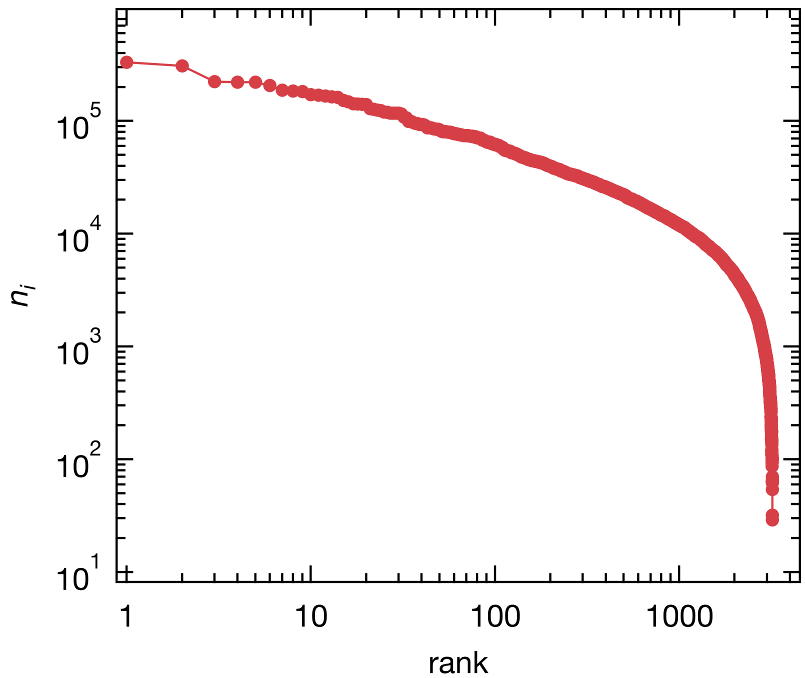


Figure 2.2: Abundance-rank plot of the classification codes in J-Global. The plot shows the number of articles for each 3367 classification codes in descending order of the number of articles. The number of articles with the most popular classification code in this database is about 3×10^5, while there are some rare classification codes with less than 1000 articles. The proportion of such rare classification codes is less than 10%.

Table 2.5: Example of JST classification code.

Rank	**Number**	**Example**	**Index**
1	24	B	Physics
		E	Biology
2	155	BM	Electronic structure, electrical, magnetic and optical properties
		EG	Microbiology, virology
3	533	BM04	Superconductivity
		EG04	Virology
4	133	BM0404	Superconducting materials and applications
		EG0404	Virus physiology
5	3,367	BM04042N	Superconducting magnets
		EG04042Y	Physiology and pathogenicity of virus infection

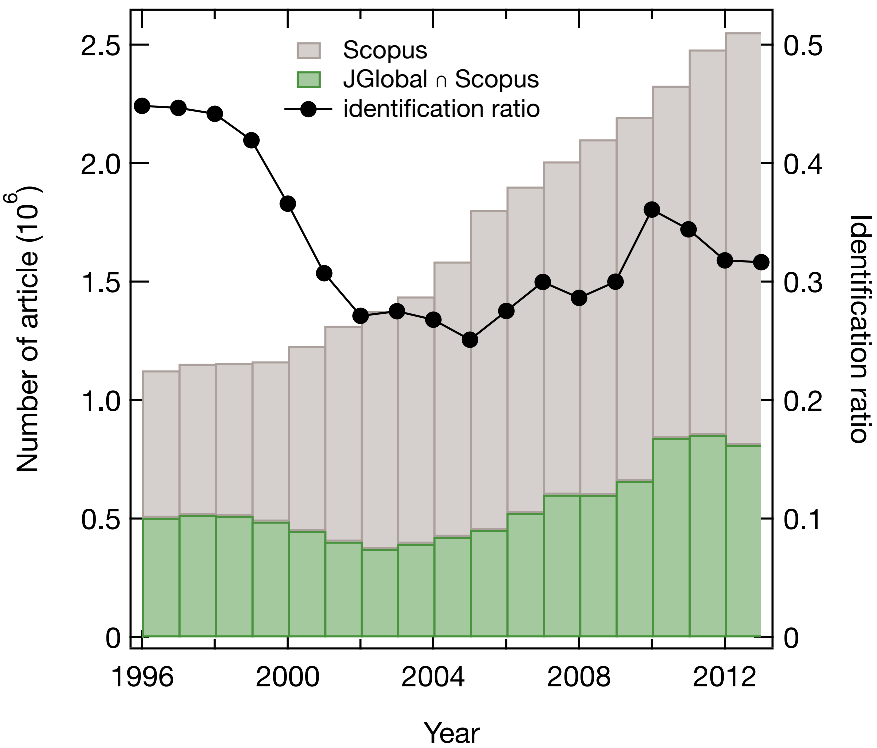


Figure 2.3: Comparison between J-Global and Scopus.
The number of articles recorded in Scopus (including proceedings and reviews) is shown in the gray bar. Among those articles, the number of articles recorded also in J-Global is plotted by the green bar. This identification of articles was done by the journal name, the title of articles, and the author name. The identification rate indicated by black circles (read on right axis) is approximately 30% in this time window.

The information of the expenditure and human resources of Japanese national universities used in Chapter 4 are based on the microdata of "the Survey of Research and Development, Ministry of Internal Affairs and Communications". All aggregation and statistical calculation have been done by the authors, independently from the ministry. The raw data became available only after our application for "Secondary use of official statistics" was accepted by the Statistics Bureau. These input data matched the bibliometric data by the name of the university. The information of the principal investigators such as affiliation, position, the location of their affiliated institution used in Chapter 5 was retrieved from the public databases, such as the JST Project Database and the Database of Grants-in-Aid for Scientific Research, utilized by the funding agencies. Those researchers were also identified in the Scopus database by their name and affiliation.

2.4 Log-normal distribution

The log-normal distribution is most often used to describe distribution seen in ecology (Limpert et al. (2001)), and has been observed in this scientodiversity studies as shown in Chapter 3 and Chapter 5. In general, the lognormal distribution requires three parameters S_0, μ, and σ for its specification:

$$S(n) = S_0 \exp\left[-\frac{(\ln n - \mu)^2}{2\sigma^2}\right]. \tag{2.5}$$

Instead of these three parameters S_0, μ, and σ, ecologists prefer to use the richness R, $n_{\max}$, which is the number of individuals in the most abundant species, and a parameter $\gamma \equiv \log_2 n_{\mathrm{mode}} / \log_2 n_{\max}$, where n_{mode} represents the mode of $nS(n)$, which is also a Gaussian function of $\ln n$. This new set of parameters $\{R, n_{\max}, \gamma\}$ is sufficient to determine a unique lognormal distribution since both R and γ depend S_0 and σ, while $n_{\max}$ can be written as a function of S_0, μ, and σ (see Appendix A). The Gini-Simpson index can be computed from given parameters S_0 and σ, and thus it cannot be a new parameter to be added to the parameter set.

The parameter γ characterizes the shape of the lognormal distribution. A distribution with smaller γ is relatively sharper than one with larger γ. The γ is also corresponding to the positional relationship between $S(n)$ and $nS(n)$. One of the reasons why γ is so popular in ecology studies is Preston's "canonical hypothesis" which proposes $\gamma = 1$ (Preston (1962, 1980)). This empirical assumption fits many data in biodiversity study and is still an open question whether it derives from purely mathematical reason or biological and/or ecological one. One strong support for this phenomenological hypothesis is the prediction of species-area relations in power law by assuming canonical lognormal species-abundance distribution (Irie and Tokita (2012); May (1975)). Regardless of its theoretical origin, if the lognormal distribution is "canonical", *i.e.* $\gamma = 1$, the lognormal distribution can be uniquely determined by only two parameters, R and $n_{\max}$ as a purely mathematical consequence.

In Chapter 3, we estimate γ from histograms of $S(n)$ and $nS(n)$ (see Figure 3.5). We count the number of papers n_i for each subject i, and then create histogram of the subjects with bin width of 0.5 on a scale of logarithms to the base 2 (Preston's octave), *i.e.* count the number of subjects each of which contains between $n/\sqrt[4]{2}$ and $n \times \sqrt[4]{2}$. We compute the histogram of $nS(n)$ from $S(n)$ for each bin, and then estimate γ by observed $n_{\max}$ and n_{mode}. We also count the total number of subjects, *i.e.* the richness R, and the Gini-Simpson index is calculated from n_i.

The generative model for lognormal distribution is random multiplicative process (Render (1990)). Consider a binary sequence in which the positive numbers z_1 and z_2

($z_1 > z_2$) independently appear with probabilities p and $q = 1 - p$, respectively. When there are S elements in this sequence, the probability function $p(k)$ can be written as the Gaussian form,

$$p(k) = \frac{1}{\sqrt{2\pi\sigma_{\mathrm{norm}}^2}} \exp\left(-\frac{(k-\mu_{\mathrm{norm}})^2}{2\sigma_{\mathrm{norm}}^2}\right) \tag{2.6}$$

where an average $\mu_{\mathrm{norm}} = Sp$ and variance $\sigma_{\mathrm{norm}}^2 = Spq$, as the continuum limit approximation of a binary product

$$p(n) = \left(\begin{array}{c} S \\ \\ k \end{array}\right) p^k q^{S-k} \tag{2.7}$$

By using the generic value of the product $X = z_1^k z_2^{S-k} = (z_1/z_2)^k z_2^S$, the number of z_1 in the sequence can be rewritten as

$$k = \frac{\log \frac{X}{z_2^N}}{\log \frac{z_1}{z_2}} = \log\left(\frac{X}{b}\right)^{\frac{1}{a}} \equiv \log m \tag{2.8}$$

where $a = \log z/z_2$ and $b = z_2^S$. Then, by transformation of the variable k to X, one obtains the log-normal distribution function (*i.e.* normal distribution along $\log X$)

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} p(n)dn \\ &= \int_0^{\infty} \frac{1}{m} p(\log m) \frac{dm}{dX} dX \\ &= \int_0^{\infty} \frac{1}{\sqrt{2\pi\sigma_{\mathrm{norm}}^2 a^2}} \frac{1}{X} \exp\left(-\frac{(\log X - \mu_{\mathrm{norm}} a - \log b)^2}{2\sigma_{\mathrm{norm}}^2 a^2}\right) dX \end{aligned} \tag{2.9}$$

The average $\bar{X}$ and median $\tilde{X}$ of the log-normal distribution function $p(X)$ with parameters $\mu = \mu_{\mathrm{norm}} a + \log b$ and $\sigma = \sigma_{\mathrm{norm}} a$ are computed as

$$\begin{aligned}
\bar{X} &= \exp\left(\mu + \frac{\sigma^2}{2}\right) \\
&= \exp\left(Np\log\left(\frac{z_1}{z_2}\right) + N\log z_2 + \frac{1}{2}Np(1-p)\left(\log\left(\frac{z_1}{z_2}\right)\right)^2\right), \qquad (2.10) \\
\tilde{X} &= \exp(\mu) \\
&= \exp\left(Np\log\left(\frac{z_1}{z_2}\right) + N\log z_2\right). \qquad (2.11)
\end{aligned}$$

The average $\bar{X}$,median $\tilde{X}$, a lognormal parameter σ are plotted as a function of design parameters $\{z_1, z_2, p\}$ in Figure 2.4 and 2.5.

The investment portfolio is important from the perspective of scientodiversity and science policy. For example, in the case of KAKENHI (Grant-in-Aid for Scientific Research), the relevant funding agency JSPS has adjusted the adoption rate so that there is not much difference between scientific fields, and does not define priority areas. Indeed, the adoption rate seems to be constant among all disciplines. This may be fair for all disciplines. However, the constant rate makes the adopted *number* non-constant (and thus the total amount of allocated grant) over research subjects. The number of adopted proposals simply be proportional to the number of submitted proposals which may be associated with the population of the specific research community. This setting can be understood as the Matthew effect, *i.e.* the rich discipline gets richer, and the poor get poorer. By repeating this process, the distribution of the number of papers among disciplines will be skewed.

Assume that the resource allocation $B_i(t)$ to the subject i at time t is proportional to the number of applications A_i to the call with the adoption rate p', which is independent of subject i, the number of papers produced in that field can be written in a simple linear model;

$$x_i(t+1) = \beta_i B_i(t) = \beta_i p' A_i(t), \qquad (2.12)$$

where β_i is a coefficient, which represents (effective) productivity. The number of application $A_i(t)$ may be proportional to the size of research community, which is estimated by the number of publication $x_i(t)$ with a coefficient α_i, then the microscopic process of the Matthew effect discussed above is described as a following simple time-evolution model;

$$x_i(t+1) = \beta_i p' \alpha_i x_i(t) \qquad (2.13)$$

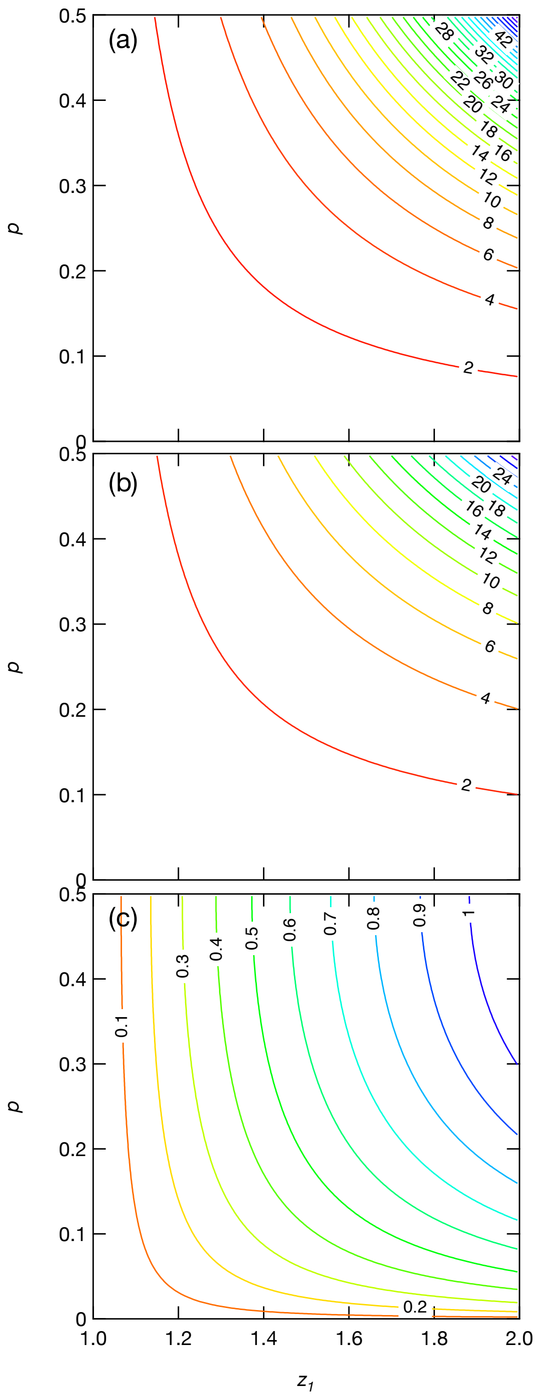


Figure 2.4: Contour plot of (a) average $\bar{X}$, (b) median $\tilde{X}$, and (c) a lognormal parameter σ for constant z_2.

Under this condition, σ hardly depends on p when z_1 is small ($z_1 \sim 1.1$), but suddenly becomes sensitive to p when z_1 increases to $z_1 \sim 2$.

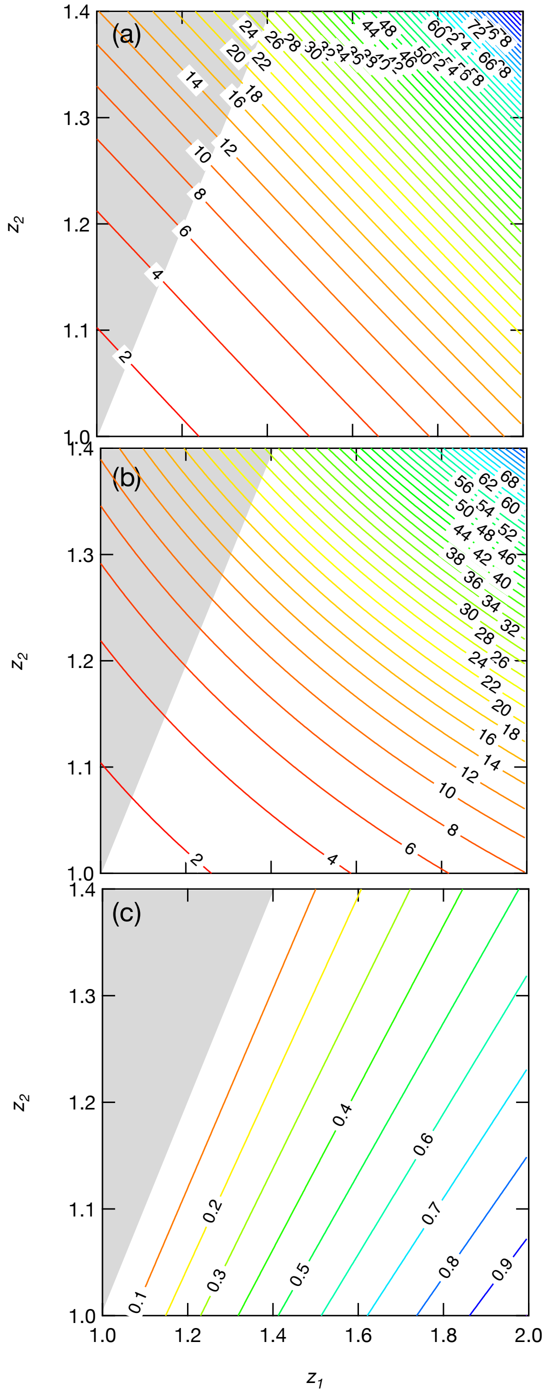


Figure 2.5: Contour plot of (a) average $\bar{X}$, (b) median $\tilde{X}$, and (c) a lognormal parameter σ for constant p.
Under this condition, the ratio of z_1 and z_2 is a parameter characterizing $\bar{X}$, $\tilde{X}$, σ, respectively, *i.e.* the contour lines are almost linear in the z_1-z_2 plane. The parameter area ($z_1 < z_2$) which is not permitted by definition is represented by a gray hatched area.

This model is known as a general random multiplicative process, which generate lognormal distribution (Limpert et al. (2001); Newman (2005); Render (1990)).

2.5 Production function

The shape of the distribution function is closely related to the shape of the production function. Given the distribution shapes of input X and output Y, the functional shape of the production function $Y = F(X)$ is identified as the followings. First, let the cumulative distribution function of X and Y be represented with at most two parameters as follows:

$$P(X \geq x) = \Phi_X(x; \mu_X, \sigma_X), \tag{2.14}$$
$$P(Y \geq y) = \Phi_Y(y; \mu_Y, \sigma_Y). \tag{2.15}$$

Then, assuming that the production function $F(X)$ monotonically increases with increase of X, the inverse function of production function $F^{-1}(y)$ is computed by use of the relationship between those two distribution functions as

$$\Phi_Y(y; \mu_Y, \sigma_Y) = \Phi_X(F^{-1}(y); \mu_X, \sigma_X). \tag{2.16}$$

When both $\Phi_X(x; \mu_X, \sigma_X)$ and $\Phi_Y(y; \mu_Y, \sigma_Y)$ are lognormal distribution, the cumulative distribution function can be explicitly wrote down as

$$\Phi_X(x; \mu_X, \sigma_X) = \frac{1}{2}\mathrm{erfc}\left(\frac{\ln x - \mu_X}{\sigma_X\sqrt{2}}\right), \tag{2.17}$$
$$\Phi_Y(y; \mu_Y, \sigma_Y) = \frac{1}{2}\mathrm{erfc}\left(\frac{\ln y - \mu_Y}{\sigma_Y\sqrt{2}}\right). \tag{2.18}$$

Then, the production function $F(X)$ can be formalized as the Cobb–Douglas production function:

$$Y = F(X) = \exp\left(\frac{\mu_X\sigma_Y - \mu_Y\sigma_X}{\sigma_X}\right) X^{\frac{\sigma_Y}{\sigma_X}} \tag{2.19}$$

Concrete derivation and cases where X and Y have other distribution shapes are summarized in Appendix B.

Table 2.6: Classification of Japanese national universities by their size.

Class	Budget size (10^{10}JPY)	Medical	STEM	Social Science	Education
1) Research	~10	O	O	O	O
2) Large	~5	O	O	O	O
3) Middle	~3	O	O	O	O
4) Tech	~1		O		
5) Non-med	~1		O	O	O
6) Social	~1			O	O

2.6 Data envelopment analysis

The data envelopment analysis (DEA) methodology was developed for empirical estimation of production frontiers in operations research and economics (Charnes et al. (1978)) and has been used to measure the productive efficiency of decision-making units. DEA is often used to assess the efficiency of not an only private company but also public and not-for-profit organizations such as hospitals (Kuntz et al. (2007)), police (Aristovnik et al. (2014); Thanassoulis (1995)), and universities (Ahn et al. (1988); Johnes (2006); Wolszczak-Derlacz and Parteka (2011)). Since DEA does not assume any particular production function, *i.e.* non-parametric, the most efficient decision-making units defined by DEA form "best-practice frontier" (Cook et al. (2014)) effective only within the data and may not necessarily form a general production frontier.

In this study, we select 69 national universities in Japan as listed in Table 2.6 and 2.7 where the universities are categorized in 6 classes according to the previous study (Toyoda (2015)). The universities in the category of 1) Research, 2) Large, and 3) Middle have the whole set of departments of science, technology, engineering and mathematics (STEM), medical, social science and education. Those three categories are classified by means of the size of the budget. The universities in 4) Tech, 5) Non-med, and 6) Social have similar size of the total budget ($\sim 1 \times 10^{10}$ JPY/year) but their configuration of the department is different. The universities in 5) Non-med have the department of the STEM (Science, Technology, Engineerings, and Mathematics), social science and education, while ones in 4) Tech and 6) Social have one or two of these three.

The technical efficiency in DEA is defined as a ratio of the weighted sum of multiple outputs to a weighted sum of multiple inputs, and the weights are calculated by solving a linear programming in a configuration and an assumption model on the structure of return to scale (Cooper et al. (2011)). In this study, we use the output-oriented model where

Table 2.7: Japanese national universities.

Class	Universities	Number
1) Research	Hokkaido University, Kyoto University, Kyushu University, Nagoya University, Osaka University, The University of Tokyo, Tohoku University, Tokyo Institute of Technology	8
2) Large	Chiba University, Hiroshima University, Kanazawa University, Kobe University, Okayama University, Tokyo Medical and Dental University, University of Tsukuba	7
3) Middle	Akita University, Asahikawa Medical College, Ehime University, Gifu University, Gunma University, Hamamatsu University School of Medicine, Hirosaki University, Kagawa University, Kagoshima University, Kochi University, Kumamoto University, Mie University, Nagasaki University, Niigata University, Oita University, Saga University, Shiga University of Medical Science, Shimane University, Shinshu University, The University of Tokushima, Tottori University, University of Fukui, University of Miyazaki, University of the Ryukyus, University of Toyama, University of Yamanashi, Yamagata University, Yamaguchi University	28
4) Tech	Japan Advanced Institute of Science and Technology, Kitami Institute of Technology, Kyoto Institute of Technology, Kyushu Institute of Technology, Muroran Institute of Technology, Nagaoka University of Technology, Nagoya Institute of Technology, Nara Institute of Science and Technology, Obihiro University of Agriculture and Veterinary Medicine, The University of Electro-Communications, Tokyo University of Agriculture and Technology, Toyohashi University of Technology	12
5) Non-med	Ibaraki University, Iwate University, Nara Women's University, Ochanomizu University, Saitama University, Shizuoka University, Utsunomiya University, Yokohama National University	8
6) Social	Fukushima University, Hitotsubashi University, Osaka Kyoiku University, Shiga University, Tokyo Gakugei University, Wakayama University	6

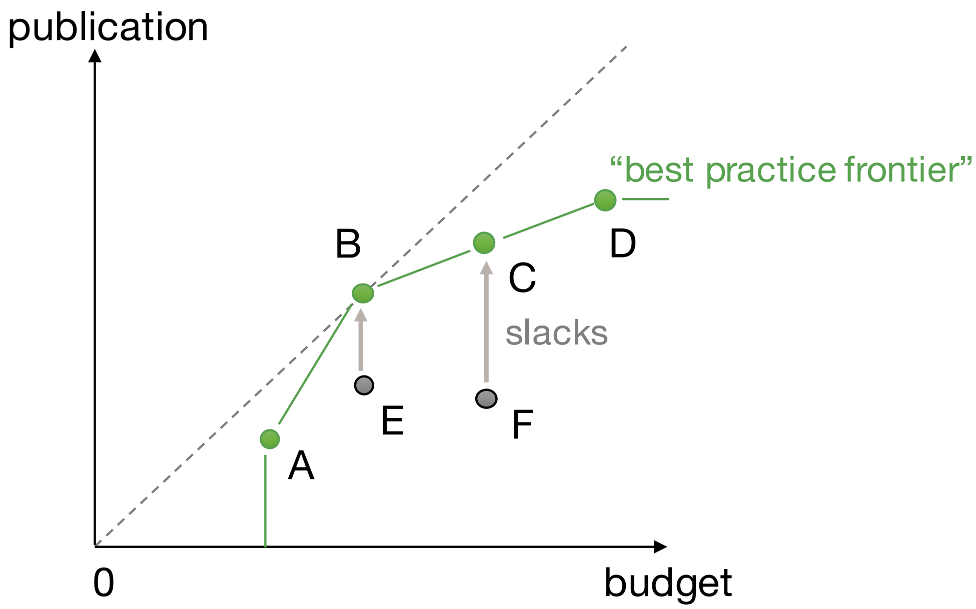


Figure 2.6: Schematic diagram of "best practice frontier" in data envelopment analysis.
As an example of 1-input 1-output model, the number of papers and the budget of the six universities (A to F) are plotted. The productivity of each university can be calculated by the output per input (*i.e.* publication per budget). The university B has the highest productivity among these six universities, and its productivity is equal to the slope of the straight line connecting the origin and the point B as indicated by a dashed line. For the output-oriented DEA, universities A, B, C, and D are defined as the best performer in this dataset thus they form the best practice frontier as illustrated in the green line. DEA efficiency of these universities is set to 1 by definition. The DEA efficiency of the university E, which shows smaller output than B although E has same input as B, is calculated by the ratio between the productivity of university E and B. It is always smaller than 1.

the linear programming model is configured to determine a potential output of decision-making unit given its inputs if it operated as efficiently as units on the best-practice frontier do as shown in Figure 2.6. We also assume variable returns to scale (Banker et al. (1984)).

As the benchmarking of universities, number of students, number of staffs including professors, operational expense, and investment to infrastructure are commonly used as inputs, and number of graduates, gross revenue, publication, and total amount of research grants are often used as outputs (Ahn et al. (1988); Avkiran (2001); Bhattacharyya and Chakraborty (2014); Castano and Cabanda (2007); Johnes (2006); Sinuany-Stern et al. (1994); Wolszczak-Derlacz and Parteka (2011)). In this study, we use number of principal investigators (including full professors, associate professors, assistant professors, and lecturers), number of PhD students, number of researchers (including post-docs), and R&D expenditure on goods as inputs, and total publication count and ratio between the number of the top 10% highly cited papers and total number of papers measured by Scopus as output. The number of paper published by a specified national university is counted as the whole count, *i.e.* a paper written by two authors whose affiliations are different but both are one of 69 national universities is double counted in our dataset. As well as the DEA efficiency ϕ^p for publication, we also compute DEA efficiency ϕ^d for diversification with use of diversity indices namely richness and evenness as listed in Table 2.8 as outputs. The scale economy of each university is also estimated by comparison of efficiencies calculated by different assumptions such as variable returns to scale ϕ_i, constant returns to scale ϕ_i^{CRS}, increase returns to scale ϕ_i^{IRS} and decrease returns to scale ϕ_i^{DRS}, where i represents a university ($i = 1 \sim 69$). When $\phi_j = \phi_j^{\mathrm{CRS}}$, the scale economy of university j is estimated as constant returns to scale. For the case of $\phi_j \neq \phi_j^{\mathrm{CRS}}$, the scale economy of university j is estimated as decrease (increase) returns to scale for $\phi_j^{\mathrm{DRS}} > \phi_j^{\mathrm{IRS}}$ ($\phi_j^{\mathrm{DRS}} < \phi_j^{\mathrm{IRS}}$). We set a one-year time lug between the three-year moving average of inputs and outputs.

Since the survey of expenditures and human resources of national universities has been done in every Japanese fiscal year (begins from April), not in the nominal year (begins from January), the actual minimum lug is 0.75 year. For example, the annual average of expenditure and number of researchers between April 1, 2001, to March 31, 2004, and the annual average of the number of papers published between January 1, 2002, and December 31, 2004, are used to compute efficiencies of the term 2002. In order to estimate the impact of R&D budget to the publication and diversification efficiencies, we attempted panel Tobit regression with several variables listed in Table 2.8. A schematic diagram of our data analysis is shown in Figure 2.6. Since the efficiencies calculated by DEA is regulated as $0 \leq \phi_i \leq 1$ by definition, we use a two-limit Tobit model (by setting the value for upper and lower limit)

$$\phi_i = \begin{cases} 0 & (\phi_i^* \leq 0) \\ \phi_i^* & (0 < \phi_i^* \leq 1) \\ 1 & (\phi_i^* > 1) \end{cases} \tag{2.20}$$

with a latent variable ϕ_i^* written as a quadratic model;

$$\phi_i^* = \beta_1 x_i + \beta_2 x_i^2 + e_i, \tag{2.21}$$

where β_1 and β_2 determine the relationship between independent variables and the latent variable and e_i is a normally distributed error. We use the university classification dummy and year dummy as control variables. It is confirmed that there is no influence of multicollinearity in the Tobit regression since correlation coefficients between any two of the five variables are less than 0.6.

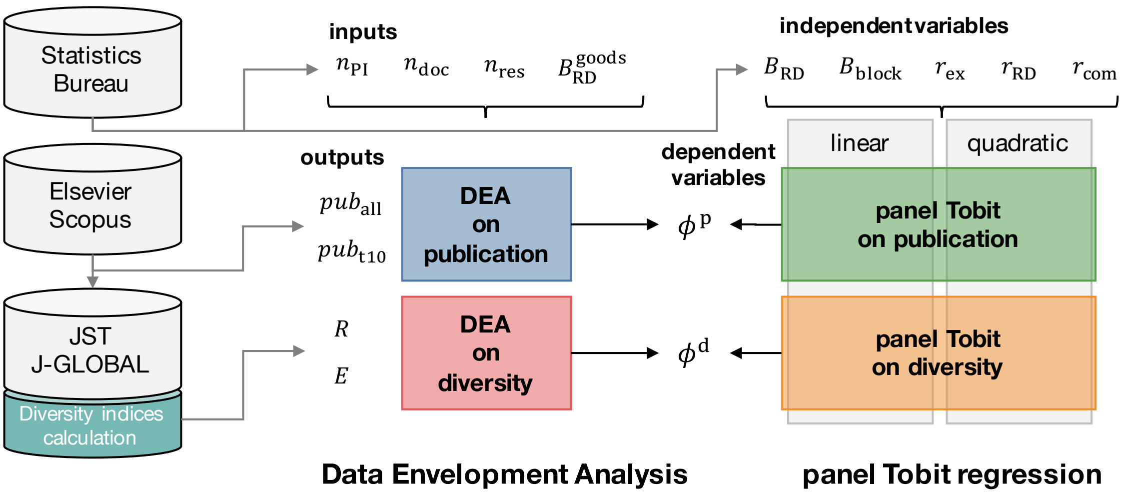


Figure 2.7: Outline of analysis.

All input variables, *i.e.* expenditures and human resources of national universities used in this study are based on the microdata of "the Survey of Research and Development, Ministry of Internal Affairs and Communications". All aggregation and statistical calculation have been done by the authors, independently from the ministry. The articles used here are identified by the affiliation (*i.e.* the name of national university) of the authors. In this thesis, only the articles commonly recorded in both Scopus and J-Global were analyzed. The number of papers and the top-10%-cited articles in each field defined by Scopus was counted in Scopus for each national university. The richness and evenness indices are computed by the number of classification codes and the number of papers with each JST classification codes in J-Global. First, two set (*i.e.* DEA on publication and on scientodiversity) of 4-input 2-output DEA are performed. Then, the panel Tobit regression on those two DEA efficiencies as dependent variables. Notice that the independent variables used here are different from that used for inputs of DEA. We used both a linear and quadratic model.

Table 2.8: Definition of variables used in the data envelopment analysis and the panel Tobit regression.

	Variable	Definition	Source
Data Envelopment Analysis			
Inputs	n_{PI}	Number of principal investigators	Statistics Bureau*
	n_{doc}	number of PhD students	
	n_{res}	number of post-docs & researchers	
	$B_{\mathrm{RD}}^{\mathrm{goods}}$	R&D expenditure on goods	
Outputs	pub_{all}	total publication count	Scopus & JGLOBAL
	pub_{t10}	Ratio of publication with top 10% citation to total publication	
	R	Richness (see the Method section)	
	E	Evenness (see the Method section)	

Panel Tobit regression			
dependent variable	ϕ^p	Technical efficiency for publication	(calculated by DEA)
	ϕ^d	Technical efficiency for diversification	
independent variable	B	Total expenditure	Statistics Bureau*
	B_{RD}	R&D expenditure	
	$B_{\mathrm{RD}}^{\mathrm{ex}}$	R&D expenditure supported by external grants	
	B_{block}	block grant per PI; $(B - B_{\mathrm{RD}}^{\mathrm{ex}})/n_{\mathrm{PI}}$	
	$B_{\mathrm{com}}^{\mathrm{ex}}$	external grants from private company	
	r_{ex}	ratio of external grant to the whole R&D budget; $B_{\mathrm{RD}}^{\mathrm{ex}}/B_{\mathrm{RD}}$	
	r_{RD}	ratio of R&D budget to total budget; $B_{\mathrm{RD}}^{\mathrm{ex}}/B$	
	r_{com}	ratio of external grants from private company; $B_{\mathrm{com}}^{\mathrm{ex}}/B_{\mathrm{RD}}^{\mathrm{ex}}$	

*All variables of budgets and human resources of national universities used in this study are based on the microdata of "the Survey of Research and Development, Ministry of Internal Affairs and Communications". All aggregation and statistical calculation have been done by the authors, independently from the ministry.

2.7 Propensity score matching

An approach often used in quantitative analysis on the policy effect is the Difference in Differences (DID) technique. One way of obtaining that treatment effect on publication performance by research grants would be to subtract publication count for non-awarded researchers from one for awardees. However, the allocation of research grants is unlikely to be at random. Indeed, in the case of KAKENHI and CREST program (as discussed in Chapter 5), awardees are selected by the peer-review of proposals and oral presentations. The awardee researchers may be more productive, have more reputations, or have better research career and position, and those variables can be referred at some point in the selection process. Therefore, the positive effect would be easily estimated by such simple comparison (cross-section estimation) although it is difficult to distinguish net treatment effect of the research grant from other unintended external causes as shown in Figure 2.8. The simple comparison of the performance of treatment group samples between before and after the specific grant awards cannot be understood as the treatment effect, too. To this extent, the direct comparison between the awardee and non-awardee researchers may be biased in terms of an intrinsic difference between the natures of researcher groups.

In order to mitigate this sampling bias, we attempted the propensity score matching (PSM) method to set appropriate samples. First, we ran Probit regression for the probability of adoption to the CREST program with several variables, such as the position, affiliation, publications and citations before the application to the grant, which seems to be relevance to adoption. The definition of variables is listed in Table 2.9. The participation probabilities, or the propensity scores, were calculated for each individual researcher. Second, we matched CREST awardees (treatment group) to one or more KAKENHI awardees (control group) by comparing their propensity score as illustrated in Figure 2.9. We dealt with the matching in two ways: the caliper matching (the caliper size is 0.03) and the kernel matching (the bandwidth is 0.001) under the common support condition. Then, we statistically tested the balance of the covariates across treatment (CREST awardee) and control (KAKENHI awardee) groups in the matched sample. We verified the balance by t-test on each variable as well as the likelihood-ratio test over all variables used in the Probit regression. The parameters used in the matching process written above were determined from the detailed analysis on the parameter dependency of the p-value of the likelihood-ratio test.

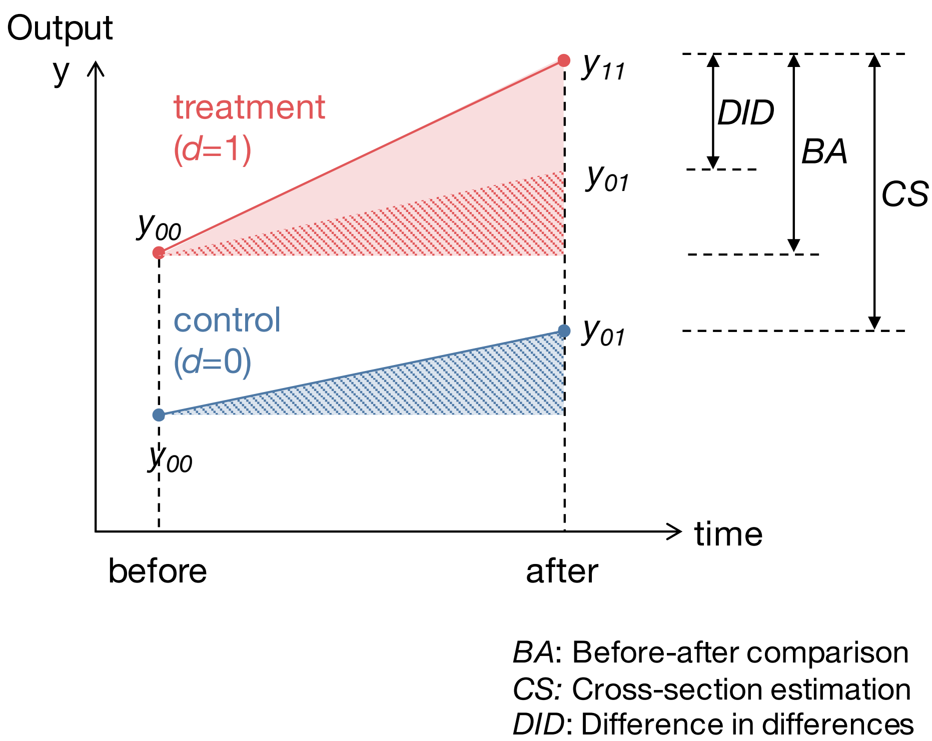


Figure 2.8: Schematic diagram of the difference in differences.
The simplest way to estimate the treatment effect is to calculate the difference between before and after treatment for the treatment group $(y_{11} - y_{00})$. However, this before-after comparison (BA) also includes the effect of the trend in a whole sample including the treatment group. The effect of the trend can be removed by calculating the difference between the BA of the treatment group and the control group, that is, the difference in differences (DID), namely $(y_{11} - y_{00}) - (y_{01} - y_{00})$. Notice that it is assumed that the effect of the trend of the treatment group and the control group are same (as depicted in the blue and red hatched triangles). The value y_{01} cannot be observed because this cannot happen, *i.e.* counter-factual. The cross-section estimation (CS) is also inappropriate to estimate the treatment effect as illustrated in the figure.

Table 2.9: Definition of variables.

variables	definition
d_{UT}	If the affiliation of researcher is the University of Tokyo, d_UT=1.
d_{FIU}	If the affiliation of researcher is either the former Imperial University, such as the University of Tokyo, Kyoto university, Tohoku university, Kyushu University, Hokkaido University, Osaka University, or Nagoya University, d_FIU=1.
d_{prof}	If the position of researcher is professor, d_prof=1.
d_{kanto}	If the affiliating institution is located in Kanto area, d_kanto=1.
d_{kansai}	If the affiliating institution is located in Kansai area, d_kansai=1.
$pub^{\mathrm{ave}}_{\mathrm{before}}$	Average publication count in three years prior to participation.
$pub^{\mathrm{ave}}_{\mathrm{going}}$	Average publication count in five years of participation.
$pub^{\mathrm{ave}}_{\mathrm{after}}$	Average publication count in three years after finishing the program.
$cite^{\mathrm{ave}}_{\mathrm{before}}$	Average citation count in three years prior to participation.
$cite^{\mathrm{ave}}_{\mathrm{going}}$	Average citation count in five years of participation.
$cite^{\mathrm{ave}}_{\mathrm{after}}$	Average citation count in three years after finishing the program.
$top10^{\mathrm{ave}}_{\mathrm{before}}$	Average publication count of top 10 % cited paper in three years prior to participation.
$top10^{\mathrm{ave}}_{\mathrm{going}}$	Average publication count of top 10 % cited paper in five years of participation.
$top10^{\mathrm{ave}}_{\mathrm{after}}$	Average publication count of top 10 % cited paper in three years after finishing the program.
$cpp^{\mathrm{ave}}_{\mathrm{before}}$	Average of citation per publication count in three years prior to participation.
$cpp^{\mathrm{ave}}_{\mathrm{going}}$	Average of citation per publication count in five years of participation.
$cpp^{\mathrm{ave}}_{\mathrm{after}}$	Average of citation per publication count in three years after finishing the program.
$cite^{\mathrm{max}}_{\mathrm{before}}$	Maximum of citation count in three years prior to participation.
$cite^{\mathrm{max}}_{\mathrm{going}}$	Maximum of citation count in five years of participation.
$cite^{\mathrm{max}}_{\mathrm{after}}$	Maximum of citation count in three years after finishing the program.

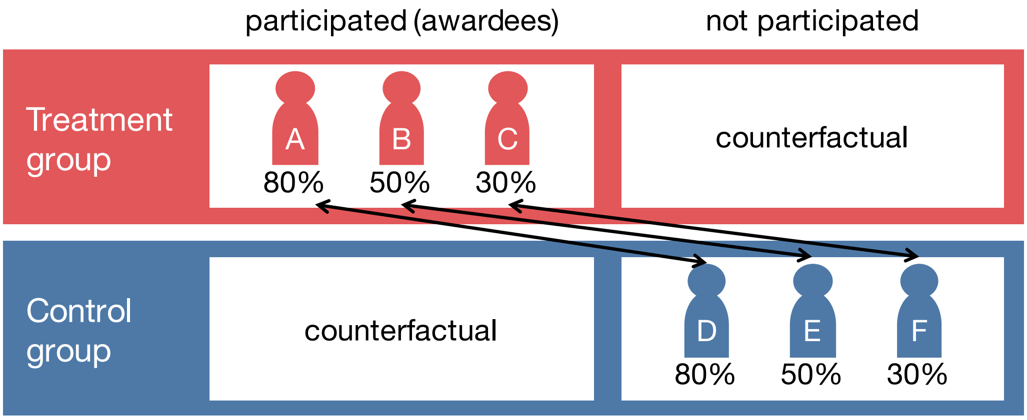


Figure 2.9: Schematic diagram of the propensity score matching.
If we can compare the treated and non-treated cases among the treatment group samples, the treatment effect can be estimated correctly. However, the former does not occur in reality (*i.e.* counterfactual). The treatment effect can be estimated by comparison between samples with similar probability of treatment, *i.e.* propensity score. In our case, the propensity score is equivalent to the awarding probability of CREST for each PIs. We compare the researcher A in the treatment group and the researcher D of the control group since they have same propensity score. In this case, the researcher D is a researcher who has a high probability (80%) but was not actually adopted by CREST. On the contrary, the researcher C is adopted despite the probability being estimated to be as low as 30%, and will be compared with the researcher F who has the same propensity score.

Chapter 3

Country-scale analysis

3.1 Analogy between scientodiversity and biodiversity

There are several bibliometrics studies (either implicitly or explicitly inspired by biodiversity) using an approach similar to that used in ecology (Mugabushaka et al. (2016); Stirling (2007); Wagner (2010)). The application of biodiversity indices, *e.g.* Gini-Simpson index, Shannon–Wiener index, and Leinster–Cobbold index, to the distribution of papers, has been verified as a powerful tool for investigation of interdisciplinarity. Besides diversity indices, the shape of the distribution of research subjects at the country level may give us new information about science activities. The shape of the distribution function has often been regarded as reflecting its origin. Just as a normal distribution is generated by the random additive process, the lognormal distribution and the power-law distribution both have generative processes (Limpert et al. (2001); Newman (2005)).

However, only a few studies have mentioned the shape of the distribution of research subjects as classified by sufficiently granular category codes, and thus understanding of the key mechanisms of scientodiversity has been limited. For example, studies of the statistical distribution of research subjects, the universality of their distribution over time and space, and mathematical models that generate observed distributions may contribute to an understanding of this 'science of sciences.' The mathematical framework for scientodiversity is still based on a naïve conceptual analogy between scientodiversity and biodiversity, which, despite the weak empirical evidence, raises primitive unanswered questions, notably the following. How are scientific topics distributed over disciplines in a given country? Does that distribution differ substantially from country to country, and if so how? or is that distribution similar across numerous countries? What factors determine that distribution and how might that distribution change over time? What kind of policy tool is suitable for the effective and efficient maintenance of scientodiversity? Answering these open questions using an approach similar to that used in biodiversity studies may yield fruitful insights on public policy on science.

To ensure the validity of the underlying analogy between biodiversity and scientodiversity, a revisitation of salient studies on biodiversity may afford an indirect understanding of the mechanism of the ecosystem of science in terms of the generative process of distribution. The key framework for quantitative analysis of biodiversity selected here is derived from the concept of island biogeography (MacArthur and Wilson (1967)). The island has been considered the most important platform for evolutionary ecology since Darwin visited Galapagos because islands can be regarded as closed ecosystems suitable for examination of the mechanisms of biodiversity (May (1975)). The closed nature of the innovation system in terms of national (Nelson (1993)), sectoral (Malerba (2002)), and regional (Chaminade and Plechero (2014)) levels, which are a foundation of today's innovation policy in many countries, suggests a possible analogy between the ecosystems of knowledge and living organisms.

Based on the outcomes of ecology studies on biodiversity, scientodiversity is expected to have different types of distribution patterns maintained by different mechanisms distinguishable by the fine shape of research subject distribution. This difference in types is also expected to appear as a change of slope in the subject-budget relationship, which can be regarded as comparable to the species-area relationship seen in biodiversity. The species-abundance distribution and species-area relationship are the most intensively investigated patterns in both theoretical and empirical studies of ecology (Hubbell (2001); May (1975)). The neutral model predicts both species-abundance distribution and species-area relationship base on realistic assumptions on population dynamics (Hubbell (2001)). This theory explains inflection points in the species-area curve by the change of dominant determinant of biodiversity. In a local spatial scale, the rate of encounter with new species mainly determines (the observed value of) biodiversity. This sampling process is sensitive to universality and rarity of species; thus the species-area curve presents a relatively steep slope. However, at the subcontinental scale, the encounter rate does not depend so much on relative species abundance, but rather on speciation rates, dispersal rates, and extinction rates and their equilibrium. Moreover, the slope of the species-area curve will recover on a much larger scale, *i.e.* continental and global scale, due to overcoming of dispersal biogeographical barriers formed by their evolutional history over the long term.

In this study, we investigated the distribution of research subjects in a bibliographic database to develop a framework of scientodiversity by comparison with that of biodiversity. We examine the following three hypotheses derived from an analogy with biodiversity. First, we examine the distribution of research subjects and test whether log-normal is the most appropriate function for describing that distribution. Although the skewed distribution of papers for specific subject areas has been reported and explained as a bias in databases (Bosman et al. (2006)), which may reflect an imbalance of author population, concentrated investments, and collection policy of databases, distribution shape has never been analyzed on the basis of a fine classification scheme. If research subjects can be regarded as analogous to biological species, the distribution of subjects might conform to the species-abundance distribution seen in biodiversity. Therefore, we also compute representative diversity indices, such as richness and Gini-Simson index (see the method section), which are often used for measurement of biodiversity, to demonstrate the feasibility of their application to the measurement of scientodiversity. We also examine whether or not the log-normal subject distribution is commonly observed in the distribution of publications in a number of countries.

Second, we verify the linear dependency of the number of subjects on research budget in a double logarithmic plot. This is found to be equivalent to the power function species-area relationship seen in biodiversity. The linear relationship in log-log plot is a mathematical consequence of lognormal species-abundance distribution. Thus, it cannot be a property unique to biodiversity. The power function is asymptotically expected, given the additional assumption that spatial density of species is constant (May (1975)). Therefore, our examination of linearity of the subject-budget relationship will serve to evaluate this assumption of equal access to resources among research subjects.

Third, we confirm an inflection point on the subject-budget relationship curve corresponding to change in the dominant determinant of scientodiversity. This change in the slope may imply a change in the underlying mechanism of scientodiversity, as seen in biodiversity (Hubbell (2001)). The relationship between publication performance and research budget at the country level has been investigated, in particular, to evaluate the efficiency of public investment (OECD (2016)). Although low diversity can be easily expected in a country with a small budget *(i.e.* a small publication), there is no study on function type of subject-budget relationship over many countries.

This study shows that science policy work for the preservation and promotion of the diversity of research would benefit greatly from an assessment of the analogy between scientodiversity and biodiversity. Just as environmental policy for biodiversity preservation at the country level often differs from policy for promotion of village level biodiversity because of difference of the underlying mechanism that maintains biodiversity in each scale, successful science policy in the case of large budgets may not work as well when applied to the case of small budgets if the maintaining mechanism of scientodiversity depends on the size of research. The results of this study will provide an evidence-based clue to means of avoiding such an inefficient science policy.

Table 3.1: Fundamental statistics of diversity indices.

Index	Country	Mean	Median	S.D.	Min	Max	N
Richness	All	779.2	227	1005.5	1	3209	219
	OECD	2408.3	2578	733.6	483	3209	35
Gini-Simpson index	All	0.9444	0.9886	0.1342	0	0.9984	219
	OECD	0.9976	0.9980	0.0010	0.9939	0.9984	35

3.2 Log-normal subject distribution

We evaluated the diversity of research subjects in articles by country, *i.e.* the location of the affiliating institutions of the authors. The histogram and empirical cumulative distribution function (see Figure 3.1(a)) indicates the skewed distribution of richness over 219 countries. More than 60 percent of the countries observed in this study show the relatively small richness, *i.e.* $R < 500$. The significant difference between mean ($R = 779.2$) and median (227) richness reflects the fact that the distribution of richness over the country is more skew than the normal distribution. For the 35 OECD member countries, represented as open circles in Figure 3.1(b), the median (2578) is larger than the mean (2408). This indicates that the variety of research subjects for a few countries is much larger than for the others. The observed skewed distribution of research subject, even for the OECD member countries, implies that the promotion of scientodiversity will be influenced by that country's economic development as shown in Table 3.1(see also Figure 3.2).

From the result of this study, scientodiversity of Japanese science, by means of the richness of research subject, is not inferior to that of Germany and the UK as contrastive to the result of previous bibliometric research (Igami and Saka (2016)). This difference perhaps due to the difference of dataset we use. Therefore, the analysis applying our method proposed in this thesis to their dataset, *i.e.* highly-cited articles only, is worth to be investigated in the future.

Figure 3.1(b) suggests that the Gini-Simpson index is dependent on richness. The fact that almost all countries are plotted in one curve close to the theoretical upper limit $1 - 1/R$ (solid line) implies that distribution of subject follows one universal statistical pattern for all countries, *i.e.* the Gini-Simpson index is almost uniquely determined by richness.

We test the suitability of fit on the observed distribution of classification code for 143 countries with $R \geq 100$. Comparison of the Akaike information criterion (AIC) among four statistical models (normal, log-normal, negative binomial, and gamma distribution) suggests that for all 143 countries, log-normal is the most appropriate distribution among

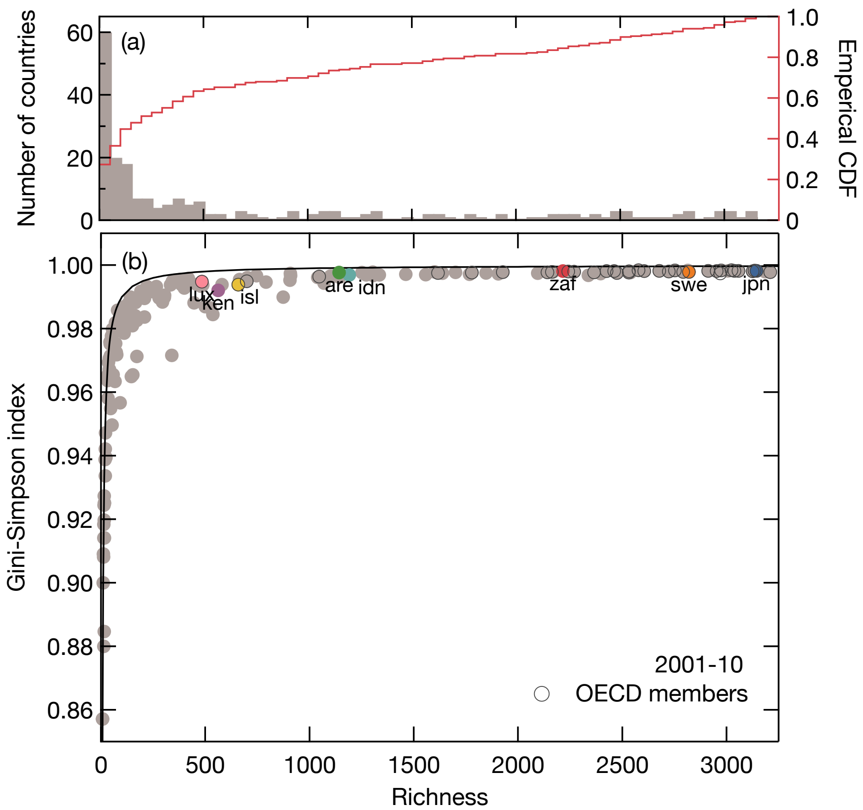


Figure 3.1: Richness and Gini-Simpson index.
(a) The histogram and empirical cumulative distribution function of the number of countries. More than 60 percent of the countries observed in this study show the relatively small richness, *i.e.* $R < 500$. (b) Gini-Simpson index as a function of richness for 219 countries. The open circles represent the 35 OECD member countries. The solid line represents the theoretical upper limit $1 - 1/R$. Richness and Gini-Simpson index are calculated based on the papers published between 2001 and 2010 registered in both J-Global and Scopus databases commonly.

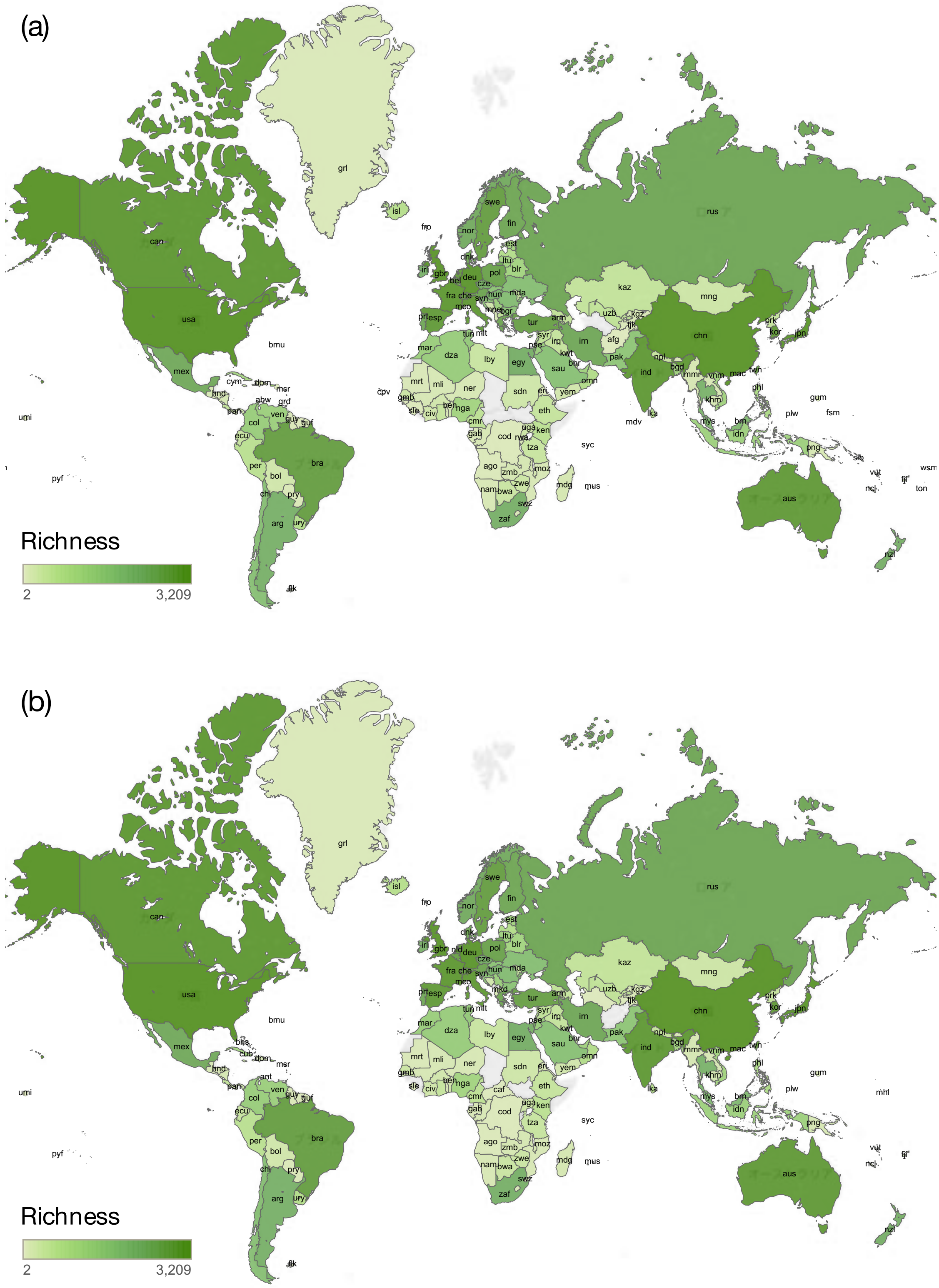


Figure 3.2: World map of richness in (a) 2000 and (b) 2010.
Richness, *i.e.* the number of classification codes attached to the papers published by a research institution in each country, is calculated based on the papers registered in both J-Global and Scopus databases commonly. As shown in the lower left legend, the shade of green indicates the magnitude of richness.

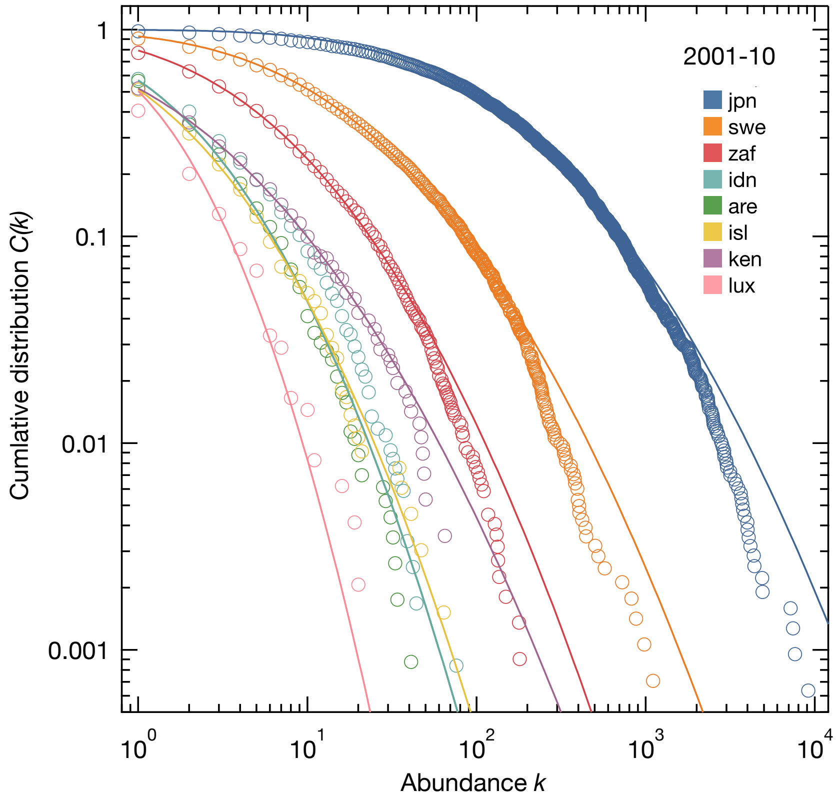


Figure 3.3: The cumulative distribution of research subject for eight example countries.
The cumulative distribution function for the classification codes is plotted by the open circles. The solid lines are the best fit functions by the lognormal distribution for each example countries as depicted by different colors. The abundance k, *i.e.* the number of papers, used in this plot is the count based on the papers published between 2001 and 2010 and registered in both J-Global and Scopus databases commonly.

the four types of distribution functions (see Table 3.2). Figure 3.3 shows the cumulative distribution of subject

$$C(k) = \frac{1}{R}\int_k^{\infty} S(n)dn, \tag{3.1}$$

where probability density function $S(n)$ represents the number of subjects containing n papers for eight example countries, Japan, Sweden, South Africa, Indonesia, United Arab Emirates, Iceland, Kenya, and Luxembourg (in descending order of R).

The log-normal distribution is the most popular tool used in biodiversity studies; this implies the coexistence of dominant species and relatively rare species in the same ecosystem. The problem from the policy point of view is whether such an inhomogeneous system is stable or not. Indeed, this question has not yet been answered; it is listed as one of the most important unanswered questions in ecology in the 20th century (May (1999)).

Table 3.2: The result of fitting.

country	Akaike information criterion (AIC)				Parameters		OECD
code	normal	lognormal	gamma	binomial	μ	σ	(M: member)
usa	55153.0	44994.0	510212.6	45523.8	5.0599	1.7008	M
chn	51238.7	40924.0	318642.7	41594.0	4.4501	1.7417	
Jpn	49248.1	40116.3	277864.7	40613.0	4.4181	1.7232	M
gbr	44096.1	35282.6	131194.5	35911.3	3.7461	1.5894	M
deu	45272.3	35446.3	158082.1	36187.8	3.7090	1.7209	M
can	40431.1	31813.6	91016.8	32473.4	3.3144	1.5991	M
fra	42044.8	32955.1	122099.5	33638.9	3.4568	1.7288	M
ita	38972.0	30723.2	84411.0	31419.4	3.1712	1.6308	M
ind	38835.5	29791.5	82734.3	30651.8	3.0211	1.6731	
aus	35280.1	27339.0	54608.9	28183.8	2.7692	1.5031	M
kor	40100.0	29533.1	88101.3	30612.5	3.0567	1.6415	M
esp	38460.5	29533.9	84118.1	30382.3	3.0733	1.6643	M
nld	34088.9	25889.5	49046.2	26807.6	2.6185	1.5026	M
twn	36050.9	26028.9	61751.3	27177.1	2.6974	1.5979	
swe	32485.2	24025.1	42366.2	25029.4	2.4500	1.4794	M
bra	31387.7	23891.8	40920.6	24810.4	2.4306	1.4843	
che	32679.2	24284.7	45106.2	25238.3	2.5137	1.5153	M
tur	29217.9	22301.5	33316.6	23179.5	2.3000	1.3940	M
bel	29864.7	22056.4	35082.7	23079.6	2.2648	1.4414	M
pol	29602.5	22227.0	37913.0	23172.3	2.3136	1.5192	M
aut	26332.2	19375.8	26311.6	20406.5	1.9992	1.3500	M
grc	25464.2	19155.1	25791.3	20117.3	1.9836	1.3659	M
rus	30238.3	21930.6	45896.4	22944.6	2.4000	1.6471	
fin	25523.0	18662.4	25740.3	19707.4	1.9349	1.3919	M
dnk	26452.0	18875.7	26666.1	19957.7	1.9802	1.3895	M

(to be continued)

country	Akaike information criterion (AIC)				Parameters		OECD
code	normal	lognormal	gamma	binomial	μ	σ	(M: member)
irn	25160.9	18355.1	25142.8	19362.1	1.9539	1.3786	
isr	26395.8	18843.6	27839.8	19911.3	2.0411	1.4155	M
prt	24585.1	18179.3	24794.6	19171.0	1.9577	1.3717	M
mex	24039.3	17707.7	24019.2	18694.1	1.9048	1.3882	M
hkg	26053.3	18336.8	28629.6	19456.6	2.0280	1.4604	
nor	22655.9	15977.3	19957.5	17094.7	1.7030	1.2924	M
sgp	26124.0	17807.2	29448.5	19050.6	1.9929	1.4875	
cze	22147.3	16134.1	21342.3	17114.6	1.8213	1.3692	M
egy	20790.8	14772.3	17799.1	15843.7	1.6151	1.2646	
nzl	20191.6	14504.6	17254.2	15548.8	1.5875	1.2589	M
zaf	18927.7	13703.0	15584.8	14725.9	1.4900	1.2013	
arg	20700.1	15067.4	19174.6	15988.5	1.7634	1.3408	
hun	19833.0	14327.5	17196.2	15280.9	1.6565	1.2676	M
irl	20002.5	14260.3	17439.9	15266.5	1.6710	1.2705	M
tha	19307.0	13319.8	16154.6	14447.2	1.5315	1.2574	
svn	15430.7	10746.3	11766.6	11745.6	1.2571	1.1177	M
mys	16416.5	11054.2	12679.7	12167.9	1.3173	1.1783	
rou	15493.5	10742.5	12173.7	11758.1	1.3291	1.1741	
chl	14609.4	10404.5	11542.1	11282.4	1.3685	1.1473	M
ukr	16534.5	11449.2	14524.8	12422.5	1.5314	1.3341	
sau	12431.6	8587.1	9262.2	9531.7	1.0750	1.0310	
bgr	13114.7	9161.4	10124.0	10040.2	1.2468	1.1403	
svk	12753.9	8953.3	9824.8	9808.3	1.2293	1.1283	M
hrv	11144.1	7793.1	8377.5	8673.6	1.0143	1.0008	
pak	11973.3	7951.8	8695.7	8890.9	1.0745	1.0606	

(to be continued)

country code	Akaike information criterion (AIC) normal	lognormal	gamma	binomial	Parameters μ	σ	OECD (M: member)
tun	11306.1	7760.5	8487.2	8588.6	1.1443	1.0956	
dza	9998.9	6837.6	7415.1	7599.7	1.0987	1.0426	
col	9237.6	6050.2	6600.0	6866.4	0.9138	0.9840	
jor	7971.3	5597.7	6048.8	6337.5	0.8411	0.8926	
mar	9013.2	6094.3	6600.7	6844.2	0.9797	0.9954	
srb	7506.1	5136.6	5593.7	5890.2	0.7830	0.8603	
ven	8143.1	5596.8	6054.1	6310.1	0.9037	0.9433	
idn	7863.1	5227.9	5710.4	5971.3	0.8475	0.9310	
ltu	8239.7	5360.9	5903.7	6130.6	0.9187	1.0037	
are	6458.0	4445.6	4844.4	5133.1	0.7342	0.8118	
bgd	7373.5	4818.0	5287.2	5538.7	0.8206	0.9206	
vnm	7168.2	4672.4	5113.4	5363.7	0.8171	0.8922	
blr	8446.6	5519.2	6165.4	6227.7	1.0660	1.0993	
nga	8231.4	4984.2	5592.4	5790.7	0.8923	1.0183	
yux	6228.8	4118.9	4525.6	4800.5	0.7102	0.8224	
est	7082.0	4707.3	5132.5	5353.6	0.8823	0.9483	M
cub	6916.4	4456.7	4900.7	5135.1	0.8244	0.9044	
kwt	5343.8	3576.8	3917.7	4147.0	0.7299	0.8323	
lbn	4674.2	3269.4	3551.3	3802.9	0.6568	0.7576	
pri	5928.4	3660.9	4090.4	4314.6	0.7364	0.8718	
phl	6671.6	3638.5	4147.1	4353.0	0.7606	0.9067	
cyp	5005.9	3018.8	3403.9	3616.7	0.6616	0.8396	
omn	3775.7	2504.1	2779.2	3011.6	0.5741	0.7297	
ury	4042.2	2786.4	3041.4	3217.9	0.7232	0.8518	
lva	4172.1	2623.9	2914.3	3105.1	0.6568	0.8173	M

(to be continued)

country	Akaike information criterion (AIC)				Parameters		OECD
code	normal	lognormal	gamma	binomial	μ	σ	(M: member)
lka	3699.3	2346.2	2633.7	2830.4	0.6056	0.7705	
isl	4130.6	2518.1	2839.6	3015.2	0.6757	0.8317	
per	3232.2	1859.0	2155.8	2352.1	0.5043	0.7232	
ken	4017.2	2539.0	2837.9	2958.3	0.8180	1.0185	
arm	3966.9	2103.7	2448.3	2586.0	0.6679	0.8797	
cmr	3606.5	1904.0	2223.0	2371.0	0.6134	0.8058	
uzb	3276.0	1687.9	1989.9	2147.0	0.5353	0.7631	
geo	2636.5	1613.5	1842.6	2000.8	0.5324	0.7417	
lux	2291.2	1389.1	1589.1	1775.2	0.4511	0.6465	
mkd	2365.2	1413.9	1640.4	1821.7	0.4334	0.6813	
syr	2557.1	1522.6	1742.3	1905.0	0.5123	0.7098	
cri	2431.4	1537.0	1737.4	1884.7	0.5458	0.7419	
kaz	2592.2	1239.8	1510.5	1689.0	0.4185	0.6363	
irq	1571.0	1046.3	1176.9	1360.1	0.3793	0.5761	
qat	1891.6	1123.9	1300.9	1476.2	0.3978	0.6205	
gha	2154.6	1451.3	1601.1	1723.1	0.5999	0.7660	
eth	2015.8	1411.3	1547.5	1666.6	0.6035	0.7553	
aze	2161.2	1378.2	1552.6	1675.4	0.5649	0.7697	
prk	2535.4	1574.7	1785.9	1886.5	0.6635	0.9287	
mda	2656.9	1709.6	1904.4	1994.6	0.7649	1.0028	
npl	1871.3	1228.0	1374.7	1505.2	0.5057	0.7124	
pse	1505.3	967.3	1098.3	1267.6	0.3842	0.5795	
tto	1398.4	940.4	1056.7	1223.4	0.3728	0.5733	
tza	2026.2	1319.4	1471.3	1576.1	0.6162	0.7959	
lby	1190.3	711.2	836.4	1025.3	0.2864	0.5110	

(to be continued)

country	Akaike information criterion (AIC)				Parameters		OECD
code	normal	lognormal	gamma	binomial	μ	σ	(M: member)
ecu	2458.1	1180.7	1444.0	1543.2	0.5461	0.7949	
bwa	1389.1	855.7	979.8	1126.4	0.3982	0.5838	
zwe	1530.4	892.9	1037.3	1153.4	0.4394	0.6695	
sen	1556.2	894.2	1034.6	1149.4	0.4493	0.6675	
bhr	1465.5	723.0	902.6	1040.2	0.3238	0.5921	
mng	1308.1	737.2	862.4	984.6	0.3887	0.6047	
uga	1533.9	994.4	1114.2	1196.3	0.5891	0.7971	
jam	992.6	641.1	725.8	848.6	0.3720	0.5549	
sdn	1035.4	649.4	745.1	851.7	0.3757	0.6105	
scg	685.4	373.4	460.3	627.8	0.2077	0.4343	
civ	850.0	577.6	644.2	742.4	0.4032	0.5688	
bol	983.5	577.1	676.5	779.0	0.3665	0.5927	
bih	616.8	346.6	421.8	584.4	0.2140	0.4256	
bfa	898.4	554.9	638.2	726.3	0.3915	0.6150	
pan	1232.9	767.9	875.8	935.7	0.5819	0.8483	
yem	727.7	444.4	517.8	627.0	0.2952	0.5246	
fji	601.6	376.5	436.4	546.6	0.2775	0.4861	
ben	955.5	605.9	685.0	737.0	0.5748	0.7835	
mac	960.5	436.1	562.0	637.5	0.3232	0.6151	
mlt	641.9	380.5	445.7	529.5	0.3196	0.5464	
ncl	885.4	598.9	664.6	712.4	0.5809	0.8302	
mus	619.6	394.9	454.5	529.7	0.3376	0.5769	
mco	989.2	537.3	642.0	687.4	0.5235	0.8285	
lie	721.3	396.6	482.8	547.7	0.3324	0.6424	
mdg	895.1	471.3	568.1	616.5	0.4490	0.7738	

(to be continued)

country	Akaike information criterion (AIC)				Parameters		OECD
code	normal	lognormal	gamma	binomial	μ	σ	(M: member)
mwi	504.7	327.1	370.4	442.2	0.3474	0.5249	
zmb	384.6	240.5	280.7	378.2	0.2409	0.4399	
rom	432.3	239.5	293.0	385.0	0.2200	0.4557	
mmr	416.4	258.9	301.9	377.4	0.2708	0.4884	
khm	518.7	337.3	383.5	439.8	0.3760	0.5986	
nic	413.1	220.3	271.9	355.6	0.2314	0.4560	
png	475.7	316.4	356.0	406.3	0.3904	0.6020	
brb	491.6	301.0	348.1	402.1	0.3649	0.5851	
glp	399.6	241.4	284.0	347.3	0.2784	0.5142	
brn	453.0	262.6	313.1	368.8	0.2969	0.5705	
gtm	384.5	246.7	285.2	341.1	0.2804	0.5454	
mli	517.3	322.7	365.8	408.3	0.4492	0.6488	
alb	273.0	122.7	166.9	272.4	0.1404	0.3624	
ner	416.7	222.7	273.8	332.3	0.2659	0.5256	
cog	345.3	238.3	265.8	314.4	0.3582	0.5273	
nam	299.0	187.4	217.9	284.1	0.2548	0.4567	
lao	392.7	235.6	277.2	328.6	0.3073	0.5539	
moz	285.8	163.4	197.3	269.0	0.2086	0.4358	

Table 3.3: Fundamental statistics of $n_{\max}$, n_{mode} and their ratio γ.

variable	country	Mean	Median	S.D.	Min	Max	N
$n_{\max}$	All	5.478	5.250	3.074	1.250	14.75	193
	Type I	3.575	3.250	1.628	1.250	7.750	120
	Type II	8.606	8.250	2.206	2.750	14.75	73
n_{mode}	All	2.297	0.250	2.970	0.250	10.25	193
	Type I	0.250	0.250	0.000	0.250	0.250	120
	Type II	5.661	5.750	2.252	1.250	10.25	73
γ	All	0.297	0.143	0.282	0.032	0.935	193
	Type I	0.090	0.077	0.049	0.032	0.200	120
	Type II	0.638	0.659	0.138	0.333	0.935	73

Contrary to the conventional theory that predicts that a complex community is unstable (May (1972)), many species coexist stably in complex networks of interspecific interactions in nature. Recently, a simulation study revealed that a certain balance of types of interaction can stabilize population dynamics of many species (Mougi and Kondoh (2012)). This stimulated ecologists to strive to identify the ecological mechanism that determines and maintains species diversity. According to studies of biological ecosystem stability, our complex research ecosystem cannot be regarded as stable without considering distribution patterns and network structures of interaction among disciplinal species. The stability of science ecosystems should be considered as a policy issue because it is one of the rationales for the promotion of the diversity of research (Stirling (2007)).

The lognormal distribution represents a system coexisting with the very rare elements and much more popular common elements. Those are of the extremely low probability in a normal distribution. Figure 3.4 shows the packed bubble representation of the number of articles in each research subject classified by JST classification code for eight example countries. A bubble represents a specific reseach subject, and its size represents the number of articles, which the classification code is attached, in a specific year indicated on the top of the figure. Thus, the number of the bubble is equal to the richness. The characteristic of the lognormal distribution is clearly visualized in Figure 3.4.

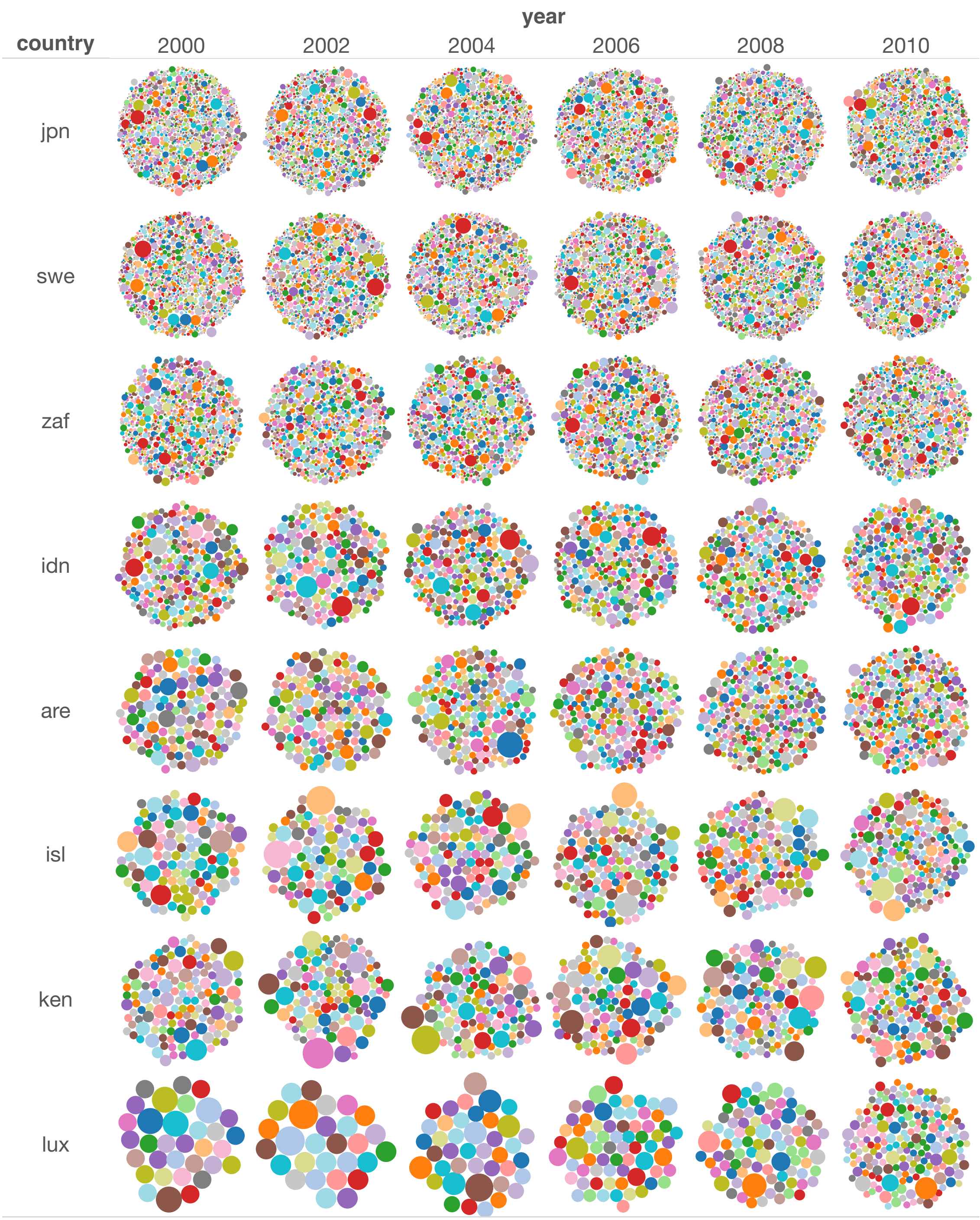


Figure 3.4: Packed bubble representation of the distribution of research subjects. Each color-coded bubble represents a classification code, and its size represents the number of articles to which the classification code is attached. This plot shows the biannual time change of the distribution of research subjects in eight example countries from 2000 to 2010. A few huge bubbles coexist with many extremely-small bubbles for the country with larger richness. The number of papers used in this plot is the count based on the papers registered in both J-Global and Scopus databases commonly.

3.3 Individual distribution

In addition to the distribution of subjects $S(n)$, the distribution of a number of papers $nS(n)$ should also be considered. The positional relationship between those two distributions is characterized by γ; the ratio between the log of the mode on $nS(n)$ and the log of the maximum number of individual papers in $S(n)$ (see also Chapter 2). The relation between γ and richness is shown in Figure 3.5(a). It is clear that the countries can be divided into two groups, namely types I and II, by means of γ. Type I countries show relatively smaller γ which decreases with the increase of R. Type II countries with larger R display wider dispersion of γ which slightly increases as R increases. Figure 3.5(b) to (q) show $S(n)$ and $nS(n)$ for 16 example countries in Preston's octaves (Preston (1980)). For the small R countries, n_{mode} is not clearly observed and perhaps the value is less than observation limit $n = 1$ (Preston's veil line). Then, the value γ for Type I countries is inversely proportional to $\log_2 n_{\max}$ which slightly increases with the increase of R.

For the Type II countries, the range of the observed value, $0.333 \leq \gamma \leq 0.935$, is consistent with the reported non-canonical ($\gamma \neq 1$) values for biodiversity cases (May (1975)), but no countries with γ value were greater than 1 were observed for distribution of research subjects. This suggests the existence of some mechanism, which regulates $n_{\mathrm{mode}} < n_{\max}$, specific to scientodiversity and not present in biodiversity. The effort to understand the biological mechanism that determines γ should be helpful to identify the mechanism of the distribution of scientific subjects. Fundamental statistics of $n_{\max}$, n_{mode}, and γ are listed in Table 3.3.

The parameter γ characterizes the shape of the lognormal distribution. A special condition defined by $\gamma = 1$ is called "canonical", in which the shape, *i.e.* S_0 and σ in eq. 2.5, will be uniquely determined by a single parameter R. This special case happens only when the most popular research subjects contain the highest number of papers in total, *i.e.* $\log_2 n_{\max} = \log_2 n_{\mathrm{mode}}$. In our observation, the parameter γ was always smaller than one. This means that the number of papers in the most popular subject i, which is expected to be only one specific subject since $S(n_i) = S(n_{\max}) = 1$, is smaller than the total number of papers in certain unpopular subjects, in which the number of papers is n_{mode}. In other word, when one picks up a paper from a specific dataset and finds it belongs in a subject i, the most probably the subject i is the most popular research subject i but the expected number of papers of the subject i cannot be $n_{\max}$ but must be smaller than $n_{\max}$.

3.4 Richness-budget relationship

We plot richness as a function of expenditure on R&D (ERD) averaged over the ten year period from 2001 to 2010, at 2010 US dollar prices, in Figure 3.6. The observed correla-

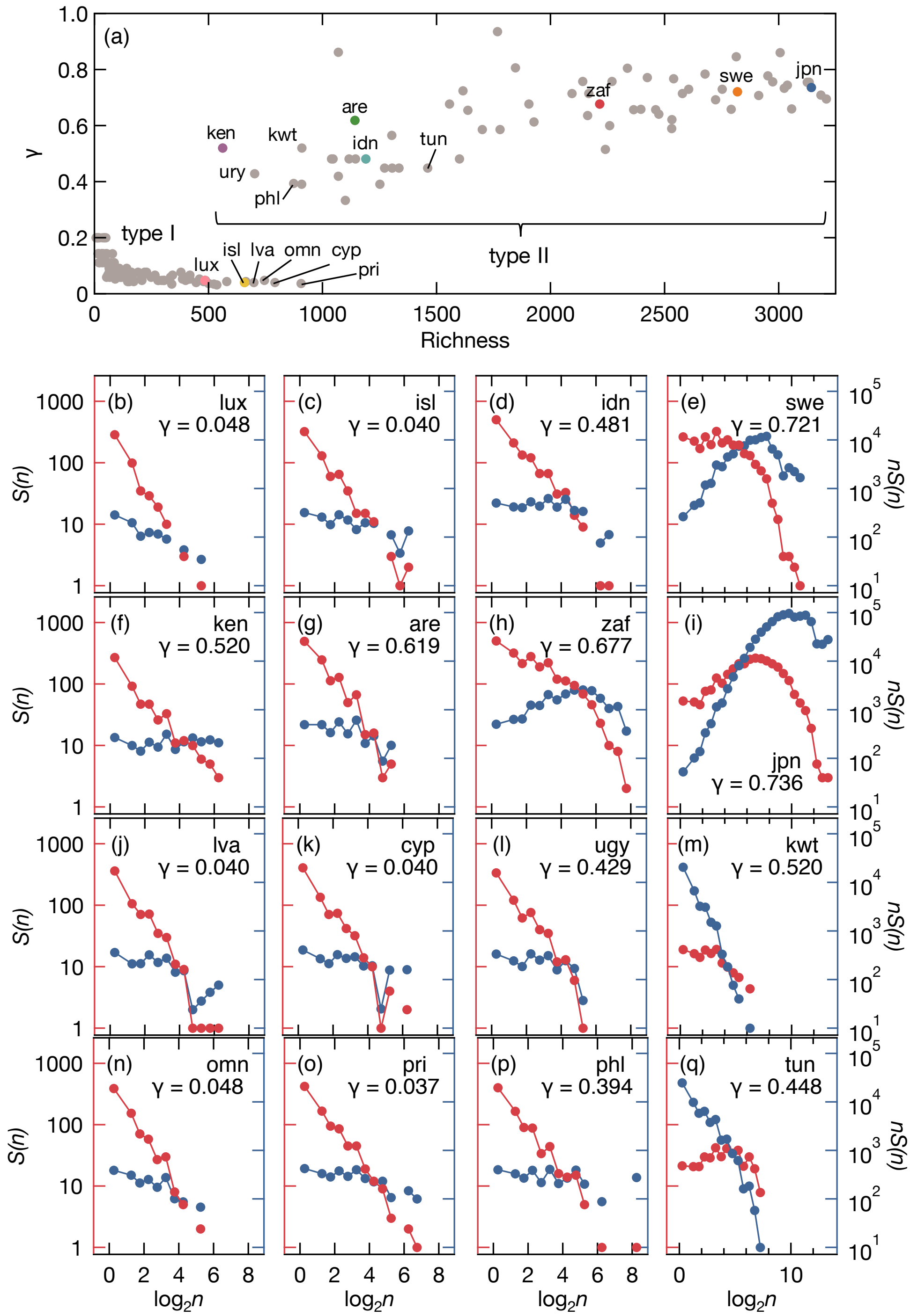


Figure 3.5: Subjects and papers distribution in the Preston's octave plot.
(a) The parameter $\gamma \equiv \log_2 n_{\mathrm{mode}} / \log_2 n_{\max}$ as a function of richness, where n_{mode} represents the mode of $nS(n)$ and $n_{\max}$ represents the log of the maximum number of individual papers in $S(n)$. The relation between γ and richness suggests that the countries can be divided into two groups, namely types I and II. (b) to (q) The distribution $S(n)$ (red marks and line, read in the left axis) and $nS(n)$ (blue marks and line, read in the right axis) for 16 example countries in Preston's octaves (Preston (1980)). The number of papers used in this plot is the count based on the papers published between 2001 and 2010 and registered in both J-Global and Scopus databases commonly.

Table 3.4: The country level comparison among example countries.

Country	R	$1-\lambda$	ERD	μ	σ	γ	Type
Japan	3,143	0.9982	183,500	4.4181	1.7232	0.736	II
Sweden	2,819	0.9979	15,800	2.4500	1.4794	0.721	II
South Africa	2,215	0.9982	2,753.3	1.4900	1.2013	0.677	II
Indonesia	1,190	0.9969	407.89	0.8475	0.9310	0.481	II
United Arab Emirates	1,142	0.9976	-	0.7342	0.8118	0.619	II
Iceland	658	0.9939	339.16	0.6757	0.8317	0.040	I
Kenya	562	0.9921	220.45	0.8180	1.0185	0.520	II
Luxembourg	483	0.9948	793.56	0.4511	0.6465	0.048	I
Tunisia	1,461	0.9971	251.05	1.1443	1.0956	0.448	II
Kuwait	909	0.9967	120.04	0.7299	0.8323	0.520	II
Philippines	873	0.9899	189.03	0.7606	0.9067	0.394	II
Uruguay	702	0.9961	124.60	0.7232	0.8518	0.429	II
Puerto Rico	906	0.9953	455.92	0.7364	0.8718	0.037	I
Cyprus	791	0.9946	84.912	0.6616	0.8396	0.040	I
Oman	745	0.9967	-	0.5741	0.7297	0.048	I
Latvia	699	0.9950	120.53	0.6568	0.8173	0.040	I

tion between R and ERD strongly resembles the species-area relationship seen in biodiversity studies (MacArthur and Wilson (1967)). Notice that ERD is defined as "current and capital expenditures (both public and private) on creative work undertaken systematically to increase knowledge, including knowledge of humanity, culture, and society, and the use of knowledge for new applications" according to the OECD's Frascati Manual (The World Bank (2017)). In this definition, R&D includes basic research, applied research, and experimental development but not higher education. Thus, this can be regarded as a proxy measure of input corresponding to the publication as an output in macroscopic viewpoint.

In the case of biodiversity, species-area relationship is derived from species-abundance distribution by assuming uniformity of density of individual organisms. For the asymptotic case ($R \gg 1$), the slope z of species-area curve $R = CA^z$ becomes $1/4\gamma$ for $\gamma > 1$ or $1/(1+\gamma)^2$ for $\gamma < 1$, where C and z are positive constants and A represents area (May (1975)). The slopes $z = 0.25$ and 0.8 are equivalent to $\gamma_{\mathrm{est}} = 1$ and 0.12, respectively (see the eye-guide lines in Figure 3.6). Those observed slopes mean that if you doubled the budget, the number of research subject will be expected to increase by 74% in a country with small R, but it will be only 7.2% in the country with larger R. The consistency among observations (Figure 3.5(a)) and $\gamma_{\mathrm{est}} \simeq 0.12$ suggests that deterministic mechanism of scientodiversity for the Type I countries can be inferred from island biogeography studies. However, the discrepancy between observation and estimation for the Type II countries implies that there exist other mechanisms, different from those for type I countries, that determine the diversity of research topics.

Preston expected the existence of biological origin for $\gamma = 1$ since only this class of lognormal model is consistent with observed species-abundance distributions and species-area relationships (Preston (1962)). However, today it is known that power law is also shown in species-area curve calculated from other types of species-abundance distribution, such as general ($\gamma \neq 1$) lognormal, broken-stick, and power function (Irie and Tokita (2012); May (1975)). Therefore, the linear dependency in log-log plot shown in Figure 3.6 cannot distinguish whether this phenomenon is derived from some mechanism embedded in science or just a mathematical consequence from the lognormal distribution. The linear dependency observed in both smaller and larger ERD countries implies that the density of research subjects may be constant within each ERD ranges. The density of research subjects, *i.e.* probability to encounter new research subjects if one expands research space, which may be determined by the budget, is expected to be higher for Type I countries than that for Type II, thus the slope of richness-ERD curve for Type I countries is steeper than that for Type II ones.

The key characteristic of the richness-ERD curve is the inflection point of its slope around ERD$\simeq$ 1 billion USD and $R \simeq 1000$. The analogy with the species-area relation-

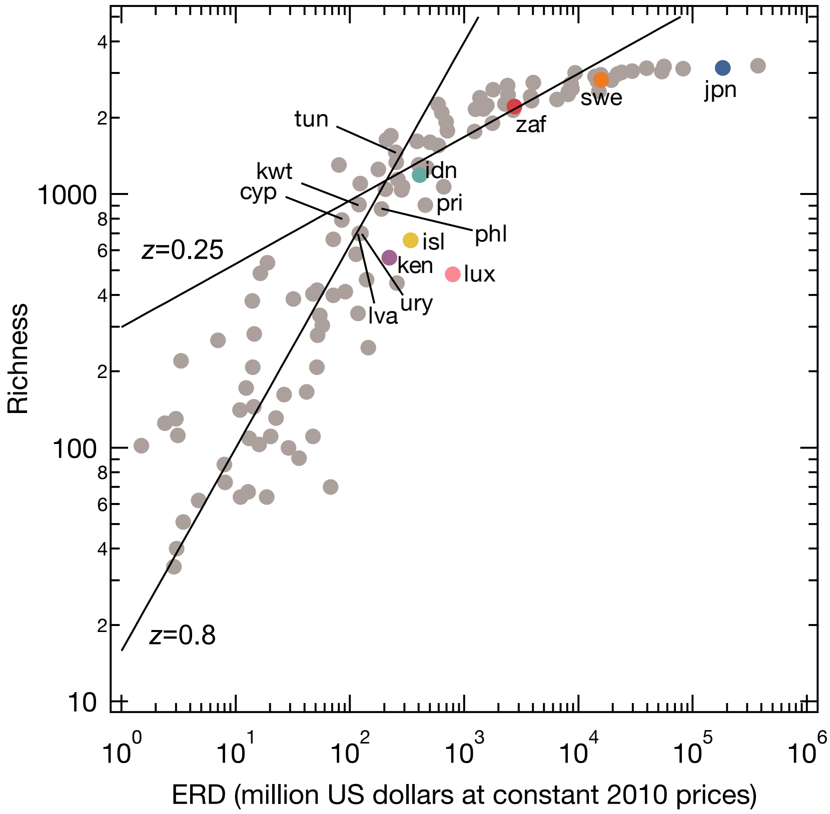


Figure 3.6: Richness-ERD relationship curve.
Country-level richness as a function of expenditure on R&D (ERD) averaged over the ten year period from 2001 to 2010, at 2010 US dollar prices. Richness, *i.e.* the number of classification codes attached to the papers published by a research institution in each country, is calculated based on the papers published between 2001 and 2010 and registered in both J-Global and Scopus databases commonly. The expenditure on research and development (ERD) has been retrieved and calculated from World Development Indicators (The World Bank (2017)). The richness for 16 example countries, such as Japan, Sweden, South Africa, Indonesia, United Arab Emirates, Iceland, Kenya, Luxembourg, Tunisia, Kuwait, Philippines, Uruguay, Puerto Rico, Cyprus, Oman and Latvia, are indicated in the colored circle. The solid lines are the eye-guide for $R \propto ERD^{0.25}$ and $R \propto ERD^{0.8}$, respectively.

ship in biodiversity suggests that the dominant determinant of scientodiversity changes at this scale. This implies that depending on the scale of the object to be managed, a different science policy approach is necessary for the promotion of scientodiversity. The boundary value (ERD$\simeq$ 1 billion USD) is equivalent to the annual budget size of funding agencies, top research universities, and institutes in developed countries. Thus, both new-subjects-oriented and equilibrium-aware research strategies are essential for the promotion of scientodiversity on different scales.

The classification of countries based on the richness-ERD curve (Figure 3.6) is mostly corresponding to the classification by γ (Figure 3.5(a)). Type I countries generally have small budget size and the number of research subjects, *i.e.* richness, but have the potential to expand their diversity efficiently by the increase of their budgets. The fitting of $S(n)$ by lognormal function, measured by AIC, seems relatively poor compared with that for Type II countries, and the distribution of individual papers $nS(n)$ tends to be in the lower right in Preston's octave plot. It is clear that there are more than 100 research subjects addressed in a small number of papers ($n < 10$) in those Type I countries. Smaller σ obtained from fitting by lognormal distribution is consistent with smaller Gini-Simpson index. Type II countries have large research budgets and number of research subjects. However, it is no longer possible to efficiently increase the number of subjects by increasing the budget. In that sense, the system appears to be in an equilibrium state. The lognormal distribution can be confirmed also in Preston's octave plot for $S(n)$, and certainly, the mode of $nS(n)$ is smaller than $n_{\max}$, *i.e.* $\gamma < 1$.

This classification of Types I and II and that shown in Figure 3.6 did not always match. For example, Kenya is classified as Type II by $\gamma = 0.520$ (Figure 3.5(a)), but seems to belong to the group with the larger budget in the richness-ERD relationship (Figure 3.6). In contrast, Iceland and Luxembourg, which have budget size (and R) comparable to that of Kenya, are classified as Type I. The method used here to calculate n_{mode} from the probability density function rather than a cumulative one is susceptible to noise. Thus, it is possible that the estimation of n_{mode} was too large for Kenya, and therefore the calculated γ was also too large.

As shown in Table 3.4, Kuwait, Philippines, and Uruguay are countries with relatively small R but large γ like Kenya (see also Figure 3.6). On the other hand, γ was estimated as small (*i.e.* Type I), even for R of about 500 to 800, for Cyprus, Oman, and Latvia. In this study, no common factors in these countries were found that could constitute a cause of observed classification mismatch between Figure 3.5 and Figure 3.6. It can be inferred from the results of this study that scientodiversity can be determined by at least two different mechanisms for two extreme cases, where the budget is quite large or quite small, and probably both mechanisms coexist in intermediate scale (*i.e.* R is between 500 and 1000). One possible explanatory scenario is that the balance between these mechanisms falls to one side by some factor other than R and budget. The result will be observed as a difference in γ. The further research, including an international comparison of science policy at the country level, is needed.

Although not discussed here, one of the most promising approaches to the promotion of research diversity is international collaborative research. A positive relationship between international collaboration, which is often measured as co-authorship involving different institutions in different countries, and citation impact have been found at the country level (Confraria et al. (2017); OECD (2016)). For example, the relatively large richness of Indonesia may be explained by the country's high rate of international collaboration compared to that of Puerto Rico, which has a similar size budget (Table 3.4). In contrast, Tunisia, with a lower collaboration intensity (Confraria and Godinho (2014)), shows higher richness than Kenya with a higher collaboration intensity despite their similarity on the ERD scale. Similar things can be seen in Type I European countries; for example, Iceland and Luxembourg, which are known for their high international co-authorship ratio (OECD (2016)), have the richness of approximately $R \sim 500$, approximately half of the value expected from the ERD scale. Thus, the impact of international co-authorship on scientodiversity is not simple in character, and a more detailed analysis of the structure and content of international collaborative network is called for.

3.5 Conclusions and policy implications

Comparison of our characterization of the distribution of research subjects over countries with the distribution of biological species suggests that scientodiversity is analogous structurally to biodiversity. We show subjects-abundance distribution and subjects-budget relationships which correspond to the species-abundance distribution and species-area relationships seen in biodiversity, respectively. We also examine the relation between subjects-abundance distribution and subjects-budget relationships, and the relation is well within that which can be inferred by analogy from biodiversity studies (May (1975)). This structural similarity between two empirical patterns in scientodiversity and two in biodiversity strongly suggests that the mechanisms underlying them should be similar. The observed change of slope of subject-budget plot (see Figure 3.6) suggests that the mechanism for determining diversity varies depending on the scale of research.

There are many important concepts in evolutionary ecology for which until now, no equivalent has been found in the domain of the science of sciences, for example, equilibrium among speciation, dispersion, and extinction, food web, ecological niches, mutualistic interspecific network in biodiversity studies. Although equivalence between the key determinant of biodiversity and those of scientodiversity is not well examined in this study, the findings may indicate a fruitful direction for understanding the system of science and the promotion of scientodiversity. One of the most highly insightful concepts for science policy is the keystone species, defined as a species that has a disproportionately large impact on its environment relative to its abundance (Paine (1995)). *Pisaster ochraceus* (predatory starfish), *Enbydra lutris* (sea otter), and *Castor Canadensis* (beaver) are well-recognized examples of keystone species (Power et al. (1996); Wagner (2010)). The innovation ecosystem will dramatically change if a keystone research subject is removed, even though relative abundance was small. Further investigation on the role of each research subjects in their inter-subject network is thus needed.

Differences between biodiversity and scientodiversity may also give important implications on scientometrics studies. A crucial question should arise regarding the definition of “species” of science. The taxonomic classification scheme used in this study resembles Linnaean taxonomy in terms of hierarchy. Other ways of defining biological species such as morphology, sexual reproduction, and ecological niche may propose a new usable definition of research subjects. Parallel to a recent success in molecular phylogenetics, a quantitative classification scheme to objectively identify science disciplines should accelerate research on scientodiversity. Bottom-up approaches, such as text mining and the clustering of scientific papers, promise to enable the construction of modern classification scheme for science research topics. The definition of species affects the distribution pattern and diversity indices. The lognormal shape of the distribution is not guaranteed for different definitions of subjects. Thus, from this study alone, it cannot be determined whether the origin of the lognormal distribution of research subjects is an intrinsic property of scientific research or a trivial mathematical consequence derived from the classification scheme on the database. Similar to that the appropriateness of a definition of species is a challenging problem in the ecological study (May (1999)), it is worth comparing the shape of the distribution of research subjects with using various classification schemes.

Further examination of the analogy between the area in biogeography and budget in science is also necessary. ERD indicates the scale of research from aspects of both numbers of researchers and size of the research budget. It is known that more than half of ERD is researcher labor costs. According to the findings of ecology studies of biodiversity, the area can be reverse defined as a unit in which one can assume homogeneity of species (Hubbell (2001)). The density of research subjects may be regarded as constant per researcher rather than per research budget. Thus, further investigation on the subjects-researchers curve promises to present new implications for the promotion of scientodiversity. The structural similarity between biodiversity and scientodiversity also proposes several policy implications. First, research universities, institutions, and funding agencies should pay attention to both new subjects and equilibrium among existing ones when they make a decision on their resource allocation with a magnitude of around 1 billion USD. The inflection point of richness-ERD curve (Figure 3.6) implies that two different factors may influence the scientodiversity in a certain range of richness (*i.e.* $500 \leq R \leq 1000$). Second, for scientifically small countries, such as the type I countries shown in Figure 3.5(a), the fact that a local region in a continent shows a larger number of species than isolated islands, even in equivalent area (May (1975)), implies that international interaction in scientific activity may improve the scientodiversity of the nations involved, even with the same R&D budget size. Extrapolation of the richness-ERD line for type II countries estimates that $R \simeq 500$ should be achievable with ERD$\simeq$100 million USD (Figure 3.6). Third, recovery of the slope of the species-area curve is known to occur in a much larger area, *i.e.* continental and global scale. This phenomenon can be explained as the overcoming of dispersal biogeographical barriers formed by their evolutional history over the long term (Hubbell (2001)). Thus, international collaboration among large countries with rich scientodiversity ($R > 1000$) should push science toward a new phase of diversification. According to the observed inflection point of the richness-ERD curve, the underlying mechanism of scientodiversity may vary depending on the budget size. Thus, promoting scientodiversity in country level may not be accomplished by simply scaling up a grant program that promotes interdisciplinarity of research teams. And vice versa, the creation of new knowledge, which often rays across the traditional disciplinal border, will not be easily stimulated by facilitating national level science policy.

In the history of scientometrics, the statistical property of research articles has attracted many researchers and provided implicative laws such as Lotka's law, Gibrat's law, and Bradford's law. These statistical laws have counterparts in other fields such as ecology, seismology, medicine, economics, and social sciences (Limpert et al. (2001); Newman (2005)). The analogy we present here is based on a quantitative approach to diversity and statistical models that have been used in both bibliometrics and evolutionary ecology for a long time. Our approach, incorporated from the studies of ecology, itself is not new for scientometrics but provides an integrative view for the science of sciences with mobilizing various disciplines. In that framework, the history, sociology, and economics of science, which until now have been based on innovation policy, should be reconsidered from the viewpoint of an "ecology" of science.

Chapter 4

University-scale analysis

4.1 Japanese national universities

The Research University plays important role in national innovation ecosystem (Cole (2009)). Benchmarking based on their research activities such as publication, patent, and university-industry collaboration have revealed a high performance of American research universities in some rankings (Quacquarelli Symonds Limited (2017); Times Higher Education (2017)). In Japan, National universities are the central player for basic research. More than 70 % of Ph.D. students belong to national universities and almost 80% of academic papers published by researchers whose affiliation is located in Japan are produced by the researcher in the national universities (The Japan Association of National Universities (2017)).

The stagnation of publication performance and decline of the diversity of research have emerged as a policy concern in Japan. The bibliometric study and a questionnaire survey of Japanese researchers indicate reconstruction of grant system may cause the negative impacts on both publication and diversification (Igami and Saka (2016); National Institute of Science and Technology Policy (2015)). A comprehensive survey on Japanese national universities in terms of both budget and publication suggested that (1) total amount and (2) distribution of R&D budgets are associated to the observed decline of publication performance of Japanese national universities (Toyoda (2015)). The report concludes that the sluggish public research funding from the late 1990s should be the primary cause of the downturn in publication while the concentration of research grants to few universities may reduce resource availability for the rest of universities and thus total publication may be influenced by a balance in the resource distribution. The importance of (1) total R&D budgets is also indicated by a study to evaluate research activity of the national universities by the growth accounting methodology (Aoki and Kimura (2014)). The study shows that distortion of resource allocation has a smaller impact on papers productivity than the total budgets shared by all national universities has. The study reveals that 2.0% net growth of publication can be explained by 14.5% declining of productivity, 16.0% growth of total research budgets, 1.1% improvement of resource misallocation, 1.3% by deflation, and 1.9% reduction of labor costs between 2005 and 2009.

The concentration of research grants happened through the two major transformations, the reduction of operating expense subsidies (*i.e.* institutional grant including personnel expenses) and the expansion of competitive research findings, in Japanese grant system. The operating expense subsidy is distributed based on the number of students and professors, while competitive research grants tend to be concentrated to specific researchers. Thus, the conversion from the institutional grant to competitive ones resulted in enlargement of financial difference among national universities. This transformation pushed Japanese researchers into the competitive environment to get external grants since internal research budget supported by their own institution was no longer enough (Sunami (2017)). Since the university can also acquire indirect expense that can be used freely by universities, the university management was often incentivized to the acquisition of external research grants.

In this study, we investigated the research efficiency of Japanese national universities in terms of the quantity and diversity of their publication using data envelopment analysis (DEA) for data between 2001 and 2012. First, we confirmed the change of (1) total amount and (2) balance of R&D budget by using government statistics. Second, we evaluate technical efficiencies of national universities in terms of both publication and diversification by DEA methodology. Finally, we test the impact of the balance of research grants to the technical efficiencies by Tobit regression. Our results confirmed that difference among universities in terms of inputs, outputs, and efficiencies. Tobit regression of efficiencies implies the importance of (2) balance rather than (1) total R&D budgets. Our results contribute to not only the foundation of a strategy for individual universities but also provide useful suggestions for nation-wide resource allocation.

4.2 Inputs and outputs

Figure 4.1 shows the time change of inputs and outputs of whole universities. All figures are the moving average over 3 years. As shown in Figure 4.1(a), R&D expenditure (B_{RD}) accounts for roughly half of total expenditure (B), and about 80% of R&D expenditure is derived from the internal budgets. This internal budget is derived from the Management Expenses Grants ("uneihi koufukin" in Japanese) by the government. Overall expenditure has increased by about 10% in the past 10 years. As shown in Figure 4.1(b), the growth rate (*i.e.* value in 2002 is set to 1) of B became 10%, while $B^{\mathrm{in}}_{\mathrm{RD}}$ decreased. Although the number of PIs and doctor students did not change much, researchers including post-doctoral fellows showed a sharp increase of about 40%. This implies the increase of the number of the research projects, funded by an external grant, that temporary hires those researchers. Meanwhile, the number of publication has increased more than 30% as a whole with a plateau as shown in Figure 4.1(e). The fact that the number of papers increased despite the fact that R&D expenditure hardly increased means that the paper productivity per research expenditure has been remarkably improved. Richness of research subjects across the university (notice that it is not the sum of richness of individual universities) has hardly changed. On the other hand, the average of richness of individual universities grew by about 20%. This implies that small universities have not worked on novel research themes (that have never been investigated by large research universities), but worked on a safe (popular) research subjects.

The increase in university total expenditure shown in Figure 4.1 is about 300 billion yen, which may be due to an increase in income of university hospitals. The expenditure on R&D B_{RD} includes the labor costs of faculty, but this contains both research and education work time, and thus B_{RD} may represent the upper bound of the input money to R&D in university. Also, since the increase in the number of papers includes the increase of international co-authoring papers, the observed increase of the number of papers will not simply mean improvement in the productivity of Japanese researchers. In addition, the increase in the number of papers since around 2010 is thought to be due to the increase in the paper in the field of clinical medicine as pointed out in the previous study (Toyoda (2015)). The increase in this clinical medicine paper is thought to be due to an increase in doctors in university hospitals included in n_{res}, which increased by about 3000 people. However, the personnel expenses for those doctors are often not recorded as the research expenses. For these reasons discussed above, the simple ratio between input and output may not be a good indicator for research productivity of Japanese national university.

The Lorenz curve of the share of publication (pub_{all}), R&D expenditure (B_{RD}), the block grant (B_{block}), and number of PIs (n_{PI}) in 2010 are shown in Figure 4.2. The distribution is skewed in the order of B_{block}, B_{RD} , and pub_{all}. This is consistent with the past report that examined the distribution of Grants-in-Aid for Scientific Research (KAKENHI) and the number of papers (Shibayama (2011)). The Lorenz curve of B_{block} is approximately equal to the curve of n_{PI}. Gini coefficient for B_{block}, estimated by the area (indicated by gray hatching in Figure 4.2) between Lorenz curve and the line of equality (dotted line), is around 0.45.

Both n_{PI}, B_{block}, and B_{RD} in Figure 4.2 contain both education and research aspects. In small and medium-sized national universities, the proportion of the education part is larger than that of large universities, so the Lorenz curve of these inputs become gentler than the curve of the number of articles, which is the output only from research activity.

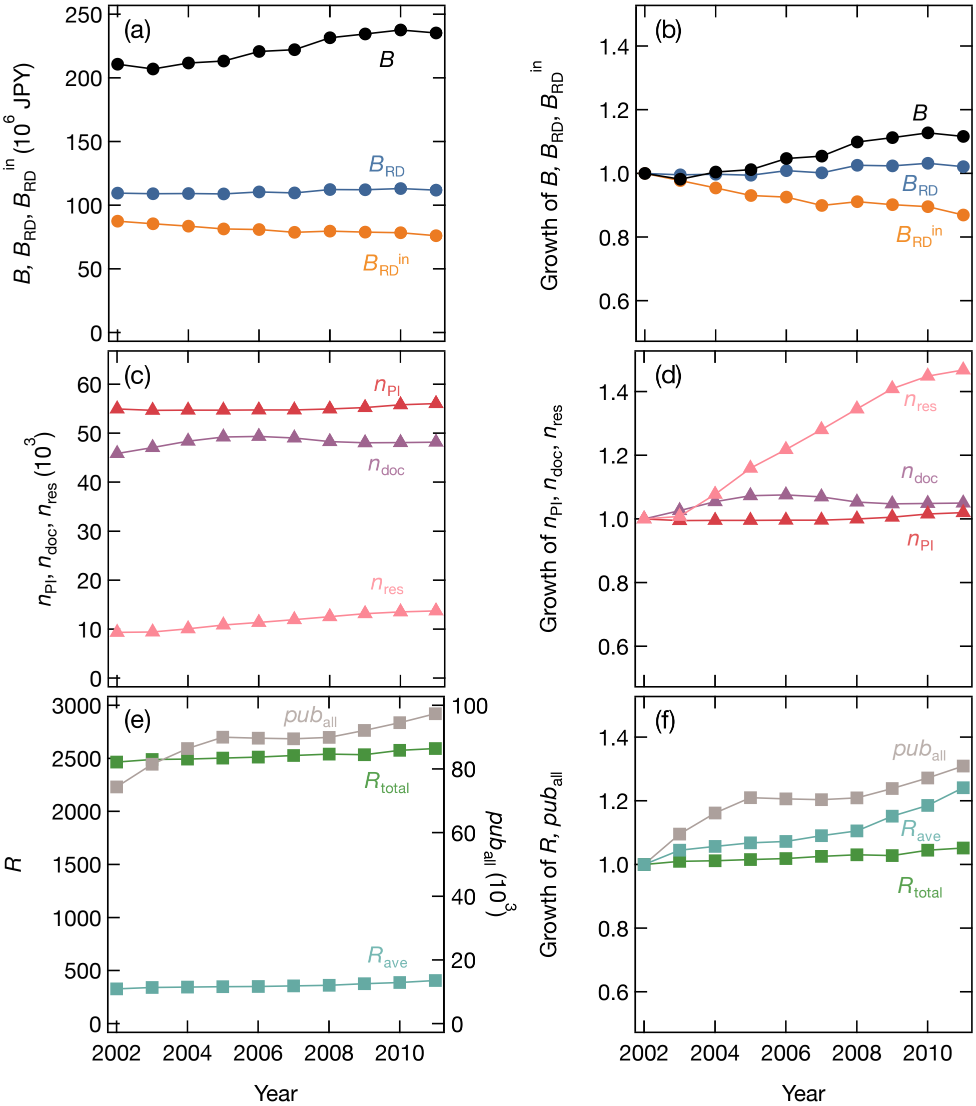


Figure 4.1: Time-change of overall inputs and outputs.

(a) Time change of the total expenditure B, the total R&D expenditure B_{RD}, and the total internally funded R&D expenditure $B^{\mathrm{in}}_{\mathrm{RD}}$ for 69 Japanese national universities. The internal fund is derived from the Management Expenses Grants ("uneihi koufukin" in Japanese) by the government. (b) The growth ratio of B, B_{RD}, and $B^{\mathrm{in}}_{\mathrm{RD}}$ compared with the values in 2002. Overall expenditure has increased by about 10% in the past 10 years. (c) Time change of the number of PIs n_{PI}, the number of doctor students n_{doc}, and the number of post-doc researchers n_{res}. Notice that those numbers are the head count. (d) The growth ratio of n_{PI}, n_{doc}, and n_{res}. compared with the values in 2002. (e) Time change of the total number of publication pub_{all}, the total richness of research subjects R_{total} and the richness averaged over 69 universities R_{ave}. Notice that R_{total} is not the sum of richness of individual universities, thus $R_{\mathrm{ave}} \neq R_{\mathrm{total}}/69$. (f) The growth ratio of pub_{all}, R_{total}, and R_{ave} compared with the values in 2002. All input data plotted in (a), (b), (c) and (d) are based on the microdata of "the Survey of Research and Development, Ministry of Internal Affairs and Communications". The number of papers and the richness index used in those plots are the count based on the papers registered in both J-Global and Scopus databases commonly. All values are moving average over 3 years.

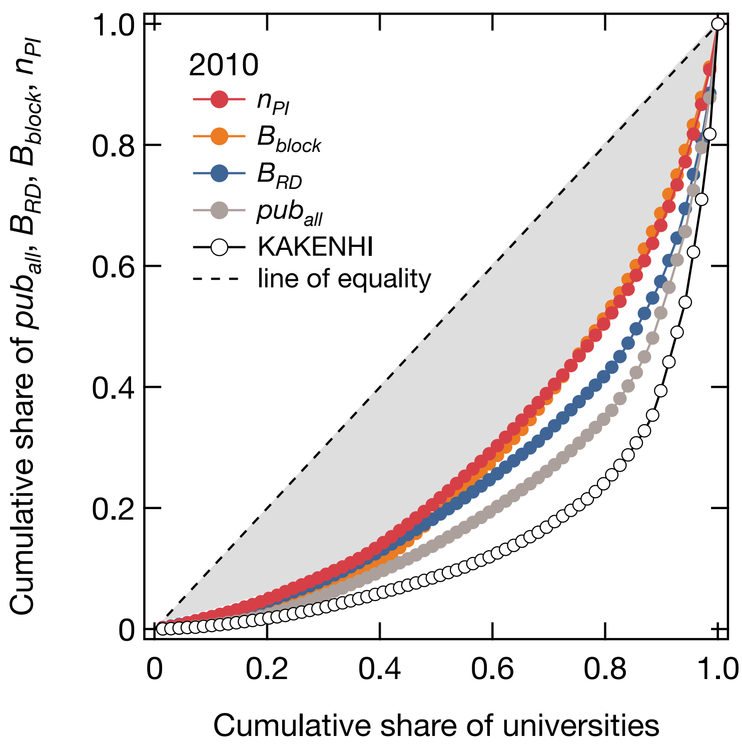


Figure 4.2: Lorenz curve of the pub_{all}, B_{RD}, B_{block}, n_{PI} and KAKENHI at university level in 2010.
The distribution is skewed in the order of B_{block}, B_{RD}, pub_{all} and KAKENHI. The skewed distribution of KAKENHI grant has been consistently reported by the previous study (Shibayama (2011)). The value B_{RD}, B_{block}, n_{PI} are based on the microdata of "the Survey of Research and Development, Ministry of Internal Affairs and Communications". The number of papers pub_{all} is the count based on the papers registered in both J-Global and Scopus databases commonly. The grant allocation amount for FY2010 of the Grants-in-Aid for Scientific Research (KAKENHI) was downloaded from the website of Japan Society for the Promotion of Science ("Japan Society for the Promotion of Science").

The productivity, *i.e.* output per input, seems higher for larger universities but it is not. We also plot the curves of the Grants-in-Aid for Scientific Research (KAKENHI) as input. The curve of KAKENHI is much skewed than that of the number of papers, as consistent with the result of the previous research (Shibayama (2011)).

The change of expenditure in terms of amount and structure have happened in differently in each type of national universities between 2002 and 2011. For the universities in 1) Research category, total expenditure has been grown but the proportion of R&D expenditure took almost flat at around 60% as shown in Figure 4.3. The breakdown of the R&D part has undergone a structural shift. The ratio of external grants has increased from approximately 30% to 40% between 2002 and 2011. Most of this increase comes from competitive research grants given by government rather than the research grants from private companies, which is less than 20% and slightly decreasing its proportion. The operational subsidies from the government to national universities has reduced approximately 1% per year since the 2004 reorganization of national universities as independent administrative institutions (Toyoda (2015)). However, according to the expenditure data shown in Figure 4.3, the universities in 1) Research category, as an average, could compensate this decrease by earning external grants mainly from the government. This shift from the institutional grant to competitive and external ones is well consistent with the policy plan which intended to make the research environment more competitive atmosphere.

For the universities in 2) Large category, the situation did not go well. The average total expenditure for the large universities, which is approximately half of one for the research universities, indeed slightly increase between 2002 and 2011, while R&D expenditure decrease at the same time, then the proportion of R&D expenditures have dropped from 60% to 50% as an average over 7 national universities in this category. The increase of weight for governmental research grant has shown as same as for the universities in 1) Research category, but the ones in 2) Large category have not been so successful to get sufficient external research grants that can compensate the loss of the budget cut of 1% per year. The observed increase in total expenditure may be associated with the growth of university hospital in terms of both income and expenditure.

For the universities in 3) Middle category, the total expenditure increased slightly between 2002 and 2011, while R&D expenditure has been almost no change, and then the proportion of R&D budgets have decreased as a result. For the universities of both 4) Tech and 5) Non-med categories indicate the relatively high ratio of R&D expenditure (~70%) simply because they do not have university hospital. The ratio of external grants to total R&D expenditure (r_{ex}) have grown between 2002 and 2011, but the contribution of private companies slightly decreases as compared to one of government. The universities of 6) Social category show relatively low dependency on the external grants ($r_{ex} \sim 0.1$) while the percentage of R&D expenditure to the total expenditure is similar to that of the universities in 4) Tech and 5) Non-med categories.

Scientodiversity also varies from university classification. The number of articles with level 1 classification code (the coarsest classification in JST classification) is summarized as a heat map for each classification of the university (Figure 4.4). 1) Research classification universities show many papers in all fields, notably B: Physics, C: Chemistry, and E: Biological sciences classification. This does not mean that these universities are concentrating on these fields. Rather, it merely reflects the characteristics of each field as the number of the article. 2) Large and 3) Middle universities have a profile similar to 1) Research university, but 4) Tech university has obviously different distribution, *e.g.* G: Medicine. This is consistent with the difference in the department composition of each university (Table 2.6). There seems to focus on I: System and Control Science, J:

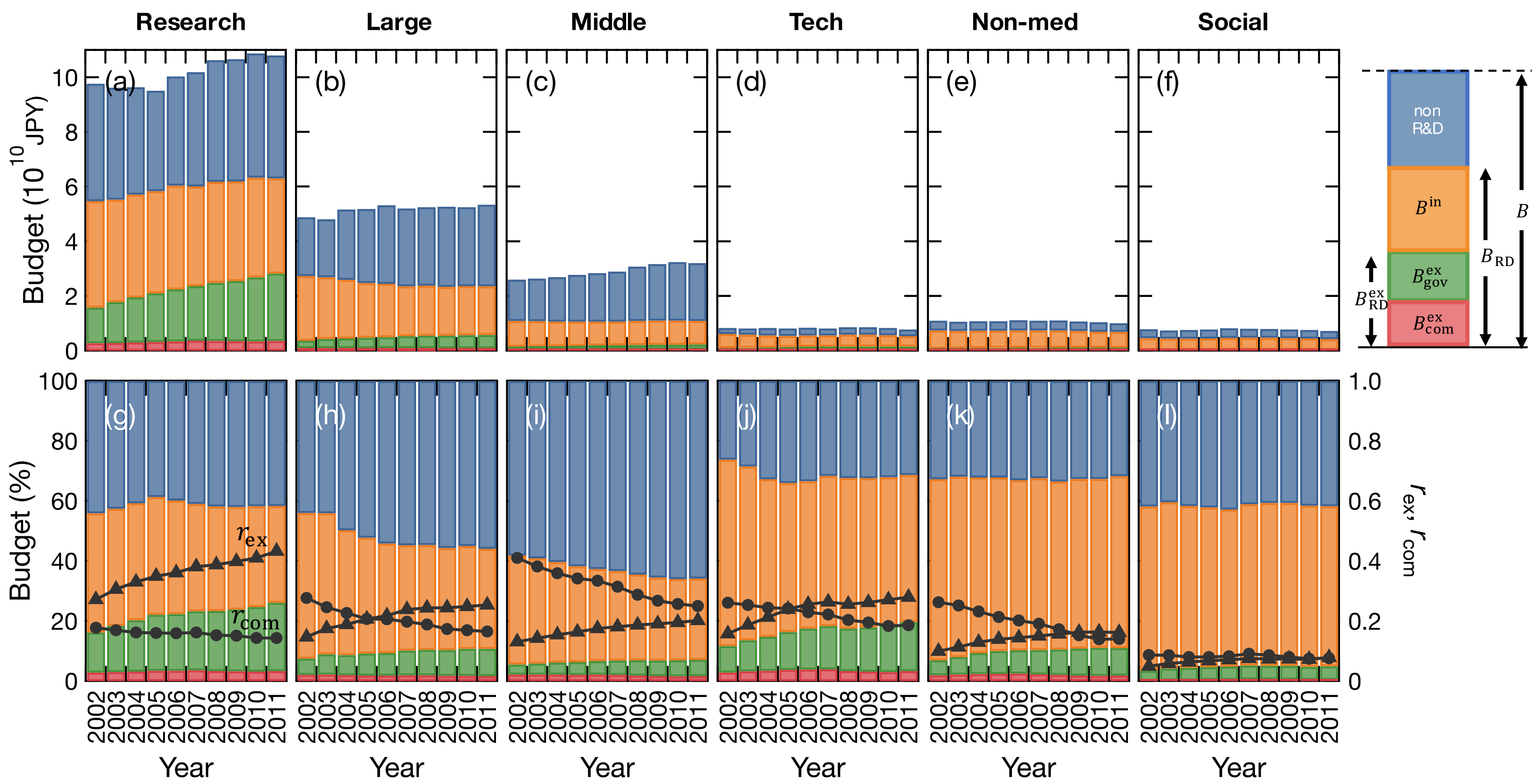


Figure 4.3: Time-change of the structure of expenditure.

The nominal value of expenditure ((a) to (f)) and the proportion ((g) to (l)) of the non-R&D expenditure (that is, $B - B_{RD}$), the internally funded R&D expenditure B^{in}_{RD}, the R&D expenditure funded by the government B^{ex}_{gov} and the R&D expenditure funded by private company B^{ex}_{com} for 6 categories of university. The ratio $r_{ex} = (B^{ex}_{gov} + B^{ex}_{com})/B_{RD}$ and $r_{com} = B^{ex}_{com}/(B^{ex}_{gov} + B^{ex}_{com})$ are also plotted for (g) to (l) (read in the right axis). The all data plotted here is based on the microdata of "the Survey of Research and Development, Ministry of Internal Affairs and Communications".

Information Engineering, and N: Electrical Engineering in 4) Tech university.

The difference between the average of richness of individual universities and richness as the whole university (as shown in Figure 4.1) can be understood by the distribution of richness for each university. As shown in Figure 4.5, the richness of research topics for individual universities indicates skewed distribution. The average richness of 69 national universities is around 400, while overall richness is around 2,500 (Figure 4.1). In other words, even for large research universities, it is difficult to follow even half of the research subjects covered by whole universities. Since there are many themes that are being studied in multiple universities, the overall richness is significantly less than the simple sum of the individual richness. Figure 4.5 also shows both the variance within each classification and the variance among each classification. The variance within 3) Middle classification is considerably gentler than that within 1) Research university classification.

Figure 4.6 shows the inputs and outputs for DEA averaged in each category of the national university. The total publication count (pub_{all}) implies skewed distribution among universities as compared to the distribution of researchers and the R&D budget. For example, the universities in 1) Research category hold approximately twice larger number of principal investigators and publish three times larger than ones in 2) Large category do. This is well consistent with the previous study on the skewed distribution of publication rather than one of the research grants (Shibayama (2011)). The ratio of publication with top 10% citation to total publication (pub_{t10}) is larger than 10% only for the university in 1) Research and 2) Large categories. This implies that there some room for improvement of citation for many national universities. The richness and evenness show opposite relationship, *i.e.* the universities with the higher richness show the lower evenness.

The growth of publication ($\Delta pub_{\mathrm{all}}$) rapidly decreases between 2002 and 2006. This trend appeared for universities in all six categories. The growth rate has been recovered but it is still under 5% level in 2011. The richness slightly increase for almost all universities while evenness shows fluctuation within ± 5. Fundamental statistics for both inputs and outputs variables are shown in Table 4.1 and 4.2, respectively.

The difference of outputs among universities may derive from the difference of inputs and/or their productivity. Data envelopment analysis computes productivity of each national university in each year by means of the weighted ratio between multiple inputs and multiple outputs and assumes that the efficiency of the most productive university (to the given inputs) in the dataset is 1.

(2002-2011)

Classification	University Research	Large	Middle	Tech	Non-med	Social
A Science and Technology in General	0.531	0.313	0.216	0.175	0.222	0.253
B Physics	1.024	0.572	0.326	0.390	0.343	0.188
C Chemistry	1.010	0.535	0.375	0.412	0.321	0.205
D Space Science and Earth Science	0.886	0.416	0.245	0.285	0.262	0.185
E Biological sciences	1.349	0.899	0.631	0.524	0.415	0.105
F Agriculture, forestry and fisheries	0.685	0.453	0.414	0.454	0.354	0.086
G Medicine	0.893	0.711	0.593	0.229	0.179	0.144
H Engineering in general	0.572	0.254	0.266	0.310	0.319	0.212
I System and Control Engineering	0.876	0.559	0.273	0.411	0.314	0.390
J Information Engineering	0.860	0.523	0.351	0.480	0.315	0.299
K Management Engineering	0.288	0.262	0.081	0.146	0.112	0.000
L Energy Engineering	0.515	0.486	0.297	0.211	0.199	
M Nuclear Engineering	0.479	0.386	0.225	0.090	0.157	0.693
N Electrical Engineering	0.822	0.499	0.341	0.445	0.380	0.264
P Thermal Engineering, Applied thermodynamics	0.485	0.245	0.273	0.313	0.209	0.000
Q Mechanical Engineering	0.353	0.198	0.173	0.212	0.240	0.077
R Construction Engineering	0.421	0.245	0.185	0.178	0.225	0.087
S Environmental Engineering	0.691	0.471	0.345	0.375	0.315	0.099
T Transport and Traffic Engineering	0.248	0.147	0.116	0.147	0.288	
U Mining Engineering	0.228	0.113	0.122	0.050	0.133	0.000
W Metallurgical Engineering	0.803	0.435	0.379	0.408	0.385	0.347
X Chemical Engineering	0.658	0.367	0.302	0.324	0.282	0.000
Y Chemical Industry	0.685	0.418	0.353	0.432	0.373	0.176
Z Miscellaneous Industries	0.147	0.249	0.096	0.162	0.323	

Log n_i 0.000 – 1.349

Figure 4.4: Heatmap of the number of publication in JST classification. The number of papers (in log-scale) for each classification codes in the most coarse class in JST classification system. The number of papers is the count based on the papers published between 2002 and 2011 and registered in both J-Global and Scopus databases commonly. The university in the category 1) Research shows many papers in all fields, notably B: Physics, C: Chemistry, and E: Biological sciences classification reflecting the characteristics of each field. The university in 2) Large and 3) Middle categories have a similar profile to 1) Research university, while 4) Tech university has obviously different from those distributions.

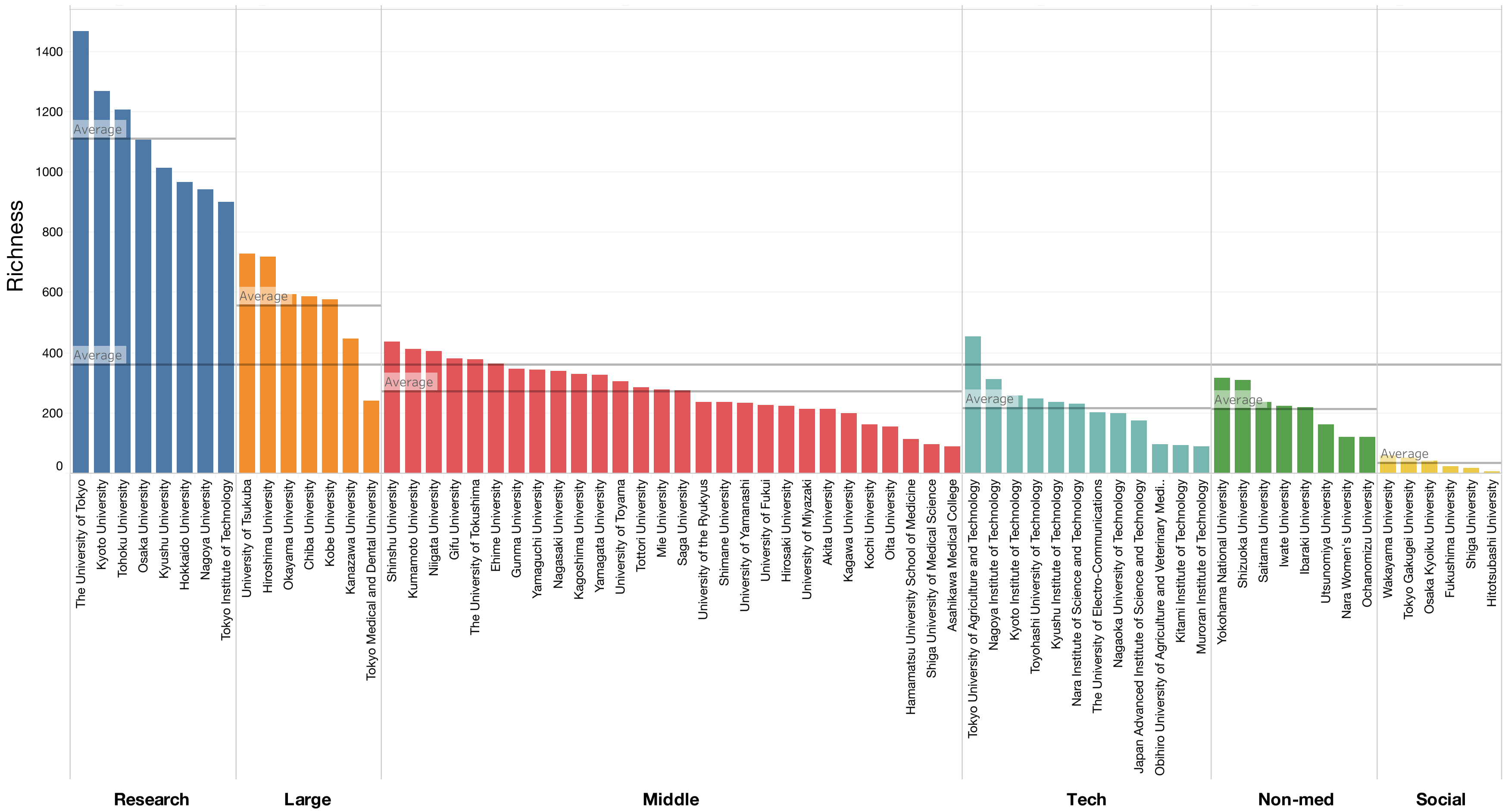


Figure 4.5: α-diversity and β-diversity in Japanese national universities.
The number of research subjects, richness, for 69 Japanese national universities in descending order of richness within each category of university. The average richess for all 69 universities and universities within each categories are shown in the solid horizontal lines, respectively. The richness index plotted here is the count based on the papers published between 2002 and 2011 and registered in both J-Global and Scopus databases commonly.

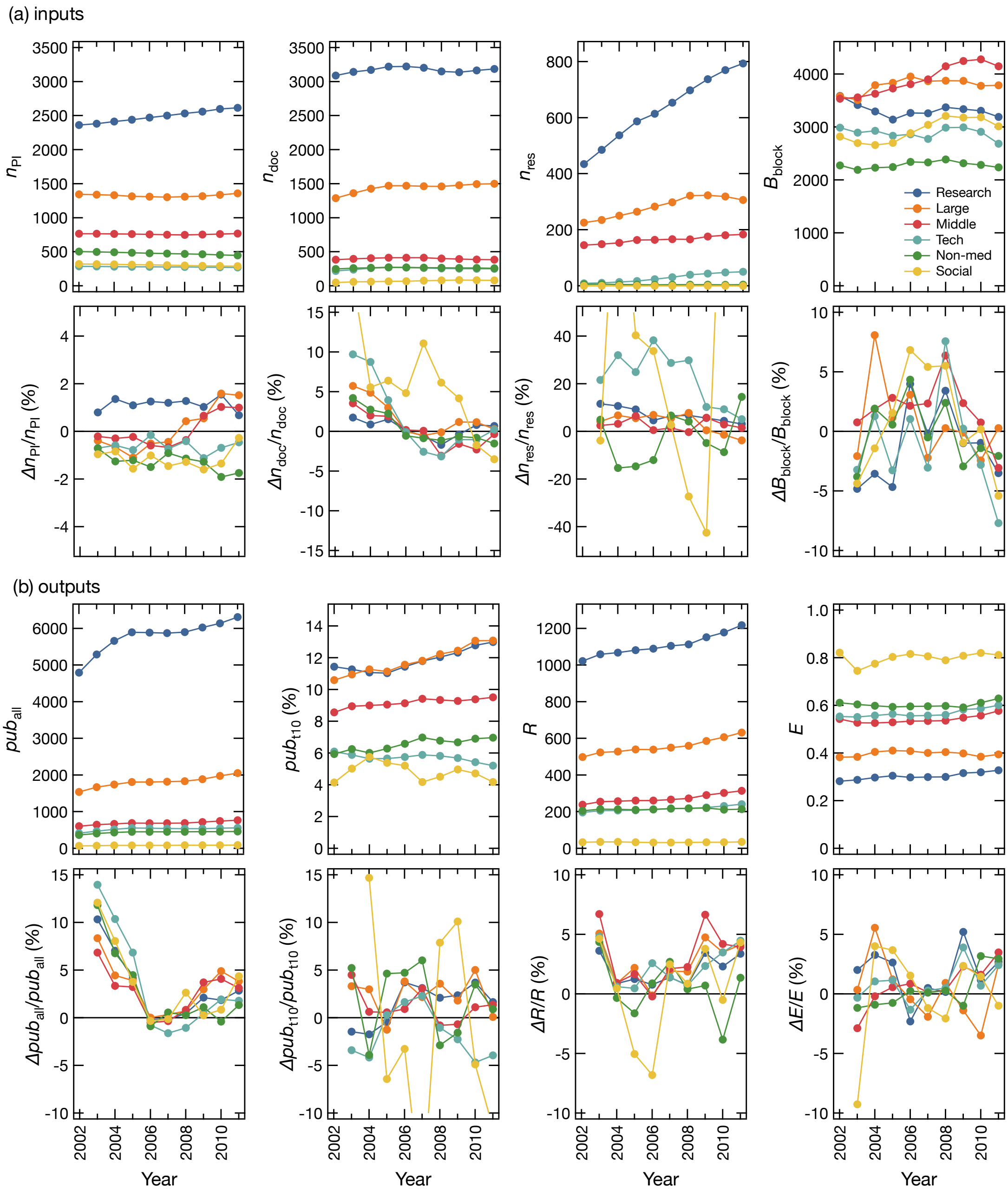


Figure 4.6: Time-change of inputs and outputs.
(a) The time-change of n_{PI}, n_{doc}, n_{res}, and B_{block} for 6 categories of national university. The growth ratio of each variables are plotted in the lower panels. Those values are based on the microdata of "the Survey of Research and Development, Ministry of Internal Affairs and Communications". (b) The time-change of pub_{all}, pub_{t10}, R and E for 6 categories of national university. The growth ratio of each variables are plotted in the lower panels. Those output variables are the count based on the papers registered in both J-Global and Scopus databases commonly.

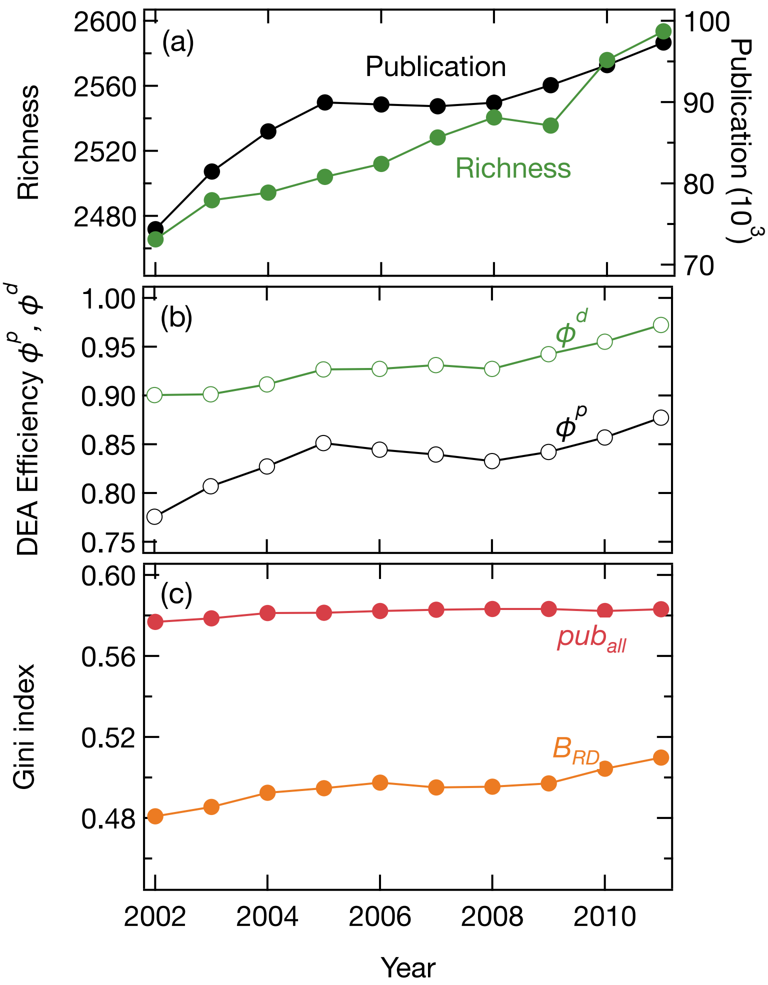


Figure 4.7: Time-change of outputs, DEA efficiencies, and Gini iindex of inputs for the whole 69 universities.

(a) The richness (R_{total}) and the number of paper (the sum of pub_{all}) for 69 national universities. Notice that R_{total} is not the sum of R for each universitiy. While publication count shows more than 30% growth during this decade, the richness slightly increases. (b) DEA efficiencies for publication ϕ^p and scientodiversity ϕ^d averaged over 69 universities. Both show improvement as a whole corresponding to the growth of publication and richness. (c) Gini index for pub_{all} and B_{RD} calculated from the Lorenz curve as shown in Figure 4.2. The input data used in this plot is based on the microdata of "the Survey of Research and Development, Ministry of Internal Affairs and Communications". The output data used here is based on the papers registered in both J-Global and Scopus databases commonly.

Table 4.1: Fundamental statistics of inputs.

		n_{PI}	n_{doc}	n_{res}	$B_{\mathrm{RD}}^{\mathrm{goods}}$	observation
1) Research	mean	2358	2985	590.9	2982446	80
	S.E.	(92.79)	(134.7)	(34.57)	(186319)	
	median	2240	2574	538.8	2810207	
2) Large	mean	1245	1366	300.2	945923	70
	S.E.	(36.17)	(47.41)	(12.89)	(62724)	
	median	1297	1152	280.5	759075	
3) Middle	mean	696.5	330.7	157.4	318280	280
	S.E.	(11.6)	(12.42)	(3.653)	(8279)	
	median	704.8	260.5	144.8	304957	
4) Tech	mean	244	205.6	24.79	262050	120
	S.E.	(8.61)	(12.14)	(2.728)	(12987)	
	median	206.2	180.8	14.83	222426	
5) Non-med	mean	425	263.3	4.608	228692	80
	S.E.	(18.08)	(16.5)	(0.626)	(16277)	
	median	427.3	214	2.167	164048	
6) Social	mean	276.9	147.4	4.6	99797	60
	S.E.	(5.459)	(33.44)	(1.413)	(9695)	
	median	272.8	29.33	0	68212	
all	mean	798.2	698	168.1	651678	690
	S.E.	(27.15)	(38.24)	(8.111)	(40265)	
	median	630.5	272	117.7	306879	

Table 4.2: Fundamental statistics of outputs.

		pub_{all}	pub_{t10}	R	E	observation
1) Research	mean	5777	0.1182	1108	0.3033	80
	S.E.	(254.4)	(0.0019)	(21.42)	(0.005)	
	median	4886	0.1151	1059	0.3003	
2) Large	mean	1814	0.1182	556.3	0.3972	70
	S.E.	(62.2)	(0.0025)	(19.36)	(0.0063)	
	median	1823	0.1126	580.2	0.401	
3) Middle	mean	690.7	0.0916	271.7	0.5414	280
	S.E.	(15.96)	(0.0013)	(5.984)	(0.0042)	
	median	659.3	0.0867	269.3	0.5451	
4) Tech	mean	524.1	0.0571	215.9	0.5671	120
	S.E.	(23.53)	(0.0024)	(9.205)	(0.0092)	
	median	521.2	0.051	218.5	0.5516	
5) Non-med	mean	438.8	0.0654	213.6	0.6029	80
	S.E.	(22.32)	(0.0023)	(8.331)	(0.0075)	
	median	392.8	0.0584	218.5	0.6046	
6) Social	mean	81.01	0.0481	33.46	0.7998	60
	S.E.	(4.775)	(0.0028)	(2.616)	(0.0124)	
	median	71.83	0.0466	31.5	0.8174	
all	mean	1283	0.0846	360.4	0.5332	690
	S.E.	(71.13)	(0.0012)	(12.09)	(0.0055)	
	median	663.3	0.0819	254.5	0.5331	

4.3 Publication and diversification efficiencies

The result of DEA for publication and diversification are shown in Table 4.3. Both efficiencies ϕ^p and ϕ^d are average within each university categories. Among 1) Research, 2) Large, and 3) Middle categories, the universities with a larger size in terms of both budgets and researchers indicate better efficiencies, while the universities in the category 4) Tech indicate the second-best efficiencies both in terms of publication and diversification. Those three categories, almost all universities are estimated as the decreasing return to scale in publication efficiency analysis. However, 30 % of the universities in the category 4) Tech are estimated as increasing return to the scale with high efficiency ($\phi^p = 0.918$). Thus, enlarging the size of those universities may be efficient policy to increase total publication. Those 4) Tech universities have relatively large pub_{t10} slacks, *i.e.* approximately 5% can be improved without increase of any inputs.

For the efficiency analysis for diversification, 18.3 % of the universities in the category 4) Tech are estimated as increasing return to the scale with high efficiency ($\phi^d = 0.932$), while all universities in the categories 1) Research, 2) Large, and 3) Middle are estimated as the decreasing return to scale for diversification. There are approximately 3.5% of E slacks for the categories 1) Research and 4) Tech universities although almost no slacks improvements are expected for the richness R. The scale expansion, in particular, of the specific universities, may not be a good policy for improving diversity as comparing with that for increase the number of papers. More than 35% of R slacks improvement is expected for the universities in the categories 6) Social, while they indicate high-efficiency $\phi^d = 0.920$ perhaps due to the good performance of E.

The standard deviation of the publication efficiency ϕ^p for 3) Middle university is slightly larger than that for other university categories despite the relatively small average efficiency $\phi^p = 0.781$. The standard deviation of the diversification efficiency ϕ^d for 2) Large classification is larger than those of other university classes. Those observed variations in the efficiency score suggest the inter-university diversity of their activity, *i.e.* some universities focused on research while others emphasize education. The difference of the research portfolio of each university may also affect the deviation of efficiency.

The time change of efficiencies ϕ^p and ϕ^d are shown in Figure 4.8. For publication efficiency, the observed improvement of ϕ^p for the universities in the categories 1) Research and 4) Tech indicates a shift of the best-practice frontier, while the universities in the categories 2) Large and 5) Non-med show catch-up in this decade. The universities in the category 3) Middle seems struggling to catch-up. For the diversification efficiency, the catching-up behavior is also observed for the category 2) Large and 3) Middle.

Table 4.3: Summary statistics for the efficiency scores.

(a) Publication

	ϕ^p				return to scale			slacks	
	mean	S.D.	min.	max.	decreasing	constant	increasing	pub_{all}	pub_{t10}
1) Research	0.939	(0.074)	0.754	1.000	97.5%	2.5%	0.0%	0.00%	0.78%
2) Large	0.830	(0.098)	0.635	1.000	100.0%	0.0%	0.0%	0.00%	0.00%
3) Middle	0.781	(0.120)	0.516	1.000	96.8%	2.5%	0.7%	0.00%	0.07%
4) Tech	0.918	(0.115)	0.504	1.000	43.3%	26.7%	30.0%	0.03%	4.97%
5) Non-med	0.850	(0.126)	0.483	1.000	75.0%	13.8%	11.3%	0.05%	0.64%
6) Social	0.777	(0.250)	0.200	1.000	70.0%	13.3%	16.7%	4.46%	9.27%
all	0.835	(0.145)	0.200	1.000	83.0%	8.7%	8.3%	0.40%	1.86%

(b) Diversification

	ϕ^d				return to scale			slacks	
	mean	S.D.	min.	max.	decreasing	constant	increasing	R	E
1) Research	0.937	(0.047)	0.839	1.000	100.0%	0.0%	0.0%	0.00%	3.60%
2) Large	0.812	(0.141)	0.486	1.000	100.0%	0.0%	0.0%	0.00%	0.00%
3) Middle	0.820	(0.083)	0.575	1.000	100.0%	0.0%	0.0%	0.00%	0.42%
4) Tech	0.932	(0.071)	0.712	1.000	62.5%	19.2%	18.3%	0.03%	3.51%
5) Non-med	0.914	(0.058)	0.803	1.000	81.3%	11.3%	7.5%	0.00%	0.43%
6) Social	0.920	(0.101)	0.641	1.000	51.7%	25.0%	23.3%	36.01%	0.06%
all	0.872	(0.101)	0.486	1.000	87.1%	6.8%	6.1%	3.14%	1.25%

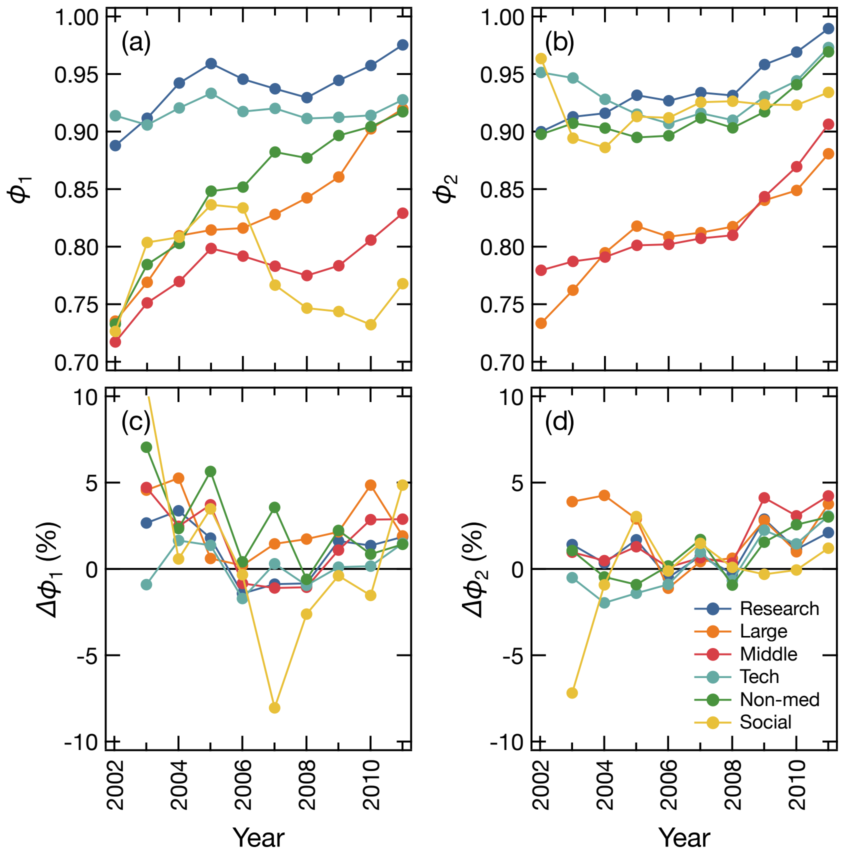


Figure 4.8: Efficiency scores for each class of university.
The time change of DEA efficiency of publication (ϕ^p shown in (a)) and that of sciento-diversity (ϕ^d shown in (b)) for 6 categories of university. The university in 1) Research and 4) Tech classification keep their high efficiency, while the university in 2) Large and 3) Middle show the catching-up behaviour for both publication and scientodiversity. The university in 5) Non-med classification indicates the high efficiency in scientodiversity while it shows the catching-up in publication. The growth ratio of ϕ^p and ϕ^d are shown in the lower panels (c) and (d), respectively.

4.4 Panel Tobit regression

The Tobit regression of efficiencies ϕ^p and ϕ^d indicate statistical significance on several variables both in linear and quadratic models as shown in Table 4.4. The publication efficiency ϕ^p depends on the total R&D expenditure B_{RD} and amount of block grant per single PI B_{block} within a concave-up relationship, *i.e.* coefficients for the quadratic term is positive for both variables. This implies that the publication efficiency can be improved by increase of B_{RD} and/or B_{block} and the larger amount of the increase of budget can lead the better improvement of the efficiency. This dependency on the size of the budget may be associated with the observed dependence of efficiency on the size of universities, *i.e.* the size of the total expenditure of universities is in the order of categories 1) Research, 2) Large, and 3) Middle as shown in Figure 4.3. The observed good efficiency ($\phi^p > 0.9$) of the universities in the category 4) Tech and 5) Non-med may be explained the relatively large amount of B_{RD} as shown in Figure 4.3.

According to the observed dependency of ϕ^p on the proportion of the external grant r_{ex} to the total R&D expenditure and the ratio of external grant from private company r_{com}, changing balance of budget is one possible strategy to improve the publication efficiency without an increase of the total amount of R&D budget. The publication efficiency ϕ^p indicates quadratic dependency both on r_{ex} and r_{com} with positive coefficients. The local minimum of the estimated latent variable ϕ^{p*} is realized for $r^*_{\mathrm{ex}} \simeq 0.225$ and $r^*_{\mathrm{com}} \simeq 0.427$ where $|\partial\phi^{p*}/\partial r_{\mathrm{ex}}|_{r_{\mathrm{ex}}=r^*_{\mathrm{ex}}} = 0$ and $|\partial\phi^{p*}/\partial r_{\mathrm{com}}|_{r_{\mathrm{com}}=r^*_{\mathrm{com}}} = 0$. Then, the increase of r_{ex} may improve publication efficiency for most of universities. In contrast, the increase of r_{com} may decrease the efficiency since r_{com} is smaller than r^*_{com} for most of universities as shown in Figure 4.3.

The diversification efficiency ϕ^d depends on B_{RD}, B_{block} and r_{RD} within the linear model but this dependency has not seen in the quadratic one. This implies that the increase in the number of research topics, *i.e.* richness, is not straight-forwardly accomplished by the increase of paper even richness depends linearly on the number of paper.

The efficiency ϕ^d can be explained only by r_{com} and its quadratic term within the statistical significance of $p < 0.001$. The estimated negative coefficient for the quadratic term implies that ϕ^d shows local maximum at $r^*_{\mathrm{com}} \simeq 0.313$ where $|\partial\phi^{d*}/\partial r_{\mathrm{com}}|_{r_{\mathrm{com}}=r^*_{\mathrm{com}}} = 0$. Then, one possible policy to improve diversification efficiency ϕ^d is an optimization of the ratio r_{com}.

4.5 Conclusions and policy implications

The efficiency of Japanese national universities in terms of the quantity and diversity of their publication has been quantitatively assessed by using data envelopment analysis

Table 4.4: Tobit regression of efficiency scores.

	Publication		Diversification	
	Linear	Quadratic	Linear	Quadratic
$B_{\mathrm{RD}}(10^6)$	-0.007	-0.059**	0.015***	0.011
	(0.006)	(0.019)	(0.003)	(0.011)
$B^2_{\mathrm{RD}}(10^{12})$		0.005***		0.000
		(0.001)		(0.001)
$B_{\mathrm{block}}(10^3)$	0.018*	-0.227***	-0.027***	0.006
	(0.007)	(0.028)	(0.004)	(0.017)
$B^2_{\mathrm{block}}(10^6)$		0.02393***		-0.00297
		(0.003)		(0.002)
r_{ex}	0.253**	-0.696*	-0.040	-0.119
	(0.093)	(0.271)	(0.053)	(0.168)
r^2_{ex}		1.548**		0.262
		(0.567)		(0.339)
r_{RD}	0.022	-0.673*	-0.215***	-0.207
	(0.070)	(0.304)	(0.040)	(0.195)
r^2_{RD}		0.356		0.024
		(0.234)		(0.151)
r_{com}	-0.160*	-0.751***	0.016	0.433***
	(0.068)	(0.177)	(0.039)	(0.112)
r^2_{com}		0.880**		-0.691***
		(0.280)		(0.177)
(Intercept 1)	0.666***	1.756***	0.933***	0.824***
(Intercept 2)	-1.942***	-2.076***	-2.488***	-2.506***
Log-likelihood:	193.3	281.4	570.2	579.5
AIC	-342.6	-508.8	-1096.3	-1104.9

***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, standard errors in parentheses.

(DEA).

First, we confirmed structural changes of university expenditure between 2001 and 2012. The universities in the category 1) Research successfully increase their R&D expenditure by obtaining external grants mainly from the government, while the other universities decrease their R&D expenditure even they increase dependency on external grants from the government. The government thought that the reduction of subsidies for operating expense since the incorporation of the National University in 2002 was able to be compensated by expanding competitive research grants. However, from our survey, the compensation was effective only for a few national universities, and many other universities have reduced their R&D expenditure over a decade. As a result, this policy made the budget distribution of national universities more skewed, *i.e.* a few rich universities get richer. From the viewpoint of paper production, the concentration of budget to some highly productive universities is not necessarily unfair. If the production function is assumed to be increasing returns to scale for all universities, the number of paper will take the maximum when all budget is given to one university with the highest productivity. However, in actual situations production function is assumed to be the decreasing return to scale in actual situations, so the government needs to allocate resources to universities with relatively low paper productivity. It is also necessary to keep in mind that paper production is not the only function of national universities.

Second, we evaluate efficiencies of national universities in terms of both publication and diversification by DEA. Among 1) Research, 2) Large, and 3) Middle categories, the universities with larger expenditures have better efficiencies. The production functions of those universities are estimated to be decrease returns to scale as expected for both publication and diversification. The good efficiency scores for universities in the category 4) Tech suggests that the management of university hospitals may have an important role in improving the efficiency of publications. The observed catch-up behavior, i.e. improvement of efficiency scores, of the universities in the categories 2) Large, 3) Middle and 5) Non-med is explained by both decreases of inputs and increase of outputs as shown in Figure 4.6. Although it was not mentioned in this research, it has been reported that the decrease in research time deteriorated the publication performance of Japanese researchers. Improvement in efficiency of these universities is the key to improving the performance of papers production in Japan as a whole, and the government should make a profound support to those universities including human resource aspects other than budget. The diversification efficiency shows similar trends with that of publication efficiency, *i.e.* $\phi^p_{\text{Research}} > \phi^p_{\text{Tech}} > \phi^p_{\text{Non-med}} > \phi^p_{\text{Large}} > \phi^p_{\text{Middle}}$. However, there are smaller room for improvement of efficiency in terms of both scale economy and slacks improvement as comparing with that of publication efficiency.

Finally, we test the impact of the balance of research grants on efficiency scores by

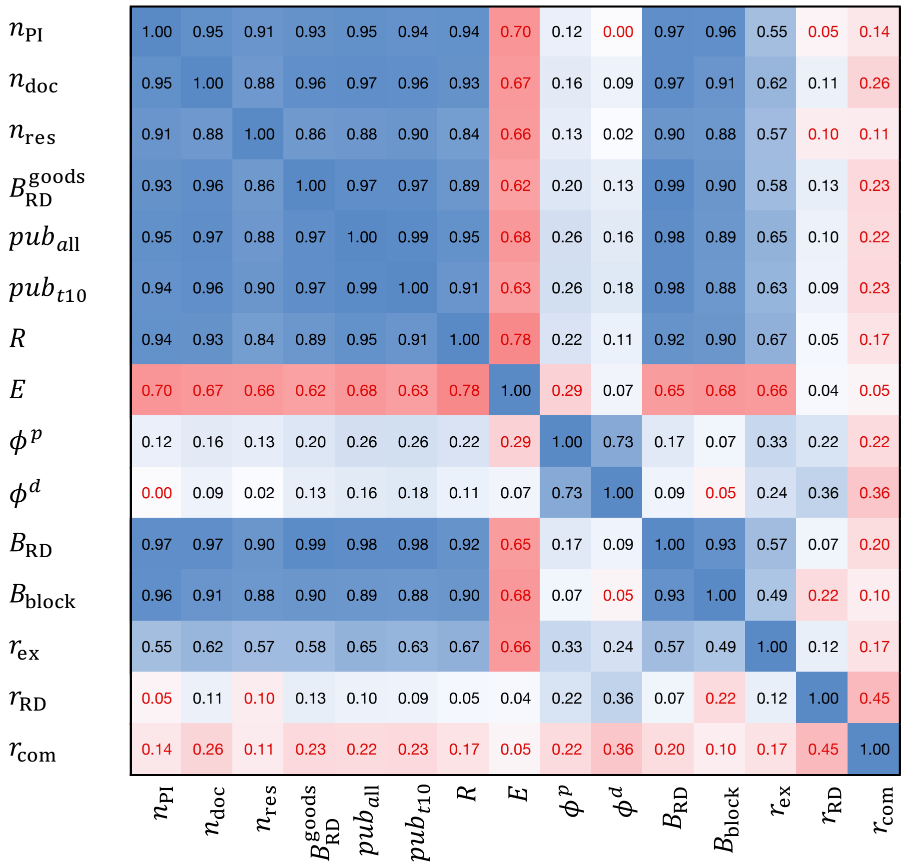


Figure 4.9: Correlation matrix of variables used in DEA and Tobit regression. The correlation coefficient for the variables used in DEA and Tobit regression in this chapter. The color indicates the sign of coefficient (i.e. blue represents positive and red does negative) and the shade of color represents the magnitude of correlation. The input data used in this plot is based on the microdata of "the Survey of Research and Development, Ministry of Internal Affairs and Communications". The output data used here is based on the papers registered in both J-Global and Scopus databases commonly.

Tobit regression. The result implies the importance of total R&D budget and block grant per researchers to publication efficiency. As well as the amount of budget, the structure of budget also affects efficiency scores both in publication and diversification. The increase of the proportion of the external research grant r_{ex} may improve the publication efficiency for most of the national universities, but not for diversification ones (within certain statistical significance). In contrast, the increase of r_{com} may decline the publication efficiency scores but improve diversification ones because the estimated sign of quadratic coefficient on r_{com}^2 is opposite between them as shown in Table 4.4.

Our results contribute to not only the foundation of a strategy for individual national universities but also provide useful suggestions for nation-wide resource allocation.

Chapter 5

Team-scale analysis

5.1 Japanese grant systems

The evaluation of research grant effectiveness has recently attracted considerable attention in science policy discussion. In particular, funding agencies are being held responsible for ensuring the effectiveness of and efficient investment in basic science (Abbott (2016)). The positive impact of research grants on the performance of researchers have been investigated (Jacob and Lefgren (2011)), but the manner in which diversity changes as a result of grant distribution has not been examined. Intuitively, mission-oriented grants for research in relatively narrow research subject areas should skew the distribution of publication research subjects, *i.e.* they should decrease the diversity of science. However, this expected negative impact on diversity cannot be confirmed by the study of isolated variables. Rather, it should be examined by means of comparison of the impact of mission-oriented and conventional curiosity-driven grants.

The curiosity-driven grant is widely believed to give researchers the freedom to choose research topics freely, thus easing concentration on specific research areas, *i.e.* it should increase the diversity of research. However, in a research environment with an excessive focus on performance, the curiosity-driven grant may push research towards rather conservative, safe topics more than the mission-oriented grant does, and thus may reduce the diversity of science. It is also reported that the distribution of the curiosity-driven grant in Japan is more skewed both in terms of researchers and universities than that of the publication performance of researchers in many disciplines (Shibayama (2011)).

Performance-based funding, whether curiosity-driven or mission-oriented, tend to be allocated to certain researchers and/or universities due to the skewed distribution of their research performance, and their high selection rate is regarded as a possible cause of reduction of a variety of research subjects (Adams and Smith (2003)). However, the above is still only working hypotheses that have never been tested quantitatively. The relationship between resource allocation type and research outcome is a pressing question that must be answered in the course of evaluation of policy effect, within the realm of public policy studies.

Competitive research grants are now recognized in many countries as a popular policy tool for the promotion of fundamental research and development. The competitive research grant is believed to (a) encourage researchers to be productive and (b) stimulate entrepreneurship among researchers, although competitive grants require that researchers spend more time on proposal applications and research administration, and less on research (Stephan (1996)). Japan has recognized that the competitive research grant system might enhance the performance of Japanese researchers in the current competitive research environment, and as a result, since the establishment of Japan's Science and Technology Basic Plan in 1996, Japan has made great efforts to establish a so-called dual-support system, which utilizes both block grants and competitive grants in some form of balance.

The performance-based approach became dominant in response to socioeconomic demand, *i.e.* the demand for scientific outcomes to be translated into economic growth in the form of new products and processes. The recent history of block grants to national universities and competitive grants in Japan is shown in Figure 5.1(a). It is clear that the budget for competitive grants has increased, whereas that for block grants to national universities has decreased by approximately 1% per year since the 2004 reorganization of national universities as independent administrative institutions. This cutback of block grants in Japan is directly related to the reduction of personnel expenses in most national universities. The impact of that reduction on the performance of Japanese scientists has not been examined, but obviously a negative effect is expected.

Figure 5.1 (b) shows the growth of two typical competitive grant programs for fundamental research in Japan, namely the Grant-in-Aid for Scientific Research (KAKENHI) and the Strategic Basic Research Programs. The two types of grants, widely recognized as curiosity-driven and mission-oriented grants, respectively, are managed by different funding agencies. The societal expectations of science have grown considerably, but in many countries, substantial growth of public research investment can no longer be expected. The application of evidence-based design and the implementation of effective and efficient research funding systems are recognized worldwide as policy demands.

5.2 Datasets

This study used data from several sources. The information of the principal investigators such as affiliation, position, the location of their affiliated institution was retrieved from

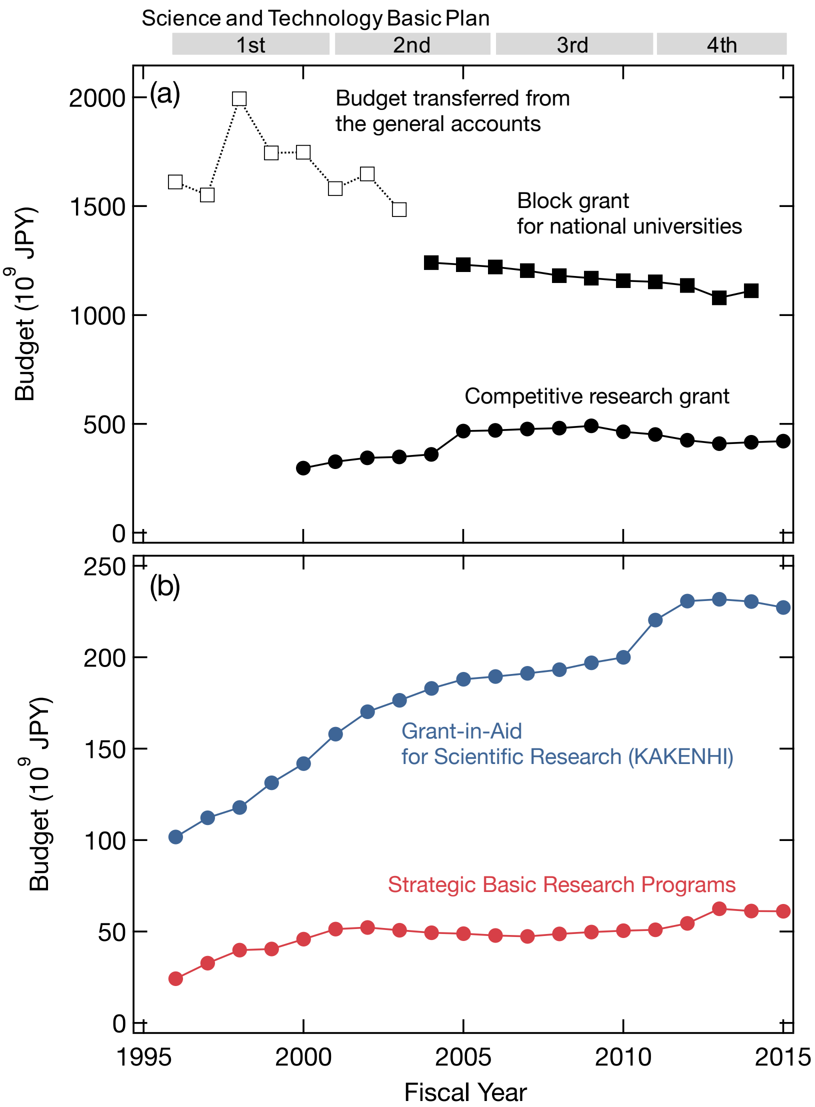


Figure 5.1: Science budget in Japan.
(a) The recent history of the block grants to national universities (open and solid square) and the competitive research grants (solid circle) after the establishment of the first Science and Technology Basic Plan in Japan. Notice that the gap observed in 2004 is caused by the change in the account system of Japanese national university corresponding to the reorganization of national universities as independent administrative institutions. (b) The growth of two typical competitive grant programs for fundamental research in Japan, namely the Grant-in-Aid for Scientific Research (KAKENHI) and the Strategic Basic Research Programs.

Table 5.1: Properties of Japanese grant programs for fundamental research.

	CREST	KAKENHI		
		(SR-S)	(SR-A)	(SR-B)
Type of grant	Mission-oriented	Curiosity-driven		
Grant (JPY per proposal)	150-500M	50-200M	20-50M	5-20M
Term (years)	5.5	5	3-5	3-5
Adoption rate (FY2015)	9.6%	13.2%	23.1%	23.0%
Funding Agency	JST	JSPS		

the public databases, such as the JST Project Database and the Database of Grants-in-Aid for Scientific Research, utilized by the funding agencies. Those researchers were identified in the Scopus database by their name and affiliation. Then, we made the publication list of each researcher from the database and counted their publications and citations as well. The articles published by the identified researchers were also retrieved in the J-Global database in order to obtain the JST classification codes given to each paper (see also the next section).

In this study, we compared two grant programs, namely the Grants-in-Aid for Scientific Research (KAKENHI) and the CREST program, as the typical examples of curiosity-driven and the mission-oriented fundamental research grants, respectively. The research subjects awarded in the CREST program are regulated in terms of the strategic target issued by the government, especially the Ministry of Education, Culture, Sports, Science and Technology (MEXT). The amount of the grant for the CREST program is larger than that for the KAKENHI program, which has several types of grant depending on its proposed budget and research terms, as listed with other properties in Table 5.1.

We identified the Japanese principal investigators ($N_0 = 348$) who were awarded to either the KAKENHI program (type (S), (A), or (B)) or the CREST program. In order to avoid the fixed-effects on publication and citation preference, or culture of the specific research community, we retrieved the principal investigators only from the field of nanotechnology and materials science. Both KAKENHI and CREST programs have supported these research areas for several decades.

5.3 The impact of grants on publication performance

Fundamental statistics of variables are shown in Table 5.2. The Probit regression of probability of participation to the CREST program indicated statistical significance on several

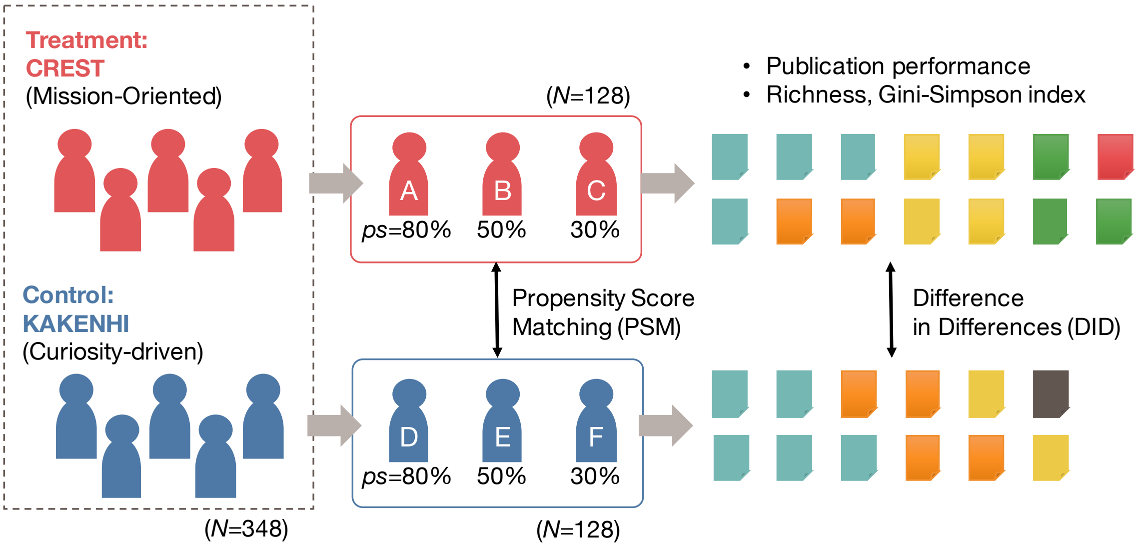


Figure 5.2: Outline of analysis.
First, we identify 348 Japanese principal investigators who were awarded to either the KAKENHI program (type (SR-S), (SR-A), or (SR-B)) or the CREST program. Then we set 128 pairs of PIs according to the propensity score (*i.e.* the probability of participation to the CREST program) calculated by Probit regression using the number of publication, citation and the affiliation information. Finally, we compare publication performance by the Difference in Differences (DID) methodology and also diversity indices (richness and Gini-Simpson).

dummy indices such as d_{UT}, d_{FIU}, d_{prof}, and d_{kanto} as well as the performance variables such as $pub^{\mathrm{ave}}_{\mathrm{before}}$ and $cite^{\mathrm{ave}}_{\mathrm{before}}$ as shown in Table 5.3. The significant impact of the affiliation and position reflected the skewed distribution of awardee of the grant. It should not be understood as the result of the biased selection process. The negative coefficient for d_{UT} can be understood as a compensation for excessive dependence on d_{FIU} and d_{kanto} since those dummy indices have value 1 simultaneously when $d_{\mathrm{UT}} = 1$, *i.e.* the University of Tokyo is one of the former Imperial Universities ($d_{\mathrm{FIU}} = 1$), and located within the Kanto region ($d_{\mathrm{kanto}} = 1$). The coefficient for $pub^{\mathrm{ave}}_{\mathrm{before}}$ indicated negative dependence on the publication although the peer-review selection process was thought to be performance-based.

This negative dependence on publication implies that evaluation measures other than the number of publications are considered in the selection process for the CREST program. On the other hand, the average citation count in the three years prior to the application to the grant program $cite^{\mathrm{ave}}_{\mathrm{before}}$ was statistically significant. The coefficients for other variables, such as d_{kansai}, $top10^{\mathrm{ave}}_{\mathrm{before}}$ and $cpp^{\mathrm{ave}}_{\mathrm{before}}$, were not statistically significant. The log-likelihood test after the propensity score matching rejected the hypothesis that for some variables there was a statistically significant difference between the control and the treatment group in both the caliper and the kernel matching method (Table 5.4). The statistical difference was relaxed, *i.e.* the p-value increased, for many variables such as d_{UT}, d_{FIU}, d_{prof}, and $cite^{\mathrm{ave}}_{\mathrm{before}}$ although the p-value for d_{kanto} and $pub^{\mathrm{ave}}_{\mathrm{before}}$ decreased slightly. The average treatment effect on the treated (ATT) of the participation to the CREST program was estimated by the difference in differences between control group and treatment group:

$$X^{s}_{\mathrm{b-g}} = X^{s}_{\mathrm{going}} - X^{s}_{\mathrm{before}},$$
$$X^{s}_{\mathrm{b-a}} = X^{s}_{\mathrm{after}} - X^{s}_{\mathrm{before}}$$

where X represents either pub, $cite$, $top10$, or cpp, and s stands for either ave or max. As shown in Table 5.5, a positive impact on the publication and citation performance was indicated. From the caliper matching samples, the ATT for publication count was statistical significant, indicating that participation in the CREST program was relevant to increase of publication by approximately one paper per year per single principal investigator during five years of participation, and increased by a further half publication count in three years after the end of the participation. This increase of the participation effect is consistent with the intuitive understanding of the time lag between the research activity and publication. The positive impact on citation count was estimated with relatively small statistical significance, *i.e.* the p values of ATT during participation were 0.108 and 0.145 for the caliper matching and the kernel matching samples, respectively. The result also

Table 5.2: Fundamental statistics.

variables	N	Mean	Std. Dev.	Min.	Max.
y_{start}	350	2005.020	3.085	1997	2010
y_{end}	350	2008.691	3.689	2000	2015
d_{UT}	350	0.171	0.377	0	1
d_{FIU}	350	0.471	0.500	0	1
d_{prof}	350	0.549	0.498	0	1
d_{kanto}	350	0.546	0.499	0	1
d_{kansai}	350	0.186	0.389	0	1
$treat$	350	0.554	0.498	0	1
$pub^{\mathrm{ave}}_{\mathrm{before}}$	348	6.646	6.709	0	45.333
$pub^{\mathrm{ave}}_{\mathrm{going}}$	350	7.081	6.950	0	44.800
$pub^{\mathrm{ave}}_{\mathrm{after}}$	350	6.439	6.557	0	40.000
$cite^{\mathrm{ave}}_{\mathrm{before}}$	348	236.091	369.728	0	2819.333
$cite^{\mathrm{ave}}_{\mathrm{going}}$	350	191.304	291.838	0	1838.400
$cite^{\mathrm{ave}}_{\mathrm{after}}$	350	90.934	190.828	0	1878.667
$top10^{\mathrm{ave}}_{\mathrm{before}}$	348	30.284	34.616	0	381.667
$top10^{\mathrm{ave}}_{\mathrm{going}}$	350	23.665	48.357	0	828.400
$top10^{\mathrm{ave}}_{\mathrm{after}}$	350	9.874	15.094	0	132.404
$cpp^{\mathrm{ave}}_{\mathrm{before}}$	348	1.367	2.088	0	18.000
$cpp^{\mathrm{ave}}_{\mathrm{going}}$	350	1.700	2.628	0	20.400
$cpp^{\mathrm{ave}}_{\mathrm{after}}$	350	1.671	2.778	0	23.500
$cite^{\mathrm{max}}_{\mathrm{before}}$	350	364.731	588.363	0	5320.000
$cite^{\mathrm{max}}_{\mathrm{going}}$	350	356.389	599.496	0	7035.000
$cite^{\mathrm{max}}_{\mathrm{after}}$	350	146.217	317.701	0	3233.000

Table 5.3: Probit regression of probability of participation to the CREST program.

	Coef.	S.E.
d_{UT}	-0.893***	(0.301)
d_{FIU}	0.655***	(0.220)
d_{prof}	0.522***	(0.153)
d_{kanto}	0.809***	(0.241)
d_{kansai}	0.294	(0.215)
$pub^{\mathrm{ave}}_{\mathrm{before}}$	-0.0448**	(0.0192)
$cite^{\mathrm{ave}}_{\mathrm{before}}$	0.00131**	(0.000664)
$top10^{\mathrm{ave}}_{\mathrm{before}}$	0.0341	(0.0850)
$cpp^{\mathrm{ave}}_{\mathrm{before}}$	-0.00619	(0.00379)
N	348	
Pseudo R squared	0.0659	
log likelihood	-223.17685	

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$, standard errors in parentheses.

indicated that there was no significant effect on the citation after finishing the program. The maximum count of citation during participation period also showed a positive ATT in both the caliper and the kernel matching samples with statistical significance. Any other of ATT was statistically insignificant.

5.4 Overview of popular research topics

We examined the diversity of research subject in the articles published between 1996 and 2013 by the researchers in the control and treatment groups, the scores being obtained after propensity score matching (caliper matching). There were 128 researchers in each group, as listed in Table 5.4. The ten most frequently occurring Scopus subject categories in the papers published by both control group experimental group researchers are listed in Table 5.6, with proportion and ranking indicated. The seven most common categories for the treatment group were exactly same, and in the same order, as those for the control group. A high concentration of papers in the category Physics and Astronomy was observed among the control group papers; this implies a smaller diversity index for the control group papers. Indeed, the Gini-Simpson index $1-\lambda$ (calculated from the distribution of Scopus classification) for the control group papers was 0.62, smaller than that for the treatment group (0.80). Note that the numerical value of the Gini-Simpson index is strongly affected by classification scheme granularity. For example, if the papers in the most popular category, Physics and Astronomy, are divided into two smaller categories accounting for equal shares of the total, the diversity index for the control group papers improves to 0.78. Thus, classification definition can interfere with detailed analysis using the Gini-Simpson index. Moreover, the Gini-Simpson index is sensitive to concentration but not dispersion. In the case reported here, the indices calculated for only the top seven categories (0.58 and 0.76 for the control group and the treatment group papers, respectively) were still within 10% of the values calculated from the whole list for both groups. This reflects the fact that the cumulative proportion of the top seven categories is greater than 90 % for both groups.

The fine structure of distribution of research subject retrieved from the J-Global database is shown in Figure 5.3, where the size of the circle represents the number of articles, with specific category codes listed in the left. The number of papers is the sum of six-year periods, such as 1996 - 2001, 2002 - 2007, and 2008 - 2013. Here, we display only the 91 categories which occurred 100 times in total for both groups over the 18 year period. The cumulative proportion of those 91 categories was 68.8% and 52.4% for the control and the treatment group, respectively. The distribution pattern of categories for the treatment group papers was different for all terms from that for the control group.

Table 5.4: Balance test after the propensity score matching.

		Before matching	Caliper matching	Kernel matching
d_{UT}	Mean (treatment)	0.165	0.180	0.164
	Mean (control)	0.182	0.180	0.193
	p value	0.680	1.000	0.466
d_{FIU}	Mean (treatment)	0.500	0.477	0.503
	Mean (control)	0.429	0.445	0.507
	p value	0.186	0.618	0.937
d_{prof}	Mean (treatment)	0.613	0.500	0.608
	Mean (control)	0.461	0.516	0.588
	p value	0.004	0.803	0.684
d_{kanto}	Mean (treatment)	0.557	0.547	0.556
	Mean (control)	0.539	0.594	0.593
	p value	0.742	0.451	0.465
$pub_{\mathrm{before}}^{\mathrm{ave}}$	Mean (treatment)	6.851	7.051	6.583
	Mean (control)	6.386	6.324	6.249
	p value	0.522	0.393	0.607
$cite_{\mathrm{before}}^{\mathrm{ave}}$	Mean (treatment)	271.930	210.610	219.490
	Mean (control)	190.950	201.480	220.620
	p value	0.042	0.793	0.967
N(treatment)		194	128	189
N(control)		154	128	154
Pseudo R squared		0.056	0.009	0.002
LR test (p value)		0.000	0.877	0.990

Table 5.5: Average treatment effect of participation in the CREST program.

		Before matching	Caliper matching	Kernel matching
$pub^{\mathrm{ave}}_{\mathrm{b-g}}$	ATT	1.469***	1.048**	1.244**
	SE	(0.415)	(0.512)	(0.572)
	p value	0.000	0.040	0.030
$pub^{\mathrm{ave}}_{\mathrm{b-a}}$	ATT	1.776***	1.536**	1.3**
	SE	(0.574)	(0.621)	(0.648)
	p value	0.002	0.013	0.045
$cite^{\mathrm{ave}}_{\mathrm{b-g}}$	ATT	17.387	44.52	68.154
	SE	(27.475)	(27.695)	(46.755)
	p value	0.529	0.108	0.145
$cite^{\mathrm{ave}}_{\mathrm{b-a}}$	ATT	-43.148	9.163	13.636
	SE	(34.686)	(28.579)	(47.141)
	p value	0.216	0.749	0.772
$top10^{\mathrm{ave}}_{\mathrm{b-g}}$	ATT	0.34**	0.232	0.423*
	SE	(0.166)	(0.213)	(0.254)
	p value	0.041	0.276	0.096
$top10^{\mathrm{ave}}_{\mathrm{b-a}}$	ATT	0.169	0.253	0.147
	SE	(0.209)	(0.247)	(0.276)
	p value	0.418	0.306	0.595

(to be continued)

		Before matching	Caliper matching	Kernel matching
$cpp^{\text{ave}}_{\text{b-g}}$	ATT	5.138	9.477	9.761*
	SE	(5.71)	(7.549)	(5.814)
	p value	0.369	0.209	0.093
$cpp^{\text{ave}}_{\text{b-a}}$	ATT	-1.69	3.723	6.496
	SE	(3.448)	(2.775)	(4.441)
	p value	0.624	0.180	0.144
$cite^{\text{max}}_{\text{b-g}}$	ATT	58.887	116.672*	144.034**
	SE	(61.8)	(65.258)	(72.837)
	p value	0.343	0.074	0.048
$cite^{\text{max}}_{\text{b-a}}$	ATT	-50.216	30.703	71.184
	SE	(57.921)	(49.73)	(76.865)
	p value	0.385	0.537	0.354
N		348	282	343

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

Table 5.6: Popular research subjects retrieved from the Scopus database.

	Control group		Treatment group	
Scopus Subject Category	Rank	Proportion (%)	Rank	Proportion (%)
Physics and Astronomy	1	57.10	1	34.18
Materials Science	2	20.65	2	18.84
Chemistry	3	8.58	3	17.34
Engineering	4	5.50	4	10.86
Biochemistry, Genetics and Molecular Biology	5	1.55	5	4.56
Chemical Engineering	6	1.25	6	4.40
Medicine	7	1.00	7	1.49
(Cumulative proportion of top seven categories)		95.63		91.66
Mathematics	8	0.98	12	0.95
Energy	9	0.92	9	1.17
Multidisciplinary	10	0.84	15	0.43
(Cumulative proportion of top ten categories)		98.37		94.21

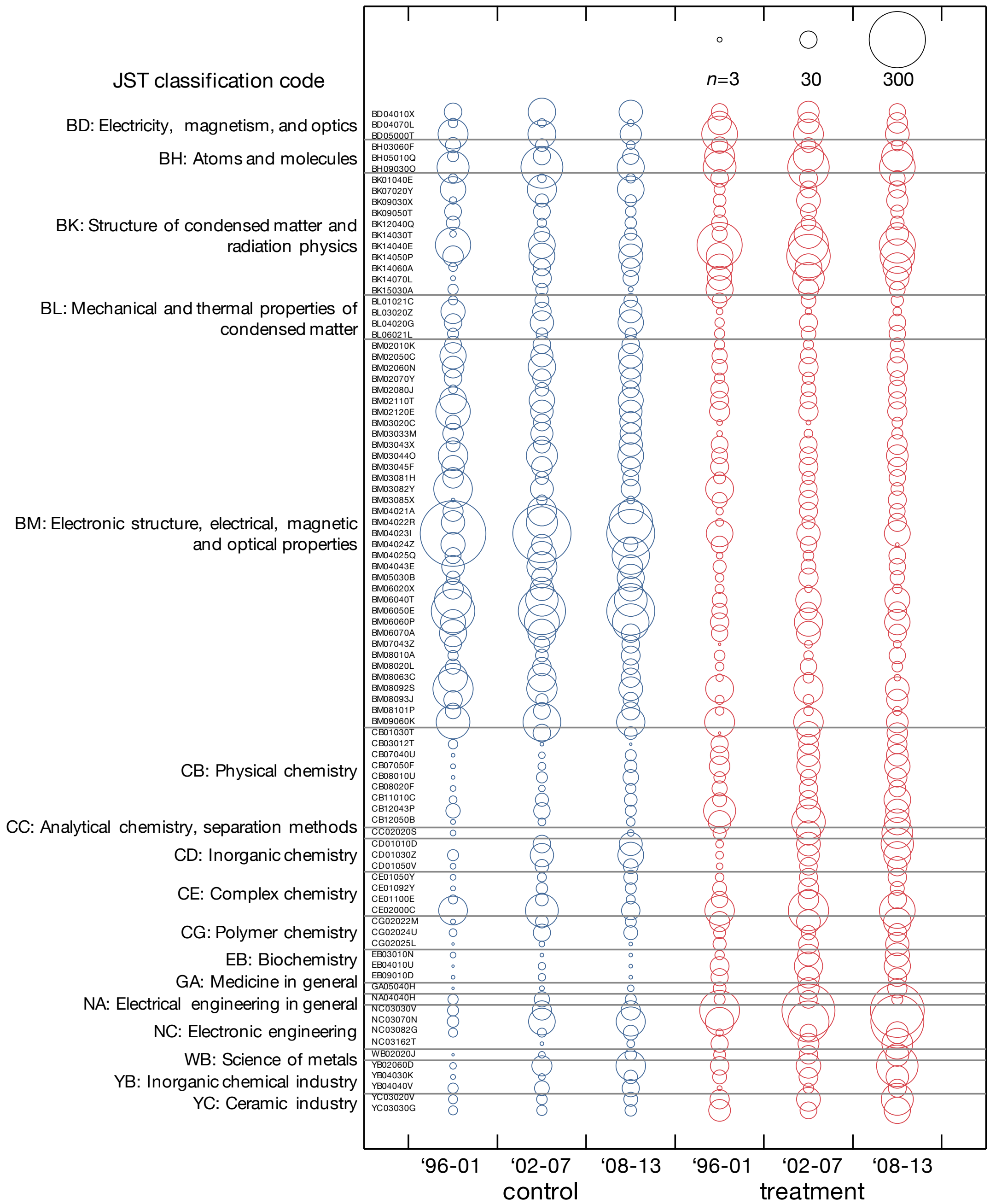


Figure 5.3: Distribution of popular research subjects in the J-Global database. The distribution of research subject for the control and treatment group retrieved from the J-Global database. The size of the open circle represents the number of articles with specific classification codes listed in the left. Those number of papers are the sum of six-year periods, 1996 - 2001, 2002 - 2007, and 2008 - 2013. We display here only the 91 classification codes attached more than 100 articles in total for both groups over the 18 year period.

The mid-level categories BK (structure of condensed matter and radiation physics) and NC (electrical engineering) were dominant among the treatment group papers, although for the control group a very large concentration was observed for category BM (electronic structure, electrical, magnetic and optical properties). The size of the above categories seems to be stable over time, which implies that the impact of the funding program on the distribution of popular research subjects is small. The observed positive impact of mission-oriented grants on publication count may reflect an increase in the number of studies on relatively minor research subjects. This difference in the distribution of research areas suggests some complementarity of role between the curiosity-driven KAKENHI and the mission-oriented CREST grant programs. The framework of a simple linear model, in which the curiosity-driven grant covers fundamental research and the mission-oriented grant covers research of a more applied nature, does not afford an understanding of the difference among research subject concentrations.

5.5 Time-series change in diversity of research subjects

In this study, we quantitatively compare the impact of mission-oriented research grants and curiosity-driven grants on the diversity of research subjects in Japan. More specifically, we identify groups of researchers whose publication performance is positively affected by mission-oriented grants (with statistical significance) and evaluate the diversity of research subjects through analysis of the distribution of the classification codes over groups of papers.

The cumulative distribution of JST classification code

$$C(k) = \sum_{n=k}^{\infty} p(n) \tag{5.1}$$

where $p(n)$ represents the proportion of classification code with abundance $n_i = n$, showed heavy-tailed distribution for both control and treatment groups in all time intervals (Figure 5.4). A few categories with large popularity occupied a relatively large share of the distribution as compared with that for normal distribution. Concentration in specific categories with large abundance, *e.g.* $k > 50$, was observed for the control group for all terms. Those distributions were well fitted by lognormal distribution function with two fitting parameters μ and σ. The lognormal distribution is widely observed in many fields of natural science and is understood as the result of a multiplicative process (Limpert et al. (2001)), and has been applied in studies of citation pattern (Albarrán et al. (2011); Radicchi et al. (2008)). The observed skewed distribution of research subjects implies that the recipients of both KAKENHI and CREST grants choose their research subjects freely. Their choice is perhaps influenced by surrounding conditions such as availability of resources, priori-

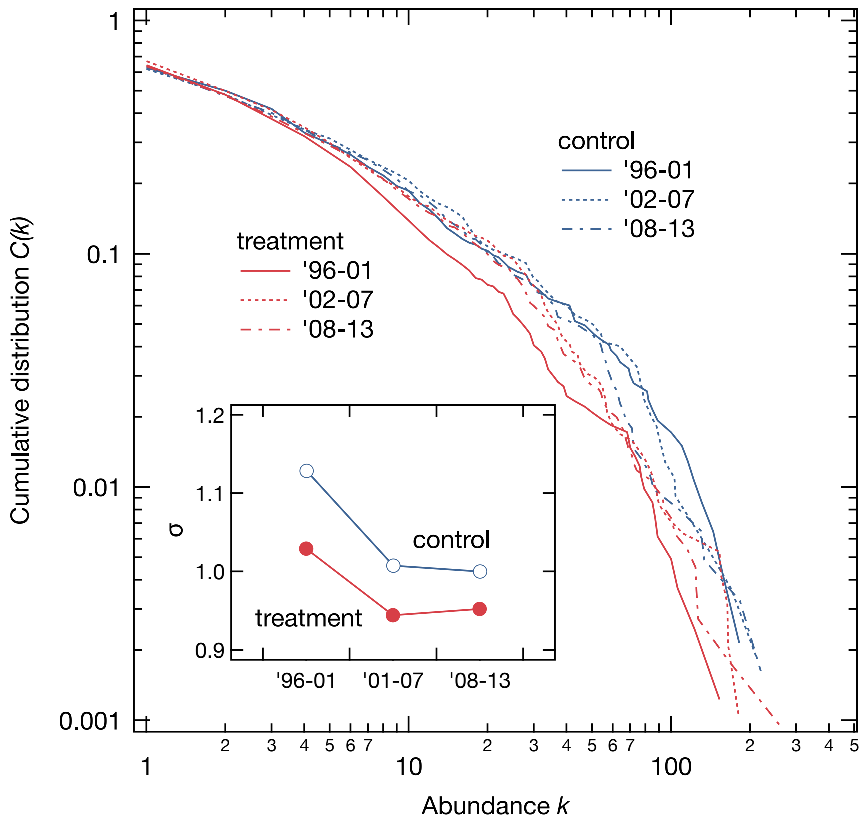


Figure 5.4: Cumulative distribution of research subjects as a function of abundance of category.
The distribution of the number of articles (*i.e.* abundance k) with each classification codes for the six-year periods, 1996 - 2001, 2002 - 2007, and 2008 - 2013 for the control and treatment groups. We display here only the 91 classification codes attached more than 100 articles in total for both groups over the 18 year period. The inset shows the fitting parameter σ of the lognormal function (see Chapter 2).

ties set in a research community, success in previous research, and government mission statement.

As shown in the inset in Figure 5.4, the parameter σ was always larger for the control group than the treatment group, and both decrease with time. This is interpreted as flat distribution, *i.e.* better diversity, of research subjects for the treatment group than for the control group, although the diversification appears to have occurred for both groups. The lognormal distribution robustly retained its shape over almost two decades of observations, for both mission-oriented and curiosity-driven research. The distribution for mission-oriented research changed slightly over time but remained well fitted to lognormal function, although the mission-oriented grant appears to have mitigated such skew distribution.

This implies that implemented investment in this area of science is less concentrated or less effective to vary the behavior of scientists. Diversification of science resulting from socioeconomic demand is rather quite limited here. Regardless of the type of research, the volume of research on most core subjects grew less concentrated, as can be seen in Figure 5.4. This does not necessarily imply a skew distribution of resource allocation: diversification of science may reflect the presence of studies in a wide range of research subjects but limited to a small number of papers for each subject. The results confirm that those relatively infrequently addressed research subjects are better served by mission-oriented grants than curiosity-driven ones.

Time-series change in richness R and Gini-Simpson index $1-\lambda$ are shown in Figure 5.6(a) and (b), respectively. It is clear that richness R, as indicated by circle makers was proportional to the total publication count for both sample groups, depicted by solid and dashed lines plotted using the right y-axis. Richness of the research subject for the treatment group was always larger than that for the control group, reflecting greater publication volume. This strong correlation between publication and richness cannot be taken to mean that diversity of subjects makes researchers more productive. There may be a third factor that correlates strongly with both publication and richness and thus gives rise to the apparent correlation between diversity and productivity. The most likely candidate is the availability of resources, which is allocated in somewhat skewed distribution as a result of funding agency policy.

Evenness of the treatment group samples is indicated by the Gini-Simpson index for all years in Figure 5.6(b). The improvement of control group diversity discussed in terms of the parameter σ in the inset in Figure 5.4 was confirmed by the increase of the Gini-Simpson index observed in around 2003. For the treatment group, in spite of a gradual increase of publication numbers around 2003, the Gini-Simpson index increased only slightly per year, in contrast with the observed minute change in publication and the relatively large increase in the index for the control group. This difference in diversification pattern, *i.e.* improved richness among the mission-oriented grant-funded studies and improved evenness among curiosity-driven grant funded-studies, is also suggestive of compartmentalization between the two types of grant.

5.6 Diversity difference between two types of grants

The linear dependency of richness R on publication number can be seen in Figure 5.7(a). The slope of the linear correlation for the treatment group sample was slightly steeper than that for the control group, indicated in solid and dashed lines, respectively, in Figure 5.7(a). However, this publication-richness linear relation can be a spurious correlation between time series data. As shown in Figure 5.7(b), there seems weak correlation (adjusted

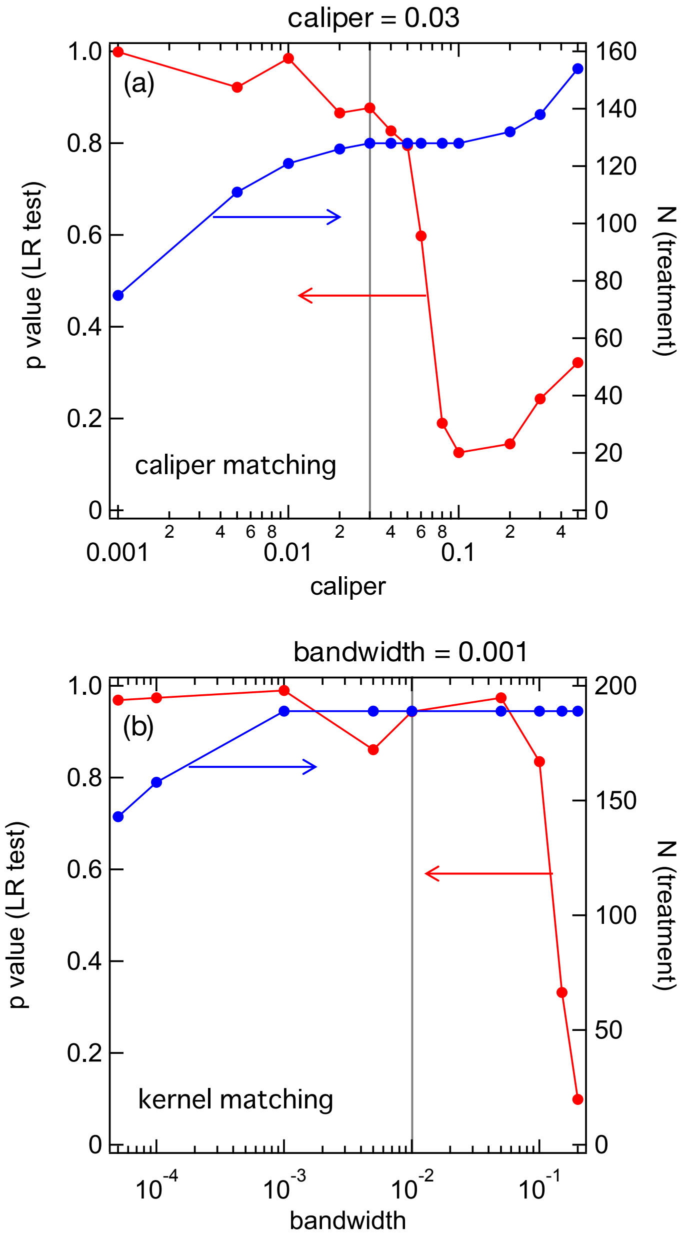


Figure 5.5: Parameter dependency of p-value of the likelihood-ratio test and number of PIs in the treatment group.
(a) The p-value of the LR test after caliper matching with respective calipers is shown by red marks and line (read in the left axis). The p-value drops with increase of the caliper size, while the number of PIs of the treatment group (shown by blue marks and line, read in the right axis), who are matched with PIs in the control group, merely change with the change of caliper between 0.01 and 0.1. (b) The p-value of the LR test after kernel matching with respective calipers is shown by red marks and line (read in the left axis). The p-value steeply drops at around the bandwidth 5×10^{-2}. The number of PIs of the treatment group keep its value with the bandwidth between 10^{-3} and 10^{-1}.

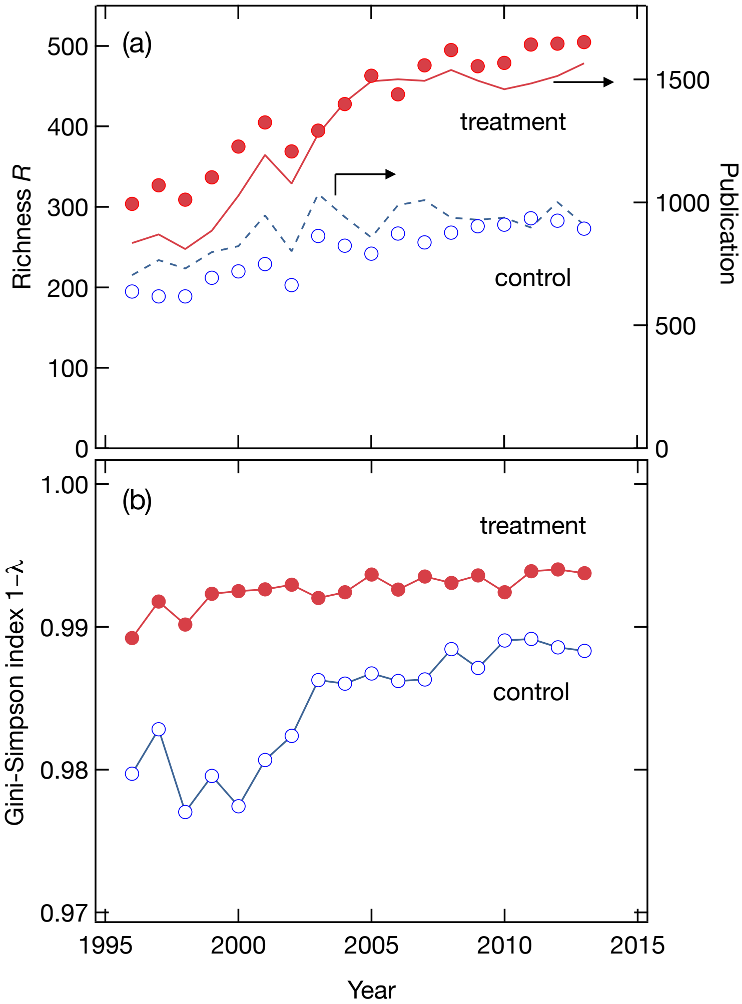


Figure 5.6: Time dependence of the richness and the Gini-Simpson index. Time-series change in richness R (a) and Gini-Simpson index $1-\lambda$ (b) between 1996 and 2013 for the treatment (shown in the red solid circles) and the control groups (shown in the blue open circles). The number of papers for the treatment and the control groups are plotted (read in the right axis))in the red solid and the blue dashed lines, respectively in the panel (a).

coefficient of determination R^2) between the first differences of richness and publication. The difference in the slope of the regression line between the control and the treatment groups (plotted in dashed and solid lines, respectively, in Figure 5.7(b)) is not statistically significant. Therefore, the treatment group (CREST) has greater richness than the control group (KEKENHI) perhaps because of the large number of papers of the treatment group, and density of research subject, *i.e.* richness divided by the number of articles, are almost same between them. This suggests that CREST program cannot be said to be a simple concentrated investment as contrary to the expectation that mission-oriented grant invests in specific research themes. Future studies will extend the analysis to a detailed examination of the publication list of each scientist; this is expected to reveal some changes in research theme in personal research histories.

The relation between richness and the Gini-Simpson index is shown in Figure 5.8. It is clear that the treatment group papers were more diverse than the control group papers in terms of both richness and evenness. The dashed line represents the maximum value of the Gini-Simpson index for each R. The relatively large deviation between the maximum (line) and the present value for the control group papers (open circles) means that there is considerable opportunity for improvement of diversity, even for the same richness. Notice that the observed low richness is not itself problematic, but maybe a matter of inefficient resource allocation, *i.e.* surplus and deficit are problematic from the viewpoint of public policy. Although the optimal allocation of resources over whole scientific subjects cannot be realized easily, public investment should follow a strategy based on the observation of distribution of research papers as well as measurement of diversity indices such as richness, evenness, and disparity. If publication count is regarded as a proxy measure of the amount of invested resources, our results suggest that the mission-oriented grant is the preferred tool for enhancing the diversity of science. However, our results also show complementarity of investment by the two types of grant over different research subjects, so it is unlikely that the optimal distribution of resources can be achieved by means of mission-oriented grants alone.

Agenda setting and investment tied to specific research subjects are often performed as the responsibility of the research funding agencies, but consideration is rarely given to overlapping between research subjects. The observed complementarity is merely implemented as a result, not as a design element. Hence research funding agencies should give greater consideration to the diversity of science when determining the distribution of research subjects for the creation of effective and efficient policy.

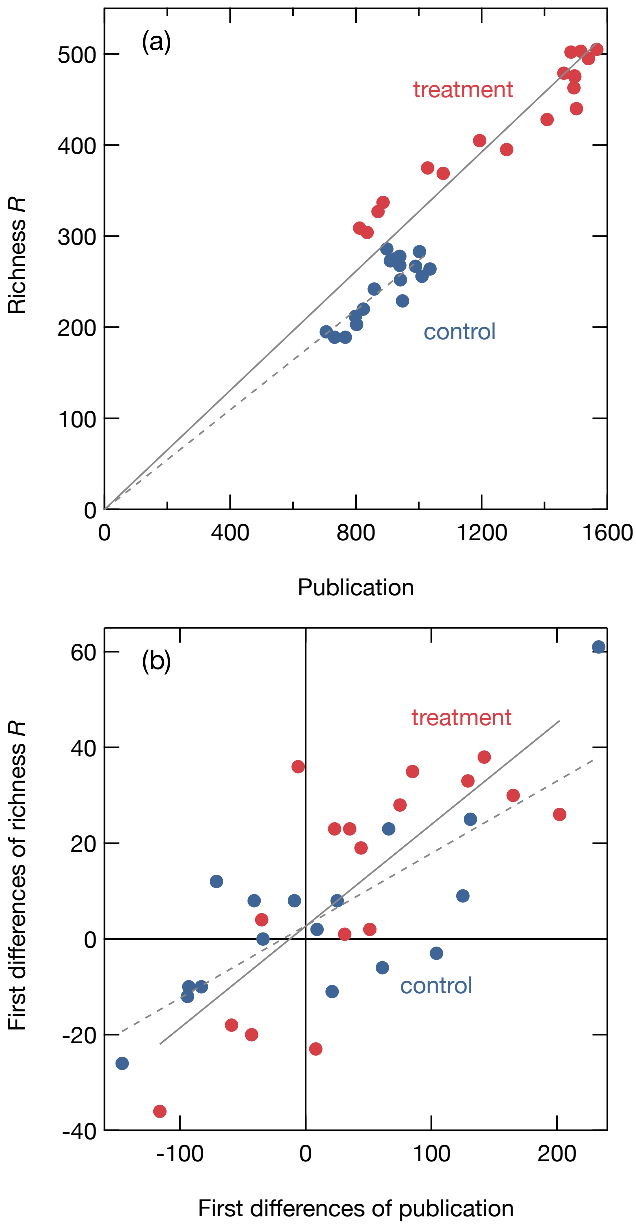


Figure 5.7: Linear dependence of richness on publication.
(a) The richness as a function of the number of paper for the treatment and the control groups between 1996 and 2013. The spurious correlation between time-series data are indicated by solid and dashed lines for the treatment and the control groups, respectively. The intercepts are set to zero for both cases. (b) The first differences of richness as a function of the first differences of publication. The difference in the slope of the regression line between the treatment and the control groups (plotted in solid and dashed lines, respectively) is not statistically significant.

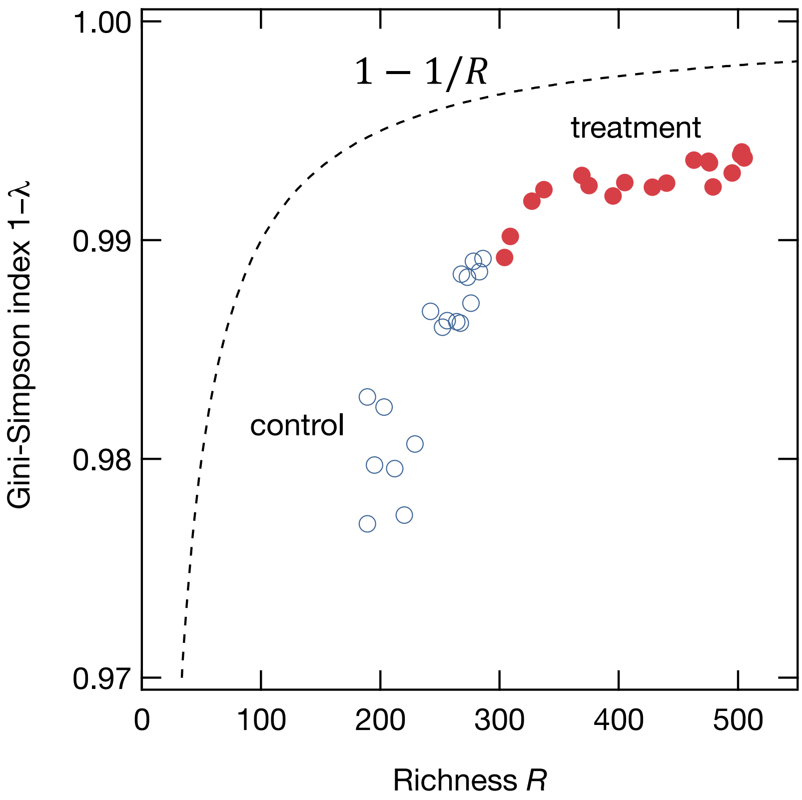


Figure 5.8: Plot of Gini-Simpson index vs. richness.
The Gini-Simpson index as a function of richness for the treatment (red solid circles) and the control groups (blue open circles) between 1996 and 2013. The dashed line represents the theoretical maximum $1 - 1/R$ of the Gini-Simpson index for each R. The observed advantage of the Gini-Simpson index for the treatment group may be associated with larger R for the group.

5.7 Conclusions and policy implications

The impact of mission-oriented research grants and curiosity-driven grants on the diversity of research subjects in Japan has been quantitatively compared.

First, the estimation of the probability of adoption has been performed by means of Probit regression using affiliation and position of applicants as well as publication performance. It should be noted that the skewed distribution of awardees in specific institutions and positions underlies the success of propensity score matching. The average treatment effect on the treated after propensity score matching strongly suggests that participation in the CREST program had a positive impact on publication performance during and after completion of the award period. This positive effect improved publication performance by more than 10% over initial publication count, perhaps reflecting an increase in the number of researchers working in teams. Citation count also increases with the participation to the CREST program. The difference between average and maximum citation number increased in proportion to the participation. Our results also suggest that the publication count of the top 10% of citations (a proxy indicator of the quality of published papers) did not correlate significantly with program participation.

Second, we have evaluated the diversity of research topics through analysis of the distribution of the classification codes applied to articles published by awardees of both grants. The group of articles published by the treatment group researchers displayed the wider variety and more equal distribution over categories of research subject than that of the control group researchers, as defined by propensity score matching. The better diversity indices are confirmed with the lognormal distribution with a smaller σ for the treatment group. The number of categories was higher for the treatment group mainly because of its larger number of articles. This implies that the treatment group, *i.e.* especially CREST program (and not mission-oriented research in general), was prone to diversification.

Contrary to an intuitive understanding of curiosity-driven research grants as a source of diversity in research, the diversity observed here appears to have been fostered by mission-oriented grants. This would imply that the concentrated investment by the CREST program in specific research targets effectively incentivized researchers to take up new research themes rather than continuing with their established research subjects, usually funded by the curiosity-driven KAKENHI program.

From the perspective of the diversification of science, research grants provide incentives for researchers to take up new research topics rather than persist with their already developed, conventional topics. Diversification of science requires certain changes in the behavior of scientists. The simplest incentive model assumed in this study is that in which allocation of budget induces a certain proportion of scientists to perform research in specific subjects. Shifts in research theme should be invoked, but the speed of change should be regulated by its own viscosity, as a reflection of individual preferences and social norms in the research community to which the researcher belongs. Our findings reveal limited diversification of science in Japan, especially in the fields of nanotechnology and materials science. The central subjects addressed by researchers appear to change little over time and to be robust against grant-based incentives. However, our results also demonstrate that the number of subjects increases with increased number of publications, thus detailed analysis on the shift of research subject by individual basis will be beneficial for understanding the mechanism of incentive.

In this survey, we only investigated the relatively large programs in KAKENHI such as SR-S, A, and B as listed in table 5.1, although these are still much smaller in budget size than CREST program. Therefore, this research will not answer how the difference between CREST and KAKENHI's allocation strategy (*i.e.* allocate a large research budget to one researcher or allocate a small budget to many researchers) affects the paper productivity and scientodiversity. Future research will be necessary. Notice that since we retrieved articles by the name of researchers, our dataset included the articles that are not related to CREST and KAKENHI grants. However, with assuming that the number of those unrelated articles to both CREST and KAKENHI are of the same level and their time change is on the same trend, it will be removed by the DID methodology to some extent.

According to this study on the impact of research grants on publication performance and topic diversity, several policy implications can be proposed as the followings.

First, scientists and funding agencies should consider the distribution of high-performers when they worry about the concentration of grants into a specific institution or research topic. The observed strong correlation between affiliation and awarding of a grant, *i.e.* concentrated resource allocation, is not immediately problematic because the concentration of researchers with relatively higher performance, *i.e.* high probability of awarding, in a specific institution provides an alternate explanation for the skewed distribution of awardees. On the other hand, performance distribution over topics cannot be easily managed in the peer review within a single discipline, thus it would be advisable for funding agencies to incorporate some sort of balance mechanism over research areas both in the agenda-setting and the selection process.

Second, the efficiency of grant programs should be evaluated after normalization; at least publication should be divided by an effective number of participating researchers. A simple explanation for the increase in publication numbers associated with the participation of grant programs is the increase of resources, especially human resources, resulting from the larger amount of research grant (the CREST program awarding more than the KAKENHI program). The increase in the number of collaborators and post-docs may have had a significant effect on the increase in the number of papers by the principal investigator, due to co-authoring.

Third, government research targets must be carefully designed based on the quantitative analysis of the current diversity of research subjects and evidence-based estimation of policy effect on diversity. To the best of our understandings, mission-oriented grants more promote diversity of science than curiosity-driven grants. However, the difference in the number of popular subjects, as presented in Figure 5.3, indicates the importance of complementarity between the two types of grant. Intensive discussion and development of appropriate indicators for program evaluation are clearly needed.

Furthermore, the principle of design (developed here) of incentives for researchers to focus in a specific research area should contribute to implementation work in the industry cluster. Concentrated investment in a specific research domain, *i.e.* mission-oriented grants, evokes urban economics and industry clustering (Delgado et al. (2012); Feldman and Kogler (2010)). This concentration improves the efficiency of transaction, transfer, and communication cost, and knowledge spillover can be expected. The observed co-occurrence of rich diversity and high productivity in mission-oriented research must be understood in the framework of the economy of urbanization (Jacobs (1969)) which utilizes diverse entities with different expertise and function.

Chapter 6

Discussion and Conclusions

6.1 Three aspects of scientodiversity

Management and promotion of scientodiversity have long been a prominent issue in the context of sociology and economics of science. Despite years of research, the precise formulations of scientodiversity and understandings of the relationship between resource allocation and scientodiversity are still limited, while the decline of research diversity of Japanese science has emerged as a policy concern in Japan.

This study has investigated the relationship between the diversity of research subjects, or scientodiversity, and resource allocation on science aiming a new design of diversity-aware resource allocation, especially as a research grant program. Our approach based on three different aspects of resource allocation, such as total amount, distribution, and types, has revealed positive and negative impacts of resource allocation on scientodiversity in the corresponding three scales, *i.e.* country, university, and research team. It has also been established quantitatively methodology to measure scientodiversity inspired by biodiversity studies through this research.

Diversity, in general, represents three different aspects, such as variety, balance, and disparity. This study mainly focused on variety and balance aspects of scientodiversity measured by the richness, *i.e.* the total number of research subjects, and Gini-Simpson index, corresponding to the shape of the distribution function of subject classification codes attached to research articles. Promotion of scientodiversity is intuitively understood as an increase in variety, but it does not mean improvement in balance. On the contrary, an excessively skewed distribution is often regarded as problematic implicitly, but skewness can be necessary to some extent from the viewpoint of overall optimization.

From this viewpoint, this chapter first discusses two aspects of diversity, *i.e.* variety and balance, separately, then considers policy implications by the integrated view. The third aspect, disparity, is discussed as a limitation in section 6.5.

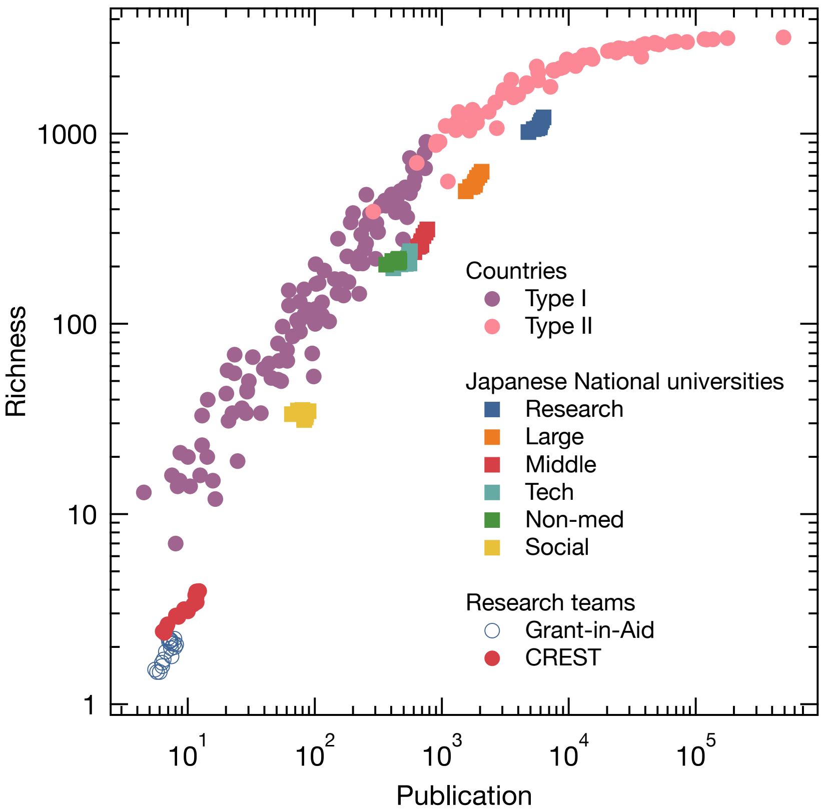


Figure 6.1: Multi-scale comparison of scientodiversity.
The richness as a function of the number of publication. The purple and pink solid circles represent the number of publication and richness for for each countries. All values are averaged over 10 years between 2001 and 2010. The solid squares represent the number of publication and richness for each Japanese national university in a specific year between 2002 and 2011. The colors of square are corresponding to the classification of university discussed in Chapter 4. The red solid and blue open circles represent the number of papers and richness in a specific year between 1996 and 2013 for the treatment and the control groups, respectively. Both values are the average over corresponding PIs.

6.2 How to increase richness?

Richness is determined by the total amount of resources. This is observed remarkably on the country scale as discussed in Chapter 3. Figure 6.1 shows a comprehensive summary of richness-publication plots for the scale of the country, university, and research teams. The richness shows the exponential dependence of publication like the richness-ERD plot (Figure 3.6). The slope varies with an increase of the scale. It can be understood as the diminishing return to scale for richness production (with regarding publication as an input). Therefore, in order to improve richness, basically it is only necessary to increase the total amount of resources, but concentrating resources into specific universities/teams and/or research fields becomes inefficient.

As shown in Figure 6.1, the observed richness of Japanese national universities is always smaller than that of the countries having a comparable amount of publication. While the richness-publication curve of Japanese national universities is always below that of countries, the slope of them is relatively similar. This implies the density of research subjects is almost universal, *i.e.* the macroscopic mechanism for fertilizing scientodiversity may be independent of whether the considering group is university or country.

The gap between country's curve and university's one implies a possible improvement of richness without increase publication. However, since the richness of the whole system cannot be improved by the summation of local richness for each elemental units, improvement of Japan's (or the whole Japanese national university's) richness cannot be accomplished by just increasing the richness of individual university as an average. This observed gap (Figure 6.1) perhaps reflects the university's scientific role sharing. In the industrial cluster policy, the role sharing is regarded as an implementation of the economy of scale. In this sense, variety within each cluster is not so important but overall necessary role must be shared by each cluster. Thus, richness within a specific cluster can be much smaller than that of an entire system of clusters. On the other hand, the cross-fertilization within a cluster is also regarded as an essential mechanism for the successful cluster. Therefore, it is necessary to consider scientodiversity, in particular, the variety of research topics, in terms of both the role sharing and interdisciplinarity. The role share behavior of Japanese national universities is also evident as shown in Figure 4.5.

The sum of budgets for individual universities is equal to the overall budget, but richness cannot be summed in the same manner. Because there are many overlapping research subjects among universities, the university's overall richness is always smaller than the sum of the richness of individual universities. In other words, the overall diversity is represented by both diversity *in* individual universities and diversity *among* those universities. Such a situation can be formalized by the concept called β-diversity in biodiversity studies. In ecology, the species diversity in a specific ecosystem as a whole is often regarded containing diversity at the habitat level and the variety among those habitats. The former diversity is called α-diversity and the latter is called γ-diversity. The β-diversity (or so-called "absolute species turnover") is quantified as the difference between γ and α-diversity. In our case, the richness of individual university represents α-diversity, and that of the whole country is γ-diversity. In order to promote γ-diversity (diversity at country level), we must consider α and β diversity at the same time.

The role of the university is not only research, as intensively investigated and discussed in this thesis, but also education. In many countries, higher education is a top priority of the university's role. Japanese national universities have been established in all prefectures as shown in Table 2.7 and are responsible for an equal opportunity for higher education. In this thesis, we mainly focused on the aspect of research (see also Table 2.8), but variety and balance of education should also be discussed at the same time. From this viewpoint, the observed role-sharing over research subjects may prevent the equal opportunity for higher education. Although the paper productivity or employment rate of graduates can be one metric for evaluation of outcome, it still needs to measure by other metrics. The β-diversity among universities is such candidate index for understanding and monitoring of the activity of universities as an ecosystem.

The type of budget may also associate with the richness. The importance of external grants from private companies has been pointed out because of the limitation of government's investment in particular for Japanese national universities (National Institute of Science and Technology Policy (2015); Sunami (2017)). However, in many cases about Japanese universities, the motivation for obtaining external grants from companies derives from the limited investment by the government.

The result of this study implies that the acquisition of external budget has a great influence on the promotion of scientodiversity. In the university scale analysis, a statistically significant correlation was found between the DEA efficiency of diversification and the ratio r_{com} of external grants from private companies. This should be understood as a bi-directional relationship, not as a one-directional causal relationship. It can be the direction that the external funds promote scientodiversity and the direction that universities with diverse research portfolio are more likely to acquire external funds from companies. Also seen in the team scale analysis, the mission-oriented grant is found to be more effective in promoting scientodiversity than the curiosity-driven grant. This study suggested that external funds play an important role in promoting scientodiversity, not just for the amount of budget.

6.3 How to ease skewed distribution?

As introduced in Chapter 2, the relationship between resource allocation and scientodiversity can be written as a simple binomial distribution process. The balance of research subject is represented by a lognormal parameter sigma, which can also be measurable by Gini-Simpson index as discussed in Appendix A.

In order to improve balance, the program design of (A) increasing the adoption rate or (B) decreasing difference between awardee and non-awardee (like a basic income) are considered as a possible policy option. Using the simple binomial distribution effective model introduced in Chapter 2, we set the operational design parameters z_1, z_2, and p of an effective grant program, and change of those parameters so that the lognormal parameter σ decrease (*i.e.* balance is improved) by keeping the total allocation budget constant (*i.e.* the average $\bar{X}$ is constant).

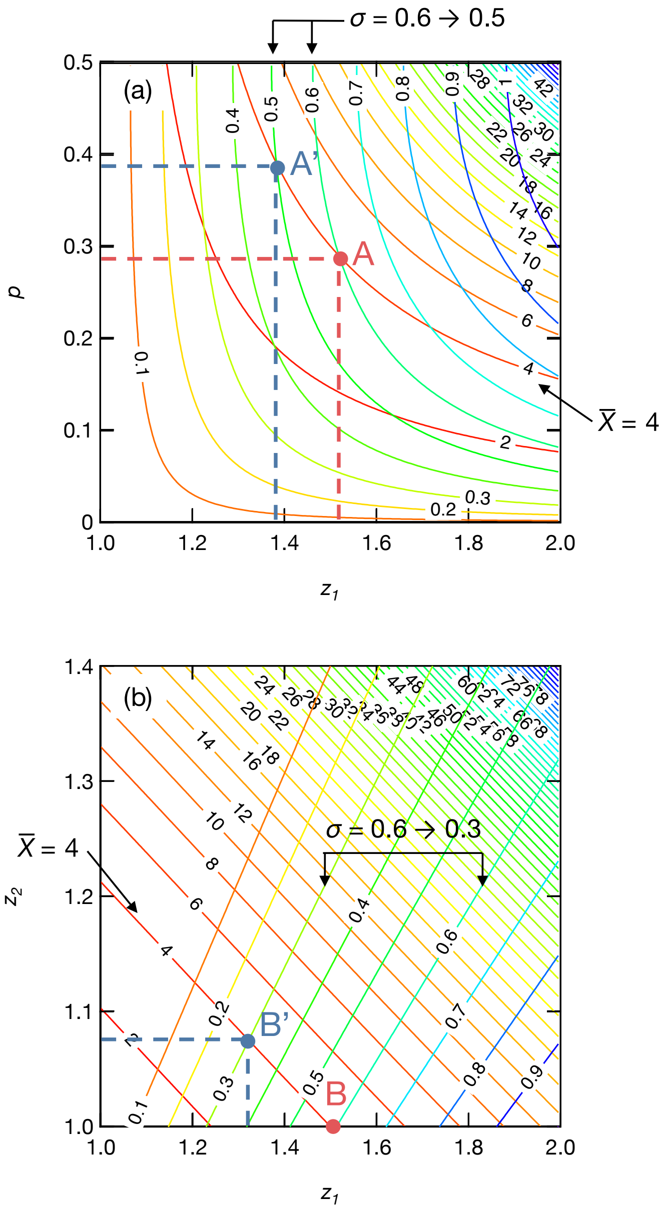


Figure 6.2: Possible improvement of evenness by (a) constant z_2 and (b) constant p. (a) With keeping $z_2 = 1$ and $\overline{X} = 4$, the increase of probability p from 0.288 to 0.385 and the decrease of the value z_1 from 1.521 to 1.387 at the same time can decrease the lognormal parameter σ from 0.601 to 0.503 shown as the point A and A'. (b) As an alternative option, we consider the increase of z_2 instead of p from 1 to 1.074 and the decrease of z_1 from 1.499 to 1.323 at the same time may reduce σ from 0.578 to 0.302 with keeping $p = 0.3$ and $\bar{X} = 4$.

The increase of adoption rate is equivalent to increase of probability p. For example, when one increase p from 0.288 to 0.385 and the realization value z_1 is reduced to 1.521 to 1.387 at the same time to keep $X = 4$, the lognormal parameter sigma decrease from 0.601 to 0.503 shown as the point A and A' in Figure 6.2(a) (see also Figure 2.4). As an alternative option, we consider the increase of z_2 instead of p. Figure 2.5 shows the average $\bar{X}$, the median $\tilde{X}$, and sigma are plotted by changing z_1 and z_2 with $p = 0.3$. The hatched area in Figure 2.5 is excluded by definition $z_2 > z_1$. By increasing z_2 without changing, the lognormal parameter σ will decrease, but the average $\bar{X}$ also increases unintendedly. Thus, z_1 should be reduced to keep $\bar{X}$ constant. If one increase z_2 from 1 to 1.074 and decrease z_1 from 1.499 to 1.323 at the same time, σ will be reduced from 0.578 to 0.302 with constant adoption rate $p = 0.3$ and average $\bar{X} = 4$. This situation ($z_2 > 1$) means that a certain (relatively small) grant will be granted to all applicants from any area of science, even their proposal is not accepted by a funding agency. In other words, a part of the resource originally planned to be allocated to awardees will be distributed to non-awardees, and thus skewness of the distribution is mitigated. This design options (B) contributes more effectively to the improvement of balance than that an option (A) does, and it can be expected to promote research activity in the relatively unpopular research subjects, i.e. the research area where the number of articles is small at present. Although criticism on the apparent inefficiency of resource allocation (so-called "baramaki" in Japanese) may not be avoidable, the establishment of a funding program that supports a few research expenses but holds a high adoption rate is strongly required for promotion of scientodiversity. Such (almost) unconditional allocation of research grants can also reduce the management cost of the funding agency. The institutional grant scheme represented by the Management Expenses Grants for Japanese national university should be an option to be reevaluated from the scientodiversity viewpoint.

This model simulation and results from our empirical study in team-scale suggest that the program design and investment strategy by the government and/or funding agencies have a great influence on scientodiversity. In the case of Grants-in-Aid for Scientific Research (KAKENHI), a key program design that affects scientodiversity is that they provide constant acceptance rate for all research disciplines.

This constant-rate strategy, *i.e.* $p_i = p$ for all subject i in our model (see Chapter 2), seems quite fair for all research subjects but is eventually becoming a strategy to focus on the popular research areas. As discussed in Chapter 2, the random multiplicative process with assuming $p_i = p$ makes the distribution log-normal shape even starting from the flat or normal distribution. If one can change acceptance ratio p by some broad category (not necessary to be equivalent to scientific discipline), the skewed distribution of research subjects can be manageable. Of course, heavily skewed distribution of p_i should be problematic because it can again generate a skewed distribution of subjects as a result.

Table 6.1: Possible improvement of evenness.

The example parameter set discussed in Figure 6.2. The points A, A', B, and B' represent each possible parameter sets. The lognormal parameter σ can be improved from 0.601 to 0.503 by the increase of probability p from 0.288 to 0.385 (and the decrease of the value z_1 from 1.521 to 1.387) with keeping $z_2 = 1$ and $\overline{X} = 4$ as described by the point A and A'. The lognormal parameter σ can also be improved from 0.578 to 0.302 by the increase of z_2 from 1 to 1.074 (and the decrease of z_1 from 1.499 to 1.323 simultaneously to keep $p = 0.3$ and $\bar{X} = 4$) as shown in the point B and B'. Notice that this analysis is a simulation study under specific conditions. Thus the interpretation of the result needs special care on the definition of each parameters. In particular, the parameter p is not the adoption rate of a grant but the *effective* adoption rate parameter in our model.

Point	z_1	z_2	p	$\bar{X}$	$\tilde{X}$	σ
A	1.521	1.000	0.288	4.00	3.34	0.601
A'	1.387	1.000	0.385	4.00	3.52	0.503
B	1.499	1.000	0.300	4.00	3.37	0.587
B'	1.323	1.074	0.300	4.00	3.82	0.302

Therefore, a program with a large ratio between z_1 and z_2 and a small acceptance rate p, like JST CREST program (see Table 5.1) should be careful not to be fixed priority field in terms of scientodiversity. To evaluate those funding activities and program designs, it is necessary to monitor both inputs and outputs by the scientodiversity indicators such as richness and Gini-Simpson index.

The model used in the above discussion relies on many assumptions. The model assumes the linear relationship between R&D expense and publication, *i.e.* the constant returns to scale. In a real situation, the production function often shows decreasing returns to scale as shown in Figure 3.6 in the country-scale analysis. Moreover, the research productivity, which may be measured by the proportionality coefficient between R&D expenses and the number of articles, may vary depending on the field. For example, it is easy to imagine that paper productivity of pure mathematics is much smaller than that of experimental molecular biology as illustrated in Figure 4.4.

Given decreasing returns to scale, the skewness of the distribution of research subjects generated by a certain random multiplicative process will be more moderate. Thus, the above discussion presents an upper bound of change in sigma, while keeping its lognormal shape. But, the difference in productivity among research subjects, in some case, may contribute in the opposite direction. If a specific subject has much better productivity than that of others and simultaneously shows weaker decreasing returns to scale, the real distribution may be more skewed one than the above discussion.

6.4 Comprehensive policy implication

This research aims to explore the new design of resource allocation by revealing and understanding the relationship between scientodiversity and resource allocation on three different scales. In this section, we introduce possible policy implications on the total amount, distribution, and types of resource allocation through an integration of our considerations on variety and balance discussed in previous sections.

For the impact of the total amount of resource allocation, the country-scale analysis (Chapter 3) tested the analogy between biodiversity and scientodiversity in terms of the lognormal distribution function. The result showed the positive linear relationship between richness and the total R&D expenditure in a log-log plot. Therefore, the increase of the total amount of resource, *e.g.* research budget and/or the number of researchers could be the simplest policy option to increase richness in a country-scale. The relationship between the total resource and the balance of research subjects has not been so clear in this study, but our observations suggest that the total budget may not change the balance so much, *i.e.* the shape of distribution function will be kept as lognormal form. Therefore, an increase in R&D expenditure may be a promising approach that can increase richness without affecting the balance so much.

However, it may not be easy for the Japanese government to increase total R&D investment, because they especially have to care expenditures on other than science. Thus, the efficient distribution of resource within a given total is (only) a feasible option, although the richness of country scale may not increase without increasing of the total budget in a macroscopic viewpoint. This problem setting can be formulated as a question of how the budget should be allocated to each Japanese national university in a given total R&D budget. This may be justified by the fact that national universities have quite a large part of the production of scientific articles throughout the country in particular in fundamental research. One possible approach would be the facilitation of the variety among national universities (β-diversity). Even if the sum of budgets is constant, the sum of richness (γ-diversity) may increase when the role-sharing on research subject properly happens to all national universities. The gap between the country-scale curve and the university-scale one shown in Figure 6.1 suggests that there is a room for an increase in richness for Japanese national universities even if the number of papers is constant (that is, the budget is constant). Then, the next problem should be how can we stimulate such a role-sharing (by grants)?

In order to ease the skewness of the balance of research subjects, it is necessary to reconsider the resource allocation as a system. From this study, we saw that lognormal distribution appears when we repeatedly allocate resources proportional to the number of papers, regardless of whether it is intended or not. To mitigate the skewness of this lognormal distribution, the previous section proposed two types of change in grant system. Both of the proposals reduce the difference between awardee and non-awardee. Changing the type of resource may be another promising option to change scientodiversity under the condition that the total amount of resources is fixed.

The results of chapter 4 and 5 suggest that subsidies from private companies and mission-oriented grants from the government may be associated with the promotion of scientodiversity. However, the result we have shown in this thesis is about the relationship between a specific grant program (CREST) and scientodiversity, and thus there is no guarantee that the general mission-oriented grants promote scientodiversity as well. The funding by the private company to Japanese national universities and the research efficiency in terms of publication and scientodiversity showed a positive correlation, but it is not a causal relationship. Therefore, there is no guarantee that the increase of the ratio of grants from the private company promote scientodiversity. Further research is necessary to reveal the relationship between fundings and scientodiversity in detail.

6.5 Limitation

In this research, we investigated the impact of the research funding on scientodiversity. This diversity here refers to the diversity of research subjects as indicated in Chapter 1 and we have not investigated the balance of the research team's interdisciplinarity in their expertise and the geographical and gender balance. In this respect, the policy implications suggested from this study are effective only for specific "diversity". Even about the diversity of the research subjects, or scientodiversity, various factors other than the research funding, such as the researcher's mentality, the national character, academic customs in a specific discipline, and the level of higher education, can influence scientodiversity. As a result, the effectiveness of the policy option proposed by this thesis may vary from country to country or from the field to field. In this research, we observed the inter-university diversity (or β-diversity) in the survey of 69 Japanese national universities as discussed in Chapter 4. This suggests the possibility of a new strategy to promote scientodiversity other than the resource allocation or the grant programs.

For the country-scale analysis, observed scientodiversity of Japanese science, by means of the richness of research subject, is not inferior to that of Germany and the UK from the result of this study as contrastive to the result of previous research (Igami and Saka (2016)). This difference perhaps due to the difference of dataset. In this study, the number of articles is counted based on the articles in the journals recorded in both J-Global and Scopus. The difference of the collection among three databases is shown in Figure 2.1. The counting only highly-cited articles also make dataset so different. Therefore, the analysis applying our method proposed in this thesis to the dataset that other research use, *i.e.* highly-cited articles only and the whole Scopus dataset, is worth to be investigated in the future.

In this research, we focused on variety and balance among three aspects of diversity (Stirling (2007)). The remaining aspect, disparity, which represents difference and similarity among research subjects. For disparity aspect of diversity, this study implicitly assumes $d_{ij} = 1$ (Table 2.1) for all subject i and j, by given classification code in J-Global database. However, same as Linnaean taxonomy sometimes differs from the gene-based classification, the JST classification code of the database is somewhat subjectively attached and its classification schema may not necessarily be objectively justified. Notice that the research subjects and biological species are quite different in their definitions in terms of its subjectiveness. The question "how to evaluate disparity?" is equivalent to "how to objectively classify research subject?". Methodologies to measure similarity based on co-citation and text-mining have been proposed and compared with each other recently (Dias et al. (2018)). The results of our preliminary study are shown in the Appendix C.

6.6 Future works

It is necessary to develop high-resolution classification scheme and to systematically assign them to large datasets. Text-mining of article titles and abstract is a promising approach to create a fine-granular classification code from an objective viewpoint. At the same time, this technology also enables classification of any text other than papers. The text-mining analysis of research proposals and project descriptions should be explored as a scientometric study. Fine classification of these non-articles including policy document is well expected to give important information to reveal the relationship between resource allocation and scientodiversity.

Such high-definition classification codes will also raise the problem on "species definition" in science. As seen in Chapter 3, the statistical behavior of research subjects is somewhat analogous with biological species. However, this is not trivial. Rather, it is strange that an analogy is established in the statistical distribution obtained as a result of dynamics, although there is a clear big difference between species and subject. For example, the biological species may produce another new species, but the counterpart of this biological feature has not been mentioned in this thesis. Difference between multi- and inter-disciplinary research may explain the difference between biological species and the research subject.

In this study, we did not pay much attention to the interaction between research subjects, i.e. we focused only on the dynamics of each subject which is assumed to be determined independently from other subjects. The dynamics through inter-subject interaction between subject i and j may be reflected in the off-diagonal components of disparity matrix d_{ij}. Thus, the interactive activity of research subjects perhaps has great influence on the distribution shape. This direction should be focused as the modeling of dynamics in interdisciplinarity.

In this study, we did not pay attention to the interaction between species but focused only on the dynamics of the diagonal components. As the result of the interaction between the subjects is statically reflected in disparity d_{ij}, the influence on the distribution shape is not small at all. This direction should focus on modeling dynamics as science ecology. In ecology, such species-species interaction is regarded as a key clue to understanding the stability of the ecosystem. In particular, a specific species called *keystone species*, which is defined as a species that is relatively rare but showing a large effect on the whole ecosystem (Paine (1995)), has attracted many attentions from the viewpoint of conservation of biodiversity and research and discussion on keystone species is also flourishing. In terms of scientodiversity, finding keystone science may be an important policy target at the national level.

In addition, it would also be worthwhile to consider the concept of the concentration and dispersion in the temporal dimension, not in the spatial dimension (*i.e.* the distribution among universities or research teams) and the semantic dimension (*i.e.* the distribution over research subjects). For example, even if a resource allocation is concentrated on a specific university and/or research area at some time, the average resource allocation over 10 years may not so skewed. Indeed, the priority area of CREST program, which is an iconic concentrated investment, seems to be changed periodically and covers the certain variety of research topics to some extent when looking at the 10-year span. It will be necessary to cultivate such policy options based on the evidence.

Appendix A

Distribution and diversity indices

A.1 Lognormal distribution

The lognormal distribution of research subject is written in a standard form as

$$S(n) = S_0 \exp\left[-\frac{(\ln n - \mu)^2}{2\sigma^2}\right], \tag{A.1}$$

where n represents number of paper, and S_0, μ, and σ are parameters. Following Preston's works, using logarithms to the base 2 is convenient for consistent discussion with studies of ecology. The distribution, in practice, lies between $\log_2 n_{\min}$ and $\log_2 n_{\mathrm{mode}}$ as shown in Figure A.1. The number of papers in the specific research subjects can be written as

$$\begin{aligned}
nS(n) &= nS_0 \exp\left[-\frac{(\ln n - \mu)^2}{2\sigma^2}\right] \\
&= S_0 n_0 2^r \exp\left(-a^2 r^2\right) \\
&= S_0 n_0 \exp\left(-a^2 r^2 + r \ln 2\right) \\
&= S_0 n_0 \exp\left[\frac{(\ln 2)^2}{4a^2}\right] \exp\left[-a^2\left(r - \frac{\ln 2}{2a^2}\right)^2\right],
\end{aligned}$$

where $n_0 = \exp\mu$, $r \equiv \log_2(n/n_0)$ and $a \equiv \ln 2/\sqrt{2\sigma^2}$. This distribution should be seen to be a Gaussian function with a peak displaced a distance $\ln 2/2a^2$ to the right of $S(n)$ distribution.

The parameter γ is defined as the ratio between $\log_2 n_{\mathrm{mode}}$ and $\log_2 n_{\max}$:

$$\gamma \equiv \frac{\log_2 n_{\mathrm{mode}}}{\log_2 n_{\max}} = \frac{\sigma}{\sqrt{2}\sqrt{\ln S_0}}. \tag{A.2}$$

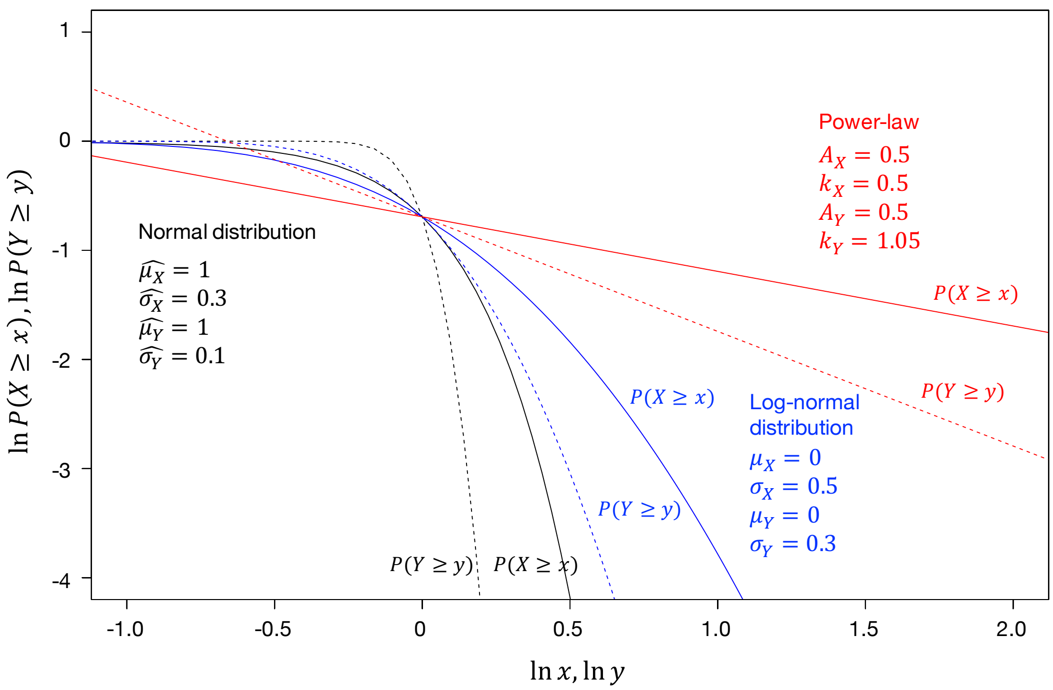


Figure A.1: Three types of distribution function.

A.2 Diversity indices

The richness is computed as a function of S_0 and σ;

$$
\begin{aligned}
R &= \int_0^\infty S(n)dn \\
&= S_0 \int_{r_{\min}}^{r_{\max}} \exp\left(-a^2 r^2\right) dr \\
&= \frac{\sqrt{\pi}}{a} \exp\left(\Delta^2\right) \operatorname{erf}(\Delta),
\end{aligned}
$$

where $\Delta \equiv \sqrt{\ln S_0}$ and the error function $\operatorname{erf}(x)$ defined as

$$
erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp\left(-t^2\right) dt. \tag{A.3}
$$

For the asymptotic case of $\Delta > 1$, the richness can be written as

$$
R \simeq \frac{\sqrt{2\pi\sigma^2}}{\ln 2} \tag{A.4}
$$

since the error function is almost equal to unity ($\operatorname{erf}(\Delta) \simeq 1$).

The Gini-Simpson index can be written in analytical form as

$$
\begin{aligned}
1-\lambda &= 1-\int_0^\infty \left(\frac{n}{N}\right)^2 S(n)dn \\
&= 1-\frac{\sqrt{\pi}\exp\left[\Delta^2(1+2\gamma)^2\right]}{2aJ^2}\left\{\mathrm{erf}\left[\Delta(2\gamma+1)\right]-\mathrm{erf}\left[\Delta(2\gamma-1)\right]\right\},
\end{aligned}
$$

where a parameter J is defined as

$$
J \equiv \frac{N}{m} = \frac{\sqrt{\pi}}{2a}\exp\left[\Delta^2(1+\gamma)^2\right]\left\{\mathrm{erf}\left[\Delta(1-\gamma)\right]+\mathrm{erf}\left[\Delta(1+\gamma)\right]\right\}, \tag{A.5}
$$

where m represents the expected number of papers in the most unpopular research subjects.

For $\gamma < 1$, the Gini-Simpson index can be approximated as

$$
1-\lambda \simeq 1-\frac{a}{2\pi\Delta(2\gamma-1)}\exp\left[-2\Delta^2(1-\gamma)^2\right]. \tag{A.6}
$$

Then, by using eq. A.4, one can write λ as a function of R,

$$
\lambda \propto R^{-2(1-\gamma)^2}. \tag{A.7}
$$

A.3 Subject-budget relationship

For the limit $R \gg 1$, the eq. A.4 can be written as

$$
\ln R \simeq \Delta^2\left[1+\frac{1}{\Delta^2}\ln\left(\frac{\sqrt{\pi}}{a}\right)+\cdots\right] \sim \Delta^2. \tag{A.8}
$$

In the same manner, we have the asymptotic approximation of J (eq. A.5) for $\gamma < 1$ as

$$
\ln J \simeq \Delta^2(1+\gamma)^2\left[1+\frac{1}{\Delta^2(1+\gamma)^2}\ln\left(\frac{\sqrt{\pi}}{a}\right)+\cdots\right] \sim \Delta^2(1+\gamma)^2, \tag{A.9}
$$

since $\mathrm{erf}\left[\Delta(1-\gamma)\right]$ lies between 0 and 1. For $\gamma > 1$, $\mathrm{erf}\left[\Delta(1-\gamma)\right]$ will be negative. Then eq. A.9 takes

$$
\ln J \simeq 4\Delta^2\gamma\left[1-\frac{1}{4\Delta^2\gamma}\ln\left(\frac{(\gamma-1)\ln 2}{\gamma}\right)+\cdots\right] \sim 4\Delta^2\gamma. \tag{A.10}
$$

From the eqs. A.8, A.9 and A.10, we have an approximate relationship between R and J by means of one parameter γ as $R \sim J^z$, where $z = 1/(1+\gamma)^2$ for $\gamma < 1$ or $z = 1/4\gamma$ for $\gamma > 1$. The subject-budget curve is obtained from the additional assumption $J = \rho B$ which represents equal accessibility of resource, where ρ and B represent density of budget per research subjects and budget, respectively. Then, one can obtain the power-law relationship $R \propto B^z$. Notice that all approximations discussed above have been done by assuming large R. For small R, *e.g.* $R < 10$, the subject-budget curve is expected to be steeper than $z = 1/4$. From eqs. A.4 and A.5, a linear relation $R \simeq J$, *i.e.* $z = 1$, is roughly estimated.

A.4 Canonical lognormal distribution

For the special case of $\gamma = 1$, the richness R and the parameter J can be calculated by the single parameter Δ:

$$R \simeq \frac{2\sqrt{\pi}}{\ln 2} \Delta \exp(\Delta^2), \tag{A.11}$$

and

$$J \simeq \frac{\sqrt{\pi}}{\ln 2} \Delta \exp(4\Delta^2) \mathrm{erf}(2\Delta). \tag{A.12}$$

Then, the subject-budget relationship can be written as

$$R \propto B^{1/4} \tag{A.13}$$

with assuming $J = \rho B$.

Appendix B

Scale economy and distributions

B.1 Lognormal distribution

The shape of production function $Y = F(X)$, which represents the relationship between input X and output Y, is determined by the shape of distribution function of X and Y. Let the cumulative distribution function of X and Y be represented with at most two parameters m and s as follows:

$$P(X \geq x) = \Phi_X(x; m_X, s_X), \tag{B.1}$$

$$P(Y \geq y) = \Phi_Y(y; m_Y, s_Y). \tag{B.2}$$

Then, assuming that the production function $F(X)$ monotonically increase with increase of X, the distribution function of Y can be rewritten as

$$P(Y \geq y) = P\left(F(X) \geq y\right) \tag{B.3}$$

$$= P\left(X \geq F^{-1}(y)\right). \tag{B.4}$$

by use of the inverse function of production function $F^{-1}(y)$. Then, the relationship between two distribution functions is defined as

$$\Phi_Y(y; m_Y, s_Y) = \Phi_X(F^{-1}(y); m_X, s_X). \tag{B.5}$$

When both $\Phi_X(x; m_X, s_X)$ and $\Phi_Y(y; m_Y, s_Y)$ are lognormal distribution with parameter μ and σ, the cumulative distribution function can be explicitly wrote down as

$$\Phi_X(x; \mu_X, \sigma_X) = \frac{1}{2}\mathrm{erfc}\left(\frac{\ln x - \mu_X}{\sigma_X\sqrt{2}}\right) \tag{B.6}$$

$$\Phi_Y(y; \mu_Y, \sigma_Y) = \frac{1}{2}\mathrm{erfc}\left(\frac{\ln y - \mu_Y}{\sigma_Y\sqrt{2}}\right), \tag{B.7}$$

where the complementary error function $\mathrm{erfc}(x)$ is defined as

$$\mathrm{erfc}(x) = 1 - \mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty e^{-t^2}dt. \tag{B.8}$$

Then, the equation B.5 can be written as

$$\mathrm{erfc}\left(\frac{\ln y - \mu_Y}{\sigma_Y\sqrt{2}}\right) = \mathrm{erfc}\left(\frac{\ln F^1(y) - \mu_X}{\sigma_X\sqrt{2}}\right), \tag{B.9}$$

thus

$$\frac{\ln y - \mu_Y}{\sigma_Y\sqrt{2}} = \frac{\ln F^1(y) - \mu_X}{\sigma_X\sqrt{2}}. \tag{B.10}$$

The inverse function of the production function can be formalized as

$$\ln F^{-1}(y) = \frac{\sigma_X}{\sigma_Y}\ln y + \frac{\mu_X\sigma_Y - \mu_Y\sigma_X}{\sigma_Y}, \tag{B.11}$$

and then we have the production function $F(X)$ as the Cobb–Douglas production function:

$$F(X) = \exp\left(\frac{\mu_X\sigma_Y - \mu_Y\sigma_X}{\sigma_X}\right)X^{\frac{\sigma_Y}{\sigma_X}}. \tag{B.12}$$

B.2 Power-law and normal distribution

When both $\Phi_X(x; m_X, s_X)$ and $\Phi_Y(y; m_Y, s_Y)$ are power-law distribution as followings;

$$\Phi_X(x; A_X, k_X) = A_X x^{-k_X} \tag{B.13}$$

$$\Phi_Y(y; A_Y, k_Y) = A_Y y^{-k_Y}. \tag{B.14}$$

The inverse function of the production function can be formulated as

$$F^{-1}(y) = \left(\frac{A_X}{A_Y}\right)^{\frac{1}{k_X}} y^{\frac{k_Y}{k_X}} \tag{B.15}$$

by using

$$A_Y y^{-k_Y} = A_X \left(F^{-1}(y)\right)^{-k_X}, \tag{B.16}$$

which is computed from equation B.5. We have the production function $F(X)$ again as the Cobb–Douglas shape:

$$F(X) = \left(\frac{A_Y}{A_X}\right)^{\frac{1}{k_Y}} X^{\frac{k_X}{k_Y}}. \tag{B.17}$$

In the case that both $\Phi_X(x; m_X, s_X)$ and $\Phi_Y(y; m_Y, s_Y)$ are normal distribution, the equation B.5 can be rewritten as

$$\operatorname{erfc}\left(\frac{y - \widehat{\mu_Y}}{\widehat{\sigma_Y}\sqrt{2}}\right) = \operatorname{erfc}\left(\frac{\ln F^1(y) - \widehat{\mu_X}}{\widehat{\sigma_X}\sqrt{2}}\right), \tag{B.18}$$

where $\widehat{\mu_X}$, $\widehat{\sigma_X}$, $\widehat{\mu_Y}$, and $\widehat{\sigma_Y}$ are parameters of normal distribution, *i.e.* $\widehat{\mu}$ and $\widehat{\sigma}$ represent the mean and the standard deviation, respectively. The inverse function of the production function is written as

$$F^{-1}(y) = \frac{\widehat{\sigma_X}}{\widehat{\sigma_Y}} y + \frac{\widehat{\mu_X}\widehat{\sigma_Y} - \widehat{\mu_Y}\widehat{\sigma_X}}{\sigma_Y}. \tag{B.19}$$

Then, we have the production function in the linear shape, i.e. the constant return to scale;

$$F(X) = \frac{\widehat{\sigma_Y}}{\widehat{\sigma_X}} X + \frac{\widehat{\mu_X}\widehat{\sigma_Y} - \widehat{\mu_Y}\widehat{\sigma_X}}{\widehat{\sigma_X}}. \tag{B.20}$$

The shape of production function with given shape of distribution functions, including the case where the functional shape of $\Phi_X(x; m_X, s_X)$ is different from that of $\Phi_Y(y; m_Y, s_Y)$, are summarized in Figure B.1.

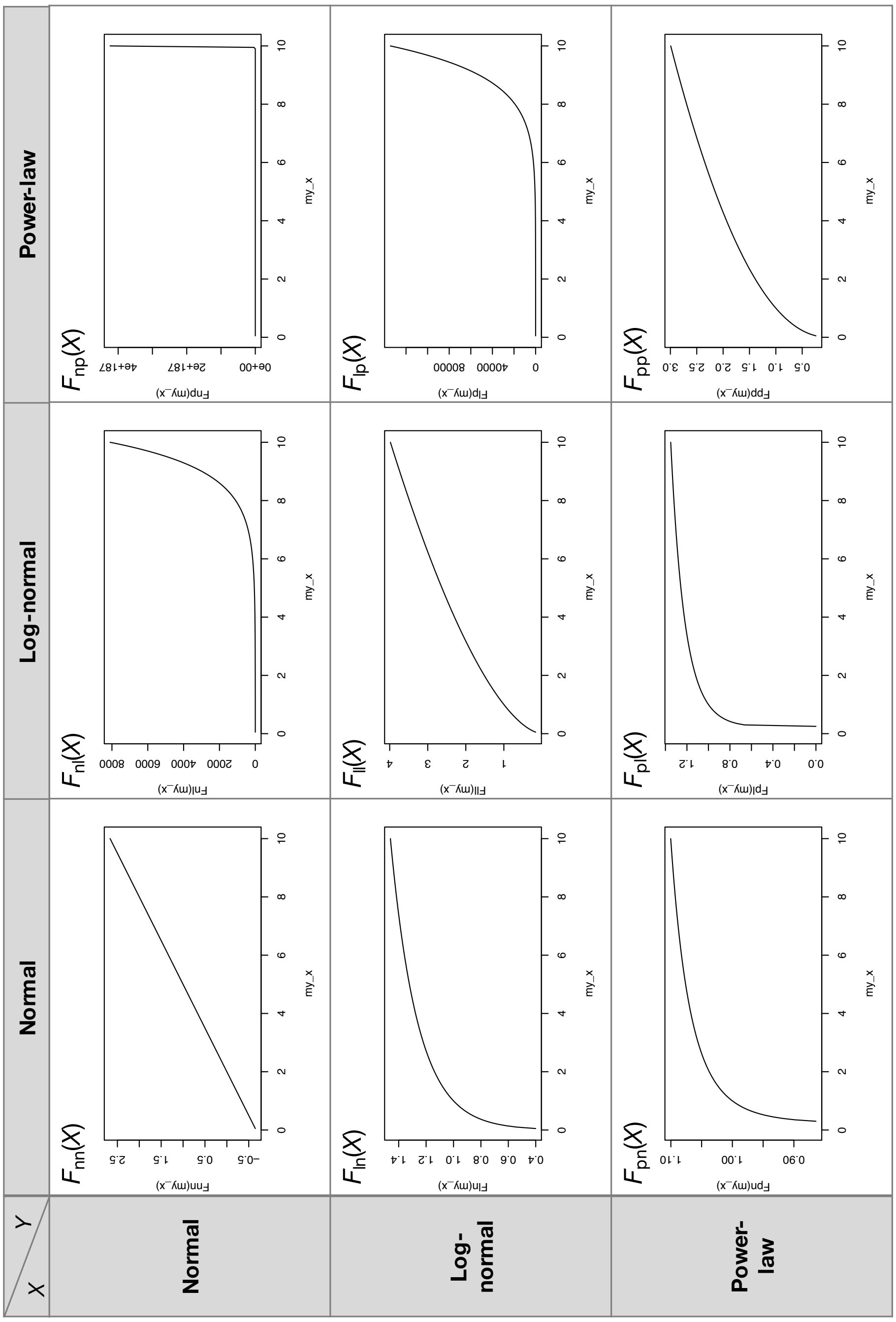


Figure B.1: Summary of shape of production function.

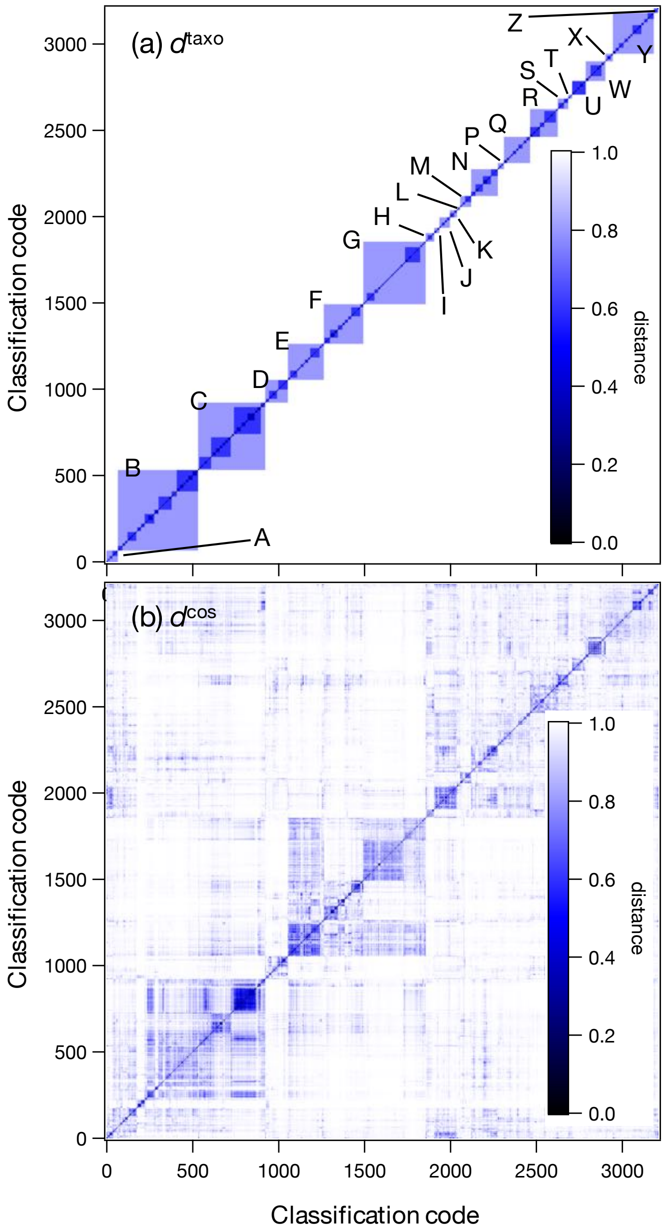
Appendix C

Disparity

There are two possible methods to formulate disparity matrix d_{ij} in a quantitative manner. One is to evaluate d_{ij} using the taxonomic distance of the classification code by using hierarchy built in the JST classification codes (Kitai (2008)). The other is to compute the similarity of co-occurrence patterns of the classification codes by using the fact that up to three codes can be attached to a single article (Chapter 2). The calculated disparity matrix d_{ij} by those two formalizations are shown in Figure C.1. Here $d_{ij} = 1$ represents subject i and j are completely different and $d_{ij} = 0$ indicates the subject i and j are equivalent. By using this disparity d_{ij} (3209×3209 matrix), it is possible to compute disparity-weighted richness and evenness. This is equivalent to the Stirling's general diversity index with $\alpha = 1$ (see Table 2.1).

The correction by disparity has been done in the team-scale analysis as shown in Figure C.2. The observed $\Delta_{00} \sim 10^5$ is equivalent to $R \sim 448$ in this case. The diversity index Δ_{10}, which represents disparity-weighted variety, indicates disparity-correction of Δ_{00} as $\Delta_{01} \sim 9 \times 10^4$, which is approximately equal to $R \sim 425$. So, the effect of disparity is at most 5% correction in R. On the other hand, $\Delta_{01} \sim 0.5$ is equivalent to the Gini-Simpson index $1 - \lambda \sim 1$ and disparity/variety-weighted balance index Δ_{11} becomes $\Delta_{11} \sim 0.45$, which corresponds to $1 - \lambda \sim 0.9$. This correction of 0.1 at Gini-Simpson index is considerably large for when one considers a certain amount of publication. This correction can be converted to the change of 0.3 in the relative abundance of paper p_i, *i.e.* for example, 150 out of 500 papers is actually a similar research topic and thus the distribution should be estimated as more skewed shape.

Correction by the disparity index has no direct influence on the main conclusion of Chapter 5. However, in considering the distribution of resource and research subjects, it is necessary to be conscious of the underlying assumption.



Classification	
A	Science and Technology in General
B	Physics
C	Chemistry
D	Space Science and Earth Science
E	Biological sciences
F	Agriculture, forestry and fisheries
G	Medicine
H	Engineering in general
I	System and Control Engineering
J	Information Engineering
K	Management Engineering
L	Energy Engineering
M	Nuclear Engineering
N	Electrical Engineering
P	Thermal Engineering, Applied thermodynamics
Q	Mechanical Engineering
R	Construction Engineering
S	Environmental Engineering
T	Transport and Traffic Engineering
U	Mining Engineering
W	Metallurgical Engineering
X	Chemical Engineering
Y	Chemical Industry
Z	Miscellaneous Industries

Figure C.1: Taxonomic distance d_{ij}^{taxo} and cosine similarity d_{ij}^{cos}.
(a) The taxonomic distance of the JST classification codes normalized by the maximum value of distance. The taxonomic distance are computed by the use of 5-class-hierarchy structure of the JST classification system. (b) The distance of the JST classification codes based on the cosine similarity of the co-occurrence patterns for each classification codes. Here the cosine similarity between classification code i and j is defined as $1-\sum_k n_{ik}n_{jk}/\sqrt{\sum_k n_{ik}^2}\sqrt{\sum_k n_{jk}^2}$, where n_{ik} represents the number of papers taht both the classification code i and k are atattched.

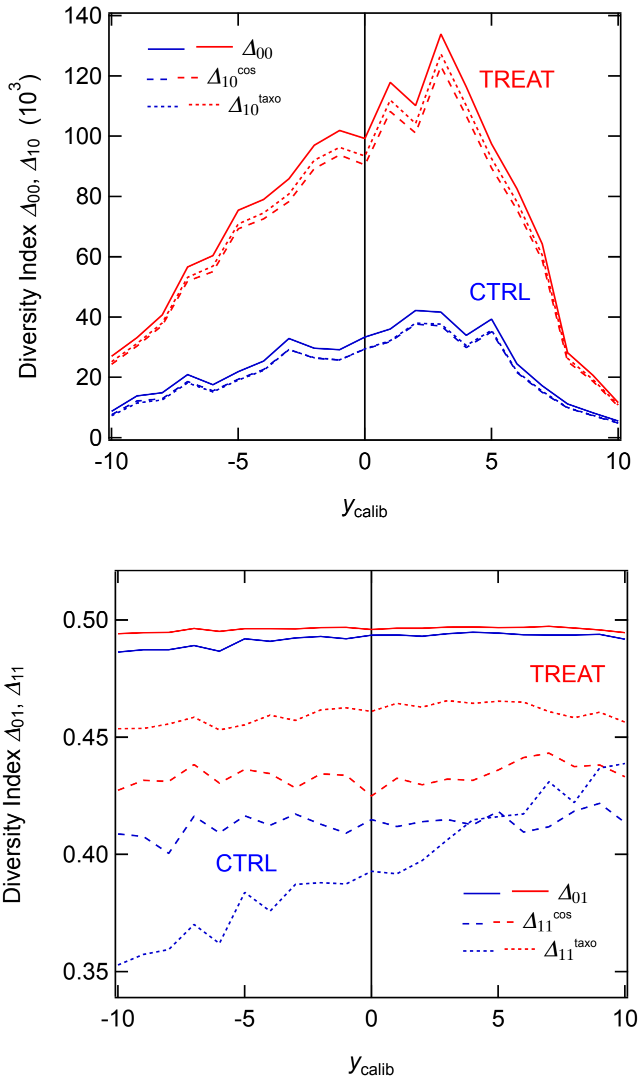


Figure C.2: Comparison between disparity-weighted and non-weighted diversity indices.

The correction by the disparity-weighted calculation for the team-scale analysis. The Stirling's general diversity index Δ_{00}, which is corresponding to richness, is slightly reduced by the corrections by both of the taxsonomic distance ($\Delta_{01}^{\text{taxo}}$) and cosine similarity (Δ_{01}^{cos}) shown in the upper panel. This disparity-correction of Δ_{00} is approximately equivalent to 5% correction in R. The variety-weighted balance index Δ_{10}, corresponding to Gini-Simpson index, is reduced by the correction as shown in the lower panel. The disparity/variety-weighted balance index computed based on both the taxsonomic distance ($\Delta_{11}^{\text{taxo}}$) and cosine similarity (Δ_{11}^{cos}) suggest that 150 out of 500 papers in those groups may be quite similar from the viewpoint of JST classification code.

References

Alireza Abbasi, Jörn Altmann, and Liaquat Hossain. Identifying the effects of co-authorship networks on the performance of scholars: A correlation and regression analysis of performance measures and social network analysis measures. *Journal of Informetrics*, 5(4):594-607, 2011. ISSN 17511577. doi: 10.1016/j.joi.2011.05.007.

Alison Abbott. Major funder tracks impact. *Nature*, 535: 477-478, 2016.

J. Adams and D. Smith. *Funding research diversity: The impact of further concentration on university research performance and regional research capacity*. Universities UK, 2003.

T. Ahn, A. Charnes, and W. W. Cooper. Using data envelopment analysis to measure the efficiency of not-for-profit organizations: A critical evaluation-comment. *Managerial and Decision Economics*, 9 (3):251-253, 1988. ISSN 10991468. doi: 10.1002/mde.4090090310.

Pedro Albarrán, Juan A. Crespo, Ignacio Ortuño, and Javier Ruiz-Castillo. The skewness of science in 219 sub-fields and a number of aggregates. *Scientometrics*, 88(2): 385-397, 2011. ISSN 01389130. doi: 10.1007/s11192-011-0407-9.

Shuhei Aoki and Megumi Kimura. Allocation of Research Resources and Publication Productivity in Japan: A Growth Accounting Approach. 2014.

Aleksander Aristovnik, Janko Seljak, and Jernej Mencinger. Performance measurement of police forces at the local level: A non-parametric mathematical programming approach. *Expert Systems with Applications*, 41(4 PART 2): 1647-1653, 2014. ISSN 09574174. doi: 10.1016/j.eswa.2013.08.061.

Avkiran. Investigating technical and scale efficiencies of Australian Universities through data envelopment analysis. *Socio-Economic Planning Sciences*, 35:57-80, 2001. ISSN 00380121. doi: 10.1016/S0038-0121(00)00010-0.

Arsev U Aydinoglu, Suzie Allard, and Chad Mitchell. Measuring diversity in disciplinary collaboration in research teams : An ecological perspective. *Research Evaluation*, 25(1):18-36, 2016. doi: 10.1093/reseval/rvv028.

Author R D Banker, A Charnes, and W W Cooper. Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis Some Models For Estimating Technical And Scale Inefficiencies In Data Envelopment Analysis *. *Management Science*, 30(9):1078-1092, 1984. ISSN 0025-1909. doi: 10.1287/mnsc.30.9.1078.

Franz Barjak. Team diversity and research collaboration in life sciences teams: Does a combination of research cultures pay off? *University of Applied Sciences Northwestern Switzerland, Series A: Discussion Paper*, page W02, 2006.

Amrita Bhattacharyya and Shankar Chakraborty. A DEA-TOPSIS-based approach for performance evaluation of Indian technical institutes. *Decision Science Letters*, 3 (3):397-410, 2014. ISSN 19295804. doi: 10.5267/j.dsl.2014.2.003.

Katy Börner. *Atlas of Science: Visualizing What We Know*. The MIT Press, Cambridge, 2010. ISBN 0262014459.

Jeroen Bosman, Ineke van Mourik, Menno Rasch, Eric Sieverts, and Huib Verhoeff. *Scopus reviewed and compared*. Utrecht University Library, Utrecht, 2006.

HJ Braczyk, PN Cooke, and M Heidenreich. *Regional innovation systems: the role of governances in a globalized world*. Routledge, Abingdon, 2004. ISBN 9780415303699.

Vannevar Bush. *Science The Endless Frontier*. United States Government Printing Office, Washington, DC, 1945.

Stephen Carley and Alan L Porter. A forward diversity index. *Scientometrics*, 90:407-427, 2012. doi: 10.1007/s11192-011-0528-1.

Mary Caroline N. Castano and Emilyn C. Cabanda. Performance evaluation of the efficiency of Philippine Private Higher Educational Institutions: application of frontier approaches. *International Transactions in Operational Research*, 14(5):431-444, 2007. ISSN 0969-6016. doi: 10.1111/j.1475-3995.2007.00599.x.

Cristina Chaminade and Monica Plechero. Do Regions Make a Difference? Regional Innovation Systems and Global Innovation Networks in the ICT Industry. *European Planning Studies*, 23(2):215-237, 2014. ISSN 0965-4313. doi: 10.1080/09654313.2013.861806.

A. Charnes, W. W. Cooper, and E. Rhodes. Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6):429-444, 1978. ISSN 03772217. doi: 10.1016/0377-2217(78)90138-8.

Jonathan R Cole. *The Great American University*. Public Affairs, New York, 2009.

Hugo Confraria and Manuel Mira Godinho. The impact of African science: a bibliometric analysis. *Scientometrics*, 102(2):1241-1268, 2014. ISSN 15882861. doi: 10.1007/s11192-014-1463-8.

Hugo Confraria, Manuel Mira Godinho, and Lili Wang. Determinants of citation impact: A comparative analysis of the Global South versus the Global North. *Research*

Policy, 46(1):265-279, 2017. ISSN 00487333. doi: 10.1016/j.respol.2016.11.004.

Wade D. Cook, Kaoru Tone, and Joe Zhu. Data envelopment analysis: Prior to choosing a model. *Omega (United Kingdom)*, 44:1-4, 2014. ISSN 03050483. doi: 10.1016/j.omega.2013.09.004.

William W. Cooper, Lawrence M. Seiford, and Joe (Eds.) Zhu, editors. *Handbook on Data Envelopment Analysis*. Springer, 2011. ISBN 9781441961518.

Robin Cowan. Tortoises and Hares : Choice Among Technologies of Unknown Merit. 101(407):801-814, 1991.

Paul A. David and Geoffrey S. Rothwell. Standardization, diversity and learning: Strategies for the coevolution of technology and industrial capacity. *International Journal of Industrial Organization*, 14(2):181-201, 1996. ISSN 01677187. doi: 10.1016/0167-7187(95)00475-0.

Mercedes Delgado, Michael E. Porter, and Scott Stern. CLUSTERS, CONVERGENCE, AND ECONOMIC PERFORMANCE. *NBER Working Paper*, page 18250, 2012. ISSN 1098-6596. doi: 10.1017/CBO9781107415324.004.

Laércio Dias, Martin Gerlach, Joachim Scharloth, and Eduardo G Altmann. Using text analysis to quantify the similarity and evolution of scientific disciplines. *Royal Society Open Science*, 5:171545, 2018.

Maryann P. Feldman and Dieter F. Kogler. *Stylized facts in the geography of innovation*, volume 1. Elsevier BV, 1 edition, 2010. ISBN 2210-8807. doi: 10.1016/S0169-7218(10)01008-7.

Ichiko Fuyuno. What price will science pay for austerity? *Nature*, 543(7646):S10-S15, 2017. ISSN 14764687. doi: 10.1038/543S10a.

M Gibbons. Science's new social contract with society. *Nature*, 402:C81-C84, 1999.

M Gibbons, C Limoges, H Nowotny, S Schwartzman, P Scott, and M Trow. *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*. Sage, London, 1994. ISBN 0803977948. doi: 10.2307/2076669.

Gernot Grabher and David Stark. Organizing diversity: Evolutionary theory, network analysis and postsocialism. *Regional Studies*, 31(5):533-544, 1997. ISSN 00343404. doi: 10.1080/00343409750132315.

Diana Hicks and J. Sylvan Katz. Equity and Excellence in Research Funding. *Minerva*, 49(2):137-151, 2011. ISSN 00264695. doi: 10.1007/s11024-011-9170-6.

M. O. Hill. Diversity and Evenness: A Unifying Notation and Its Consequences. *Ecology*, 54(2):427-432, 1973. ISSN 00129658. doi: 10.2307/1934352.

Stephen P. Hubbell. *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press, Princeton, 2001. ISBN 0691021287.

Masatsura Igami and Ayaka Saka. Decreasing diversity in Japanese science, evidence from in-depth analyses of science maps. *Scientometrics*, 106(1):383-403, 2016. ISSN 15882861. doi: 10.1007/s11192-015-1648-9.

Haruyuki Irie and Kei Tokita. Species-area relationship for power-law species abundance distribution. *International Journal of Biomathematics*, 5(3):1260014, 2012.

Brian A. Jacob and Lars Lefgren. The impact of research grant funding on scientific productivity. *Journal of Public Economics*, 95(9-10):1168-1177, 2011. ISSN 00472727. doi: 10.1016/j.jpubeco.2011.05.005.

Jane Jacobs. *The Economy of Cities*. Vintage, New York, 1969. ISBN 9780394705842.

"Japan Society for the Promotion of Science". Grants-in-Aid for Scientific Research (KAKENHI). URL https://www.jsps.go.jp/j-grantsinaid/27{_}kdata/.

Jill Johnes. Measuring efficiency: A comparison of multilevel modelling and data envelopment analysis in the context of higher education. *Bulletin of Economic Research*, 58(2):75-104, 2006. ISSN 03073378. doi: 10.1111/j.0307-3378.2006.00238.x.

Lou Jost. Entropy and diversity. *Oikos*, 113(2):363-375, 2006. ISSN 00301299. doi: 10.1111/j.2006.0030-1299.14714.x.

Takashi Kitai. Construction of JICST Scientific Technological Classification 1993. *Journal of Information Processing and Management*, 35(11):967-974, 2008. doi: 10.1241/johokanri.35.967.

Ronald Kostoff. Text mining for science and technology - a review part I - characterization/scientometrics. *Journal of Scientometric Research*, 1(1):11-21, 2012. ISSN 23200057. doi: 10.5530/jscires.2012.1.5.

Thomas S. Kuhn. *The Structure of Scientific Revolutions*. The University of Chicago Press, Chicago, 1970. ISBN 0226458032. doi: 10.1119/1.1969660.

Ludwig Kuntz, Stefan Scholtes, and Antonio Vera. Incorporating efficiency in hospital-capacity planning in Germany. *European Journal of Health Economics*, 8(3): 213-223, 2007. ISSN 16187598. doi: 10.1007/s10198-006-0021-6.

You Na Lee, John P. Walsh, and Jian Wang. Creativity in scientific teams: Unpacking novelty and impact. *Research Policy*, 44(3):684-697, 2015. ISSN 00487333. doi: 10.1016/j.respol.2014.10.007.

Per Levén, Jonny Holmström, and Lars Mathiassen. Managing research and innovation networks: Evidence from a government sponsored cross-industry program. *Research Policy*, 43(1):156-168, 2014. ISSN 00487333. doi: 10.1016/j.respol.2013.08.004. URL http://dx.doi.org/10.1016/j.respol.2013.08.004.

Loet Leydesdorff and Ismael Rafols. Indicators of the interdisciplinarity of journals : Diversity , centrality , and citations. *Journal of Informetrics*, 5:87-100, 2011. ISSN 1751-1577. doi: 10.1016/j.joi.2010.09.002.

Loet Leydesdorff, Stephen Carley, and Ismael Rafols. Global maps of science based on the new Web-of-Science categories. *Scientometrics*, 94:589-593, 2013. doi: 10.1007/s11192-012-0784-8.

Eckhard Limpert, Werner A Stahel, and Markus Abbt. Log-normal Distributions across the Sciences : Keys and Clues. *BioScience*, 51(5):341-352, 2001.

Lund Declaration. Europe must focus on the grand challenges of our time. In *Swedish Presidency Research Conference in Lund. New Times New Solutions*, Lund, 2009.

R H MacArthur and E O Wilson. *The Theory of Island Biogeography*. Princeton University Press, Princeton, 1967. ISBN 9781400881376. doi: 10.2307/1796430.

Franco Malerba. Sectoral systems of innovation and production. *Research Policy*, 31:247-264, 2002.

Mark Matthews and Francis Mcgowan. Reconciling diversity and scale : some questions of method in the assessment of the costs and benefits of European integration. *Revue d'Economie Industrielle*, 59:222-234, 1992.

R. M. May. Will a Large Complex System be Stable? *Nature*, 238:413-414, 1972.

R. M. May. Patterns of species abundance and diversity. In M. L. Cody and J. M. Diamond, editors, *Ecology and Evolution of Communities*, pages 81-120. The Belknap Press, Cambridge, 1975.

R. M. May. Unanswered questions in ecology. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 354(1392):1951-9, 1999. ISSN 0962-8436. doi: 10.1098/rstb.1999.0534.

R. K. Merton. The Normative Structure of Science. In N. Storer, editor, *The Sociology of Sciene: Theoretical and Empirical Investigations*, pages 267-278. The University Chicago Press, Chicago, 1973.

Oliver Mitesser, Michael Heinz, Frank Havemann, and Jochen Gläser. Measuring Diversity of Research by Extracting Latent Themes from Bipartite Networks of Papers and References. In H. Kretschmer and F. Havemann, editors, *Proceedings of WIS 2008, Fourth International Conference on Webometrics, Informetrics and Scientometrics & Ninth COLLNET Meeting*, Berlin, 2008.

A. Mougi and M. Kondoh. Diversity of Interaction Types and Ecological Community Stability. *Science*, 337(6092): 349-351, 2012. ISSN 0036-8075. doi: 10.1126/science.1220529.

Alexis Michel Mugabushaka, Anthi Kyriakou, and Theo Papazoglou. Bibliometric indicators of interdisciplinarity: the potential of the Leinster???Cobbold diversity indices to study disciplinary diversity. *Scientometrics*, 107(2):593-607, 2016. ISSN 15882861. doi: 10.1007/s11192-016-1865-x.

National Institute of Science and Technology Policy. Analytical report for 2014 NISTEP expert survey on Japanese S&T and innovation system (2014 NISTEP TEITEN survey). *NISTEP Report No. 161 (in Japanese)*, 2015.

R Nelson, editor. *National Innovation Systems : A Comparative Analysis*. Oxford University Press, Oxford, 1993. ISBN 9780195076172.

M. E. J. Newman. Power laws, Pareto distributions and Zipf's law. *Contemporary physics*, 46(5):323-351, 2005. ISSN 0010-7514. doi: 10.1016/j.cities.2012.03.001.

OECD. *OECD Science, Technology and Industry Outlook*. OECD, 2016. ISBN 9789264167001. doi: 10.1787/23129638.

Scott E. Page. *Diversity and Complexity*. Princeton University Press, Princeton, 2010. ISBN 0691137676.

R.T. Paine. A Conversation on Refining the Concept of Keystone Species. *Conservation Biology*, 9(4):962-964, 1995. doi: 10.1046/j.1523?1739.1995.09040962.x.

Raj Kumar Pan, Sitabhra Sinha, Kimmo Kaski, and Jari Saramäki. The evolution of interdisciplinarity in physics research. *Scientific Reports*, 2:1-8, 2012. ISSN 2045-2322. doi: 10.1038/srep00551.

Mary E. Power, David Tilman, James A. Estes, Bruce A. Menge, William J. Bond, L. Scott Mills, Gretchen Daily, Juan Carlos Castilla, Jane Lubchenco, and Robert T. Paine. Challenges in the Quest for Keystones. *BioScience*, 46(8):609-620, 1996. ISSN 00063568. doi: 10.2307/1312990.

F. W. Preston. The Canonical Distribution of Commonness and Rarity: Part I. *Ecology*, 43(2):185-215, 1962.

Frank W Preston. Noncanonical Distributions of Commonness and Rarity. *Ecology*, 61(1):88-97, 1980.

Quacquarelli Symonds Limited. QS World University Rankings, 2017. URL https://www.topuniversities.com/.

Filippo Radicchi, Santo Fortunato, and Claudio Castellano. Universality of citation distributions: toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences of the United States of America*, 105(45):17268-72, 2008. ISSN 1091-6490. doi: 10.1073/pnas.0806977105.

I Rafols and M Meyer. Diversity and network coherence as indicators of interdisciplinarity: Case studies in bionanoscience. *Scientometrics*, 82(2):263-287, 2010. ISSN 01389130. doi: 10.1007/s11192-009-0041-y.

S. Render. Random Multiplicative Process: An Elementary Tutorial. *American Journal of Physics*, 58(3):267-273, 1990.

Nathan Rosenberg. Uncertainty and technological change. In G. Wright R. Landau, T. Taylor, editor, *The Mosaic of economic growth*, pages 334-353. Stanford University Press, Stanford, 1996.

Y. Sakagami. JICST science and technology classification. *The Journal of Information Science and Technology Association*, 39(11):497-502, 1989.

M Schmidt, J Glaser, F Havemann, and Michael Heinze. A Methodological Study for Measuring the Diversity of Science. In *Proceedings International Workshop on Webometrics, Informetrics and Scientometrics & Seventh COLLNET Meeting*, Nancy, 2006.

Johan Schot and W Edward Steinmueller. Framing Innovation Policy for Transformative Change : Innovation Policy 3 .0. 2016.

Sotaro Shibayama. Distribution of academic research funds: a case of Japanese national research grant. *Scientometrics*, 88(1):43-60, 2011. ISSN 01389130. doi: 10.1007/s11192-011-0392-z.

Zilla Sinuany-Stern, Abraham Mehrez, and Arieh Barboy. Academic departments efficiency via DEA. *Computers and Operations Research*, 21(5):543-556, 1994. ISSN 03050548. doi: 10.1016/0305-0548(94)90103-1.

P Stephan. The Economics of Science. *Journal of Economic Literature*, 34(3):1199-1235, 1996.

Andrew Stirling. A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society, Interface*, 4(15):707-19, 2007. ISSN 1742-5689. doi: 10.1098/rsif.2007.0213.

Donald E. Stokes. *Pasteur's Quadrant: Basic Science and Technological Innovation*. Brookings Institution Press, Washington, DC, 1997. ISBN 9780815781776.

Atsushi Sunami. Comment: Drawing a new bottom line. *Nature*, 543(7646):S22, 2017. ISSN 14764687. doi: 10.1038/543S22a.

Emmanuel Thanassoulis. Assessing police forces in England and Wales using data envelopment analysis. *European Journal of Operational Research*, 87(3):641-657, 1995. ISSN 03772217. doi: 10.1016/0377-2217(95)00236-7.

The Japan Association of National Universities. The Japan Association of National Universities, 2017. URL http://www.janu.jp/.

The World Bank. World Development Indicators, 2017. URL http://data.worldbank.org/data-catalog/world-development-indicators.

Times Higher Education. World University Rankings, 2017. URL http://www.timeshighereducation.co.uk/.

Nagayasu Toyoda. Study on influence and evaluation on reduction of subsidy for operating expenses of national universities (in Japanese). Technical report, The Japan Association of National Universities, 2015.

Manuel Trajtenberg, Rebecca Henderson, and Adam Jaffe. *University Versus Corporate Patents: A Window On The Basicness Of Invention*, volume 5. 1997. ISBN 1043859970000. doi: 10.1080/10438599700000006.

B Uzzi, S Mukherjee, M Stringer, and B Jones. Atypical combinations and scientific impact. *Science*, 342:468-472, 2013. ISSN 0036-8075. doi: 10.1126/science.1240474.

Richard Van Noorden. Interdisciplinary research by the numbers. *Nature*, 525:306-307, 2015.

Ari Voutilainen and Mari Kangasniemi. Applying the ecological Shannon ’ s diversity index to measure research collaboration based on coauthorship : A pilot study. *Journal of Scientometric Research*, 4:172-7, 2015. doi: 10.4103/2320-0057.174866.

Stephan C. Wagner. Keystone Species. *Nature Education Knowledge*, 3(10):51, 2010. doi: 10.1007/978-3-642-58001-7_11.

Katherine Y. Williams and Charles A. O'Reilly. Demography and diversity in organizations: A review of 40 years of research. *Research in Organizational Behavior*, 20:77-140, 1998.

Oliver E Williamson. Transaction Cost Economics and Organization Theory. *Industrial and Corporate Change*, 2: 107-156, 1993. doi: 10.1093/icc/2.1.107.

Joanna Wolszczak-Derlacz and Aleksandra Parteka. Efficiency of European public higher education institutions : a two-stage multicountry approach. *Scientometrics*, 89: 887-917, 2011. doi: 10.1007/s11192-011-0484-9.