DISCUSSION PAPER No.95

Empirical analysis on factors associated with international scientific collaboration

2013 年 7 月

文部科学省　科学技術・学術政策研究所
第 1 調査研究グループ
加藤　真紀

東北大学大学院情報科学研究科
安藤　朝夫

DISCUSSION PAPER No.95

Empirical analysis on factors associated with international scientific collaboration

Maki KATO and Asao ANDO

July 2013

The 1st Policy-Oriented Research Group
National Institute of Science and Technology Policy (NISTEP)
Ministry of Education, Culture, Sports, Science and Technology (MEXT)
Japan

国際共著に関係する要因の実証分析～Nature & Science と化学論文の分析～

加藤真紀　文部科学省　科学技術・学術政策研究所　第 1 調研究グループ
安藤朝夫　東北大学大学院情報科学研究科

要旨

近年、日本の研究力は低下しており、その原因の1 つとして、質が高い（平均被引用数が多い）とされる国際共著論文数の増加率の低さが指摘されている。しかし、国際共著を行う場合の国の組み合わせに関係する要因や、国際共著論文の質の高さが何によってもたらされるのか、十分に解明されていない。そこで本分析はトムソンロイター社より提供されているWeb of Knowledgeから2種類のデータベースを作成し、これら課題の解明を計量的分析により試みた。この結果、NatureとScienceに発表された論文の分析からは、1) 2ヶ国共に研究開発投資が多く、留学生の交流が多く、EUに加盟している国間での国際共著が多いことや、2) 国際移動先に移動元より多くの研究者がおり、移動先と移動元の公用語の一致やEU加盟国同士の場合に、研究者の国際移動が多いことが示された。化学分野の論文データベースを用いた分析からは、国際共著論文数が多い研究者は、そうでない研究者よりも研究パフォーマンス（論文の数と平均被引用数）が高いことが示された。この傾向は日本、米国、英国、中国で共通である。

Empirical analysis on factors associated with international scientific collaboration

Maki KATO, 1st Policy-Oriented Research Group, National Institute of Science and Technology Policy (NISTEP), MEXT
Asao ANDO, Graduate School of Information Sciences, Tohoku University

Abstract

In recent years the number of academic papers being published on research and development has been increasing more slowly in Japan than in other leading nations. One of the reasons for this may be the low ratio of internationally-collaborated articles in Japan, which tend to have more citations than those written by domestic groups or individuals. However, up until now two major themes have still not been investigated: factors related to international co-authorship and the international mobility of researchers, and the reasons behind the higher citation rates of internationally-collaborated articles. Therefore, this study examines those two themes by empirical analysis using two types of datasets created through use of the Web of Knowledge provided by Thomson Reuters. One of the results gained from analysis on papers published in the past 20 years in *Nature* and *Science* using a count data regression model indicated several factors that have a positive relationship with the production of academic papers: investment on R&D and the number of researchers, the number of international students, and European Union membership. The mobility of international researchers had a positive relationship with the number of researchers at the transfer locations. Analysis using a database of papers in the field of chemistry showed positive correlations between research performance (number of papers and their times cited) and the degree of international collaboration. These tendencies appear to be common in Japan, as well as in the United States, the United Kingdom, and China.

Keywords *International co-authorship, Research collaboration Research performance, Bibliometrics, International mobility, International network*

概要

日本語概要 [1]

背景と目的

近年、日本の研究力は相対的に低下している[2]。その原因の1 つとして、質が高い（平均被引用数が多い）とされる国際共著論文数の増加率の低さが指摘されている[3]。しかし、国際共著を行う際の国の組み合わせに関係する要因や、国際共著論文の質の高さが何によってもたらされるのかは、十分に解明されていない。そこで本研究はトムソンロイター社より提供されているWeb of Knowledge (WoK)から2種類のデータベースを作成し、これら課題の解明を計量的分析により試みた。

データ

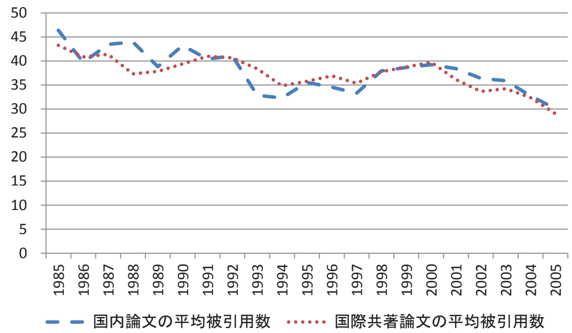
2種類のデータベースのうちの1つは、科学分野のトップクラス学術誌であるNatureとScienceに1989年から2009年までに発表されたarticleで構成され、分析の対象となる論文数は36,208本となった。もう1つのデータベースは、英語論文を重視し国際共著率向上の可能性を残す分野として化学分野を選び、同分野の論文のうちインパクトファクターの高い18雑誌の1985年から2005年までに掲載されたarticleを抽出した。分析対象論文数は188,081本、研究者数は49,599人である。

結果

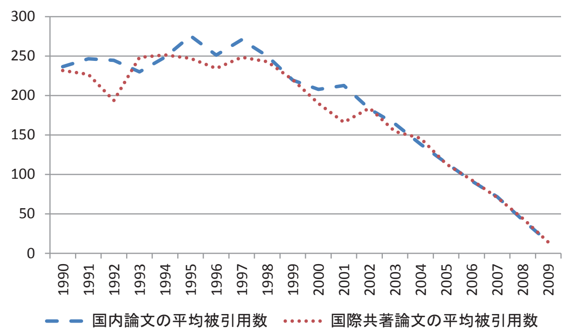
1. 国際共著論文の質（平均被引用数）

・　化学分野の論文で見ると、1ヶ国の研究機関に属する著者によって書かれた論文（国内論文）と国際共著によって書かれた論文（国際共著論文）の平均被引用数は殆ど変わらない。この傾向はNatureとScienceでも同様である（31, 32頁）。

【化学論文】



【Nature & Science】



注意：2つのグラフの対象年次（横軸）と被引用数のスケール（縦軸）は異なっている。

---

[1] 英語論文の概要に「5．論点と今後の課題」他を加筆

[2] 1999年から2001年の日本の国別論文数は米国に次いで2位だったが（世界シェア9.5%）、10年後の2009年から2011年には、米国、中国、ドイツ、英国に続き5位である（同6.6%）。トップ10%論文の割合も4位から7位へと低下している（阪・桑原, 2013）。

[3] 1999年から2001年の日本の国際共著率は18.4%であり、10年間で8.0ポイント増加した。これに対して英国、ドイツ、フランスの国際共著率は1999年から2001年において約35%から40%近くあり、10年後には50%を超えるなど3ヶ国共に2桁の伸びを示している。研究者の国際的な多様性の高い米国においても1999年から2001年の国際共著率は23.6%と日本より高く、10年間で9.8%増加している（同上）。

・ 一方で国による違いも示されている。日本、英国、中国の研究者が筆頭著者である化学論文においては、国際共著論文の平均被引用数の方が国内論文の平均被引用数よりも高い。米国では逆である（44頁）。

2. 国際共著論文を行う国の組み合わせに関係する要因

データ:Nature、Scienceの論文

- 2ヶ国共に研究開発投資が多く、留学生の交流が多く、EUに加盟している国間での国際共著が多い。また情報技術や航空システムの発達にもかかわらず、これまで指摘されたように距離が遠い国間では国際共著は少ない。この傾向は化学論文のデータを用いても同様である（16頁）。
- 国際移動先に移動元より多くの研究者がいる場合や、移動先と移動元の公用語が一致した場合や双方がEUに加盟している場合に、研究者の国際移動が多い。距離は影響しない（17,18頁）。

3. 国際共著論文の質の高さに関係する要因

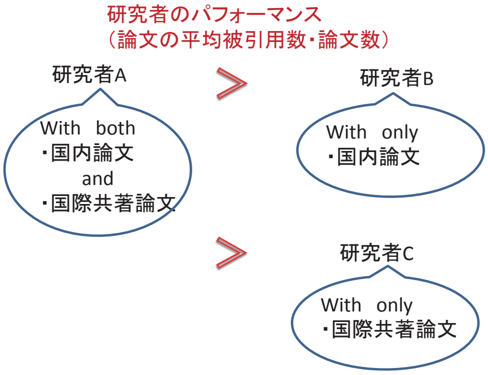
データ:化学論文

- 国際共著論文数が多い研究者は、そうでない研究者よりも研究パフォーマンス（論文数や平均被引用数）が高い（34, 35頁）。例えば研究者単位のパフォーマンスと国際共著論文数や論文率の間の相関係数を見ると、いずれも正かつ1%有意水準の係数が示されている（下表中色付け参照）。

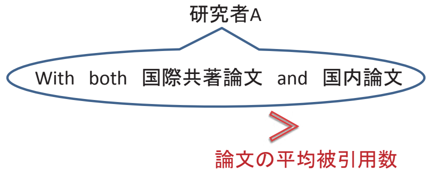
	論文数（整数カウント）	論文数（分数カウント）	平均被引用数（年で調整）	国際共著論文数	国際共著論文率	共著外国数
論文数（整数カウント）	1	0.7696***	0.2246***	0.4170***	0.2553***	0.4745***
論文数（分数カウント）		1	0.1340***	0.2439***	0.0835***	0.3439***
平均被引用数（年で調整）			1	0.1273***	0.0840***	0.1854***
国際共著論文数				1	0.9574***	0.8079***
国際共著論文率					1	0.7549***
共著外国数						1

研究者数: 49,599; 有意水準: ***p <0.01

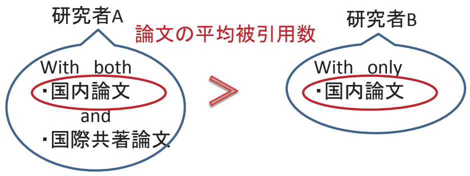
- 国際共著論文と国内論文の両方の論文を発表している研究者は、どちらか片方のみを発表している研究者と比較して、論文数が多く平均被引用数も10%以上高い（33,34頁）。



・ 国際共著論文の被引用数が多いのは、研究者のパフォーマンス以外の要因の影響も考えられる（国際共著論文と国内論文の両方を書いている著者の国際共著論文と国内論文の被引用数を比較したところ、国際共著論文の被引用数が多い）（36頁）。



・ 国際共著論文の発表は、研究者の能力および能力向上と関係する可能性がある（国際共著論文と国内論文の両方を発表している研究者の国内論文の平均被引用数は、国内論文のみを発表している研究者の論文の平均被引用数よりも高い）（36,37頁）。



・ 国際移動をした研究者は、国際移動をしていない研究者よりも論文数が多く、国際共著も多い（33，38頁）。

4. 日本の特徴

データ:化学論文

・ 日本の研究者のうち国際共著論文と国内論文の両方を発表した研究者の平均被引用数は、どちらか片方のみを発表している研究者よりも10%以上高い（下表参照）。このような傾向は米国・英国・中国でも同様である（46,47頁）。

研究者区分	数	割合	平均被引用数（年で調整）
国内論文のみ	3,240	77.70%	0.79
国際共著論文のみ	189	4.53%	0.77
国内論文と国際論文の両方	741	17.77%	0.93
合計	4,170	100.00%	0.82

注意：平均被引用数は年で基準化されているため、1が平均を表し、1.1は平均より10%多いことを意味する。

・ 日本の研究者のうち、国際共著論文をより多く発表した研究者のパフォーマンス（論文数や平均被引用数）は、発表していない研究者よりも高い（下表参照）。このような傾向は米国・英国・中国でも同様である（48,49頁）。

	論文数（整数カウント）	論文数（分数カウント）	平均被引用数（年で調整）	国際共著論文数	国際共著論文率	共著外国数
論文数（整数カウント）	1	0.8775***	0.2820***	0.4014***	0.3203***	0.5113***
論文数（分数カウント）		1	0.2229***	0.3310***	0.2484***	0.4632***
平均被引用数（年で調整）			1	0.1504***	0.1245***	0.2653***
国際共著論文数				1	0.9816***	0.7266***
国際共著論文率					1	0.6961***
共著外国数						1

研究者数: 4,170; 有意水準: ***p <0.01

・ 日本の研究者のうち国際移動をした研究者（2ヶ国以上で論文を発表している研究者）が国際共著と国内論文の両方を発表する割合は7割を超える。一方国際移動をしていない研究者のうち、国内論文のみを発表する割合は8割を超える。このような傾向は米国・英国・中国でも同様である（50頁）。

研究者の経験国数	国内論文のみ		国際共著論文のみ		国内論文と国際論文の両方		合計人数
	数	割合	数	割合	数	割合	
1	3,153	83.08%	174	4.58%	468	12.33%	3,795
2以上	87	23.20%	15	4.00%	273	72.80%	375

5. 論点と今後の課題

・ 化学分野においては、世界的な傾向と同様に、日本でも研究者のパフォーマンスと国際共著論文数は正の関係を持つ。しかし日本の国際共著率が低い背景を理解するためには、日本の研究者がどのように国際共著に関わるのか、研究資金や研究環境、パフォーマンス等の研究者個人の特性に着目した調査分析が求められる。
・ 研究者という単位に着目して分析を行った結果、国内論文（1ヶ国の研究機関に属する著者によって書かれた論文）と国際共著論文の両方を発表している研究者のパフォーマンスの高さが示された。同結果については、今後、研究者の属性・キャリアを考慮した分析を行う必要があると考えられる。

参考文献

阪彩香・桑原輝隆.(2013).『科学研究のベンチマーキング2012～論文分析でみる世界の研究活動の変化と日本の状況～』.文部科学省科学技術政策研究所 調査資料218

Contents

本文

1. International scientific collaboration and mobility of researchers: Count data empirical analysis of papers in Nature and Science

1.1. Introduction

Research is increasingly being carried out in teams across all fields, and collaborated works produce more highly cited results than works by individuals do (Wuchty et al. 2007). The reasons for scientific collaboration are varied; for instance, Beaver (2001) listed 18 reasons including access to expertise or equipment, enhanced efficiency and productivity of research, and student education. Collaboration also offers advantages such as a higher rate of acceptance of publication in leading journals because of a higher degree of competence or credit than in the case of a single author (Katz and Martin 1997). Therefore, scientific collaboration seems to be a good way for researchers to produce scientific findings and be recognized by peers.

Scientific collaboration does not always result in publication, and co-authorship of a paper does not necessarily mean real collaboration (Laudel 2002). However, there are four advantages in using co-authored papers as indicators of scientific collaboration: verifiability, stability, ease of measurement, and data availability (Katz and Martin 1997). This study also regards co-authorship as the best documented indicator of scientific collaboration currently available.

Research output is a political concern because public resources are used especially for basic research; stakeholders might be interested in the efficiency and effectiveness of such funding (Schmoch and Schubert 2008). It is worth political stakeholders paying attention to the quantity and quality of internationally co-authored papers published by teams. In terms of quantity, single-institutional co-authored papers accounted for approximately 80% of papers in the 1980s, but only 44.1% of papers in 2010, at which time internationally co-authored papers accounted for 21.6% of papers (NISTEP 2011). In terms of quality, internationally co-authored papers are more highly cited than domestic papers (Glänzel 2001; Glänzel and Shubert 2001). According to Wagner (2008), scientists in all fields are interested in networks beyond national boundaries simply to be exposed to new ideas.

Previous studies have investigated various aspects of internationally co-authored papers. Glänzel (2001) showed that the rate, number, and quality of internationally co-authored papers vary among countries and disciplines. For instance, a country with a smaller population has a higher rate of international co-authorship (Kato and Chayama 2010). Patterns and characteristics of internationally co-authored networks in particular regions or countries have also been analyzed (Ding 2011; Cardillo 2006; Leydesdorff 2008).

The selection of partners in scientific collaboration may reflect an individual researcher's interests even in international collaboration. When scientific collaboration crosses national borders, however, there may be macro-scale factors affecting the selection of collaborative partners. For instance, the pattern of international collaboration in chemistry is characterized by geopolitical, historical, and linguistic factors (Zitt et al. 2000, Glänzel and Schubert 2001).

A few empirical studies have analyzed factors of international co-authorship. For example, Nagpaul (2003) showed that three proximity measures—geographical, thematic, and socio-economic distances—have negative effects on international collaboration in an analysis of networks involving the 45 most scientifically advanced countries. Conversely, Choi (2012) found that the formation of international co-authorship network structures among 30 Organisation for Economic Co-operation and Development (OECD) member countries between 1995 and 2010 was not affected by geographical, linguistic, and economic affinities but by broadband Internet access and the international mobility of students. These studies used a quadratic assignment procedure to avoid autocorrelation in their data.

Building on the findings of previous studies, the present study examines the characteristics of pairs of countries for which there is international collaboration by considering factors related to international co-authorship. There are two new aspects to the study. The first is that we use a count data model for empirical analysis. Summarizing findings in papers published in 1978, Beaver (2001) pointed out that collaborative authorship follows a Poisson distribution and gradually tends to change towards a negative binominal distribution as collaboration becomes more frequent; these two distribution patterns are typical patterns of the count data model. From the perspective of the distribution pattern and characteristics of data, the number of papers as an independent variable can be regarded as a count variable. If count variables are analyzed using a continuous model, estimates are inefficient, inconsistent, and even biased (Long 1997). Although the use of regression models for count variables is relatively recent, we attempt to use them in this work.

The second aspect new to the study is that we use variables directly related to research and development (R&D) activities. Considering that developing countries such as China, Brazil, Turkey, and Iran have been increasing their contributions to knowledge production in the past 10 years (Saka and Kuwahara 2011), more countries need to be included in the analysis. If more developing countries are included, large heterogeneity in each developing country should be recognized. Therefore, it would seem to be appropriate to use variables directly related to R&D such as expenditure on R&D to express the proximity of the scientific levels of two countries, instead of using gross domestic product (GDP) per capita expressing the economic level. Another factor that can be considered is the international economic or political community, which encourages scientific collaboration through funding such as the Framework Program of the European Union (EU), and this factor is included as an independent variable to determine its impact on international collaboration.

Before analyzing the factors relating to international collaboration, we examine factors relating to national knowledge production because only a rather simple model is needed to understand the input and output of research activities. In principle, the national production of scientific knowledge as an output can be expressed by expenditure on R&D and the number of researchers as inputs. We add research collaboration and competition as indicators of enhanced research activity. First, we include exchange among researchers as a part of collaboration. We assume that if researchers exchange ideas more then they will form more new ideas and obtain more scientific results. In this case, greater diversity among groups generates more creativity. Although researchers can differ in many ways such as whether they belong to schools or the wider community, the present study focuses only on cultural differences using country-to-country exchange as an indicator. Competition among researchers also affects the production of knowledge; a greater number of completed doctorates intensifies competition among researchers and increases the publication of papers [4].

Another possible factor is language. English is the common language in the scientific community today, as papers are usually written in English. Therefore, if researchers are in a country where English is an official language, they have advantages in publishing papers internationally compared with colleagues who are in a country where English is not an official language but there is the possibility to publish papers in their mother tongue. However, prestigious journals distributed internationally are available only in English nowadays, and the official language of a country might not matter in the publishing of high-quality scientific papers. Building on the results of the literature survey stated above, we investigate two research questions using count data models.

Research Question 1: Are the following factors of the national production of knowledge: expenditure on R&D, number of researchers, collaboration or competition among researchers, and English being an official language?

Research Question 2: Are the following factors of proximity related to international collaboration: expenditure on R&D, number of researchers, and number of international students, proximity of languages and geographical distances, and membership in the EU?

International scientific collaboration could be related to the international mobility of researchers because face-to-face interaction is still necessary at least to initiate collaboration (Wagner 2006). Meanwhile, many studies have investigated the international mobility of

[4] For example, young researchers write more articles to obtain tenure, while researchers who already have tenure contribute to the acquisition of competitive funding. This way, competition will occur among members of the same group rather than between different groups (i.e., existing researchers vs. young researchers) and these groups interact with each other at the same time.

professionals in the context of brain drain or brain circulation, which typically involve the flow of researchers from developing to developed countries (Bhagwati 2009; Thorn and Holm-Nielsen 2008). However, owing to the limited availability of macroscopic data, there have been few studies that have analyzed the international mobility of researchers empirically.

One such work is that of Franzoni et al. (2012), who collected data from 17,182 researchers through a Web-based survey and found considerable variations in immigration and emigration patterns of researchers across countries. The same tendency was reported by by Ioannidis (2004), who examined the curriculum vitae of researchers whose papers were most cited in 1981–1999. BIS (2011) found that internationally mobile researchers performed better than less-mobile researchers in the United Kingdom using data from Scopus.

The reason why researchers are internationally mobile remains to be confirmed empirically. For instance, Franzoni et al. (2012) stated that researchers were internationally mobile mainly for professional reasons in terms of emigration, and had personal or family reasons for returning to their home countries. OECD (2008) attributed the international mobility of researchers to the presence of advanced research infrastructure and leading researchers at the destination, in addition to economic incentives. Although the selection of countries to travel to may reflect an individual researcher's interests, an individual scientific network, or family matters (Ackers and Gill 2008) to some extent, there are also macro factors. These factors might be similar to those for international collaboration between two countries because an exchange of researchers results in collaboration. Therefore, the present study will investigate factors of the international mobility of researchers by asking the following research question.

Research Question 3: Do the following factors of proximity affect the international mobility of researchers: proximity of the R&D environment such as expenditure or the number of researchers, international mobility of students, languages or geographical distances, and membership in the EU?

1.2. Models

Model of the national production of papers

The first model, which explains the national production of knowledge, builds on a simple Cobb–Douglas production function and includes indicators regarded as pertinent to the national production of papers and stated in the previous section. Although the amount of research funding or number of researchers in a country depends on the state of the academic fields, the present study does not explicitly take this effect into consideration in the models. Therefore, the model of the national production of knowledge is expressed as the following log-linear model.

$$\ln Paper_{(i,t)} = \alpha_0 + \alpha_1 \ln RFund_{(i,t)} + \alpha_2 lnRPeople_{(i,t)} + \alpha_3 \ln RExchange_{(i,t)} + \alpha_4 \ln RComp_{(i,t)} + \alpha_5 Eng_{(i)} + \varepsilon_{(i,t)}. \cdot\cdot\cdot (1)$$

The variables in the above equation are defined as follows.

$Paper_{(i,t)}$: Number of papers published in international journals by Country *i* in period *t*

$RFund_{(i,t)}$: R&D expenditure of Country *i* in period *t*

$RPeople_{(i,t)}$: Number of researchers in Country *i* in period *t*

$RExchange_{(i,t)}$: Number of foreign students in tertiary education in Country *i* in period *t* (Exchange among researchers)

$RComp_{(i,t)}$: Ratio of doctoral students to researchers in Country *i* in period *t* (Competition among researchers)

$Eng_{(i)}$: Dummy variable that takes the value 1 if English is an official language of Country *i* , and zero otherwise

$\varepsilon_{(i,t)} \sim IN(0, \partial^2)$

The signs of coefficients of all independent variables are expected to be positive.

Model of internationally co-authored papers

When scientific papers are considered an asset, international scientific collaboration can be regarded as an international activity for the production of knowledge. Tinbergen et al. (1962) proposed a gravity model that is often applied to the analysis of international trade; i.e., an international transaction of assets. Therefore, the model of internationally co-authored papers can be based on Tinbergen's gravity model. In Tinbergen's gravity model, the volume of trade between two countries is proportional to the economic size of those countries and inversely proportional to the distance between countries. Empirical analyses have been performed with this model adding variables such as linguistic relations between two countries, whether they border each other, their historical background (e.g., whether one was a colony of the other), and trade agreements [5].

In the case of gravity model for exports, the model might be expressed as follows, where exports from Country *i* to Country *j* is denoted $E_{(i,j)}$, the GDP of Country *i* is denoted $Y_{(i)}$, the GDP of Country *j* is denoted $Y_{(j)}$, and the distance between the two countries is denoted $D_{(i,j)}$.

$$E_{(i,j)} = \mu \frac{Y_{(i)}^{\rho} Y_{(j)}^{\sigma}}{D_{(i,j)}^{\tau}}$$

---

[5] Using a model based on Tinbergen's gravity model, for example, Maggioni and Uberti (2009) analyzed factors explaining the formation of bilateral knowledge networks in Europe. Funatsu (2008) carried out modeling and empirical analyses of the bilateral mobility of students.

When analyzing data, Tinbergen took the logarithm of both sides of the above equation. However, in an analysis of the number of internationally co-authored papers, it is inappropriate to directly apply an extended formula of the gravity model because there is no direction of movement in terms of imports or exports. Since countries i and j are interchangeable, the coefficients of ρ, σ cannot be distinguished. This makes it difficult to interpret their meaning. Regarding international scientific activities, it is also assumed that collaboration between countries with more resources has a synergetic effect for scientific findings. With these two reasons, the two variables of expenditure on R&D and the numbers of researchers are treated as a product of the countries. Likewise, if two countries have a common official language, internationally joint research is presumably frictionless between them. EU membership is another potential factor owing to a regional research structure providing competitive funding to encourage international collaboration.

$$\begin{aligned}\ln ICPaper_{(i,j,t)} &= \beta_0 \\ &+ \beta_1 \ln(RFund_{(i,t)} \cdot RFund_{(j,t)}) + \beta_2 \ln(RPeople_{(i,t)} \cdot R\mathrm{People}_{(j,t)}) \\ &+ \beta_3 \ln IStud_{(i,j,t)} + \beta_4 Lang_{(i,j)} - \beta_5 \ln Dist_{(i,j)} + \beta_6 EU_{(i,j,t-1)} + \varepsilon_{(i,j,t)}. \cdots (2)\end{aligned}$$

The variables in the above equation are defined as follows[6].

$ICPaper_{(i,j,t)}$: Number of papers co-authored by Country i and Country j and published in international journals in period t

$IStud_{(i,j,t)}$: Total number of foreign students who are from Country i and study in Country j and students who are from Country j and study in Country i in period t
IStud is treated as a total number of foreign students originated from the other country instead of including directions between two countries.

$Lang_{(i,j)}$: Dummy variable that takes the value 1 if there is an official language common to Country i and Country j, and zero otherwise

$Dist_{(i,j)}$: Geographical distance between the capital cities of Country i and Country j

$EU_{(i,j,t)}$: Dummy variable that takes the value 1 if Country i and Country j are both members of the EU in period t, and zero otherwise

The coefficients of all independent variables including geographical distance between two capital cities are expected to be positive. Geographical distance should have a positive coefficient

[6] In the case of variable $ICPaper_{(i,j,t)}$, it is same as $ICPaper_{(j,i,t)}$ because it means the number of internationally co-authored papers between Country i and Country j. $IStud_{(i,j,t)}$ and $EU_{(i,j,t)}$ in model(2) as well as $\ln Rmobile_{(i,j,t)}$ in model(3) are also similar.

because of the model structure, but it indicates a negative relationship.

Model of the international mobility of researchers

The model of the international mobility of researchers is also based on Tinbergen's gravity model. Although the gravity model is expressed in terms of the GDP and the distance between two countries in the example of exports between two countries, the model of the international mobility of researchers (Model 3) has independent variables related to R&D activities such as expenditure on R&D instead of GDP. If internationally collaborated research depends on the relative conditions of the paired countries, a single variable cannot express them appropriately. Therefore, variables are modified to take the greater of either the surplus of destinations or zero. A linguistic concurrence and EU membership are also added because they might lower the barrier to international mobility, as described earlier.

$$\ln Rmobile_{(i,j,t)} = \gamma_0 + \gamma_1 \ln \max\{RFund_{(j,t)} - RFund_{(i,t)}, 0\} + \gamma_2 \ln \max\{\mathrm{RPeople}_{(j,t)} - RPeople_{(i,t)}, 0\}$$
$$+ \gamma_3 Lang_{(i,j)} - \gamma_4 \ln Dist_{(i,j)} + \gamma_5 EU_{(i,j,t)} + \varepsilon_{(i,j,t).} \cdot\cdot\cdot (3)$$

The variable in the above equation is defined as follows.

$Rmobile_{(i,j,t)}$: Number of researchers who published papers in *Nature* or *Science* during the target period and had moved from Country *i* to Country *j* in period *t*

The coefficients of all independent variables including geographical distance between two capital cities are expected to be positive.

1.3. Data

Data sources and description

Data sources of the variables are given in Table 1. The present study obtained original data for dependent variables in three models from Web of Science (WoS) provided by Thomson Reuters in June 2010 focusing only on articles published in either *Nature* or *Science*[7]. These two journals are renowned in natural science and are multidisciplinary publications, and the quality of their articles is thus controlled to some extent. The target data period was from 1989 to 2009. During this period, 19,330 articles were published in *Nature* and 18,200 in *Science*, totaling 37,530. When 1322 articles without a country name were excluded, the resulting number of articles in the target data was 36,208.

---

[7] WoS contains data on natural science papers published in approximately 10,000 major peer-reviewed scientific journals. These journals are selected by criteria including the number of citations the journal has received, internationality (availability of abstracts in English), and periodical publication.

Target periods and numbers of countries vary among models (Table2) because some countries and areas that seldom appeared were excluded. Specifically, for internationally co-authored papers, countries that had 10 or more papers published from 1989 to 2009 according to integer counting were considered[8]; for the model of the international mobility of researchers, countries that were involved in the movement of researchers at least once during the same period were covered. Furthermore, some countries for which there is a lack of data for major variables were excluded from the analysis.

Table 1 Sources of variables and their description

Type of variable	Variable name	Description	Source
Dependent	*Paper*	Number of articles published in *Nature* and *Science* for each country	Compiled from Thomson Reuters' data
	ICPaper	Number of internationally co-authored articles published in *Nature* and *Science* (total for two countries)	
	Rmobile	Number of researchers who have published two or more articles in *Nature* or *Science* and whose institution was located in a different country than that for at least one of those publications.	
Independent variables	*RFund*	R&D expenditure as a percentage of GDP	WDI 2010
	RPeople	Number of researchers engaged in R&D (full time basis)	WDI 2010
	RExchange	Number of foreign students accepted for higher education (ISCED 5/6: undergraduate and graduate education)	UIS
	RComp	Ratio of doctoral students to researchers	Number of students in doctoral courses (UIS) Number of researchers (WDI 2010)
	Eng	A dummy in which the official language is English (1: English, 0: Otherwise)	CEPII data: compiled from geo-cepii
	IStud	Number of foreign students between two countries	Compiled from UIS
	Lang	A dummy in which two countries use a common official language (1: Same official language, 0: Otherwise)	CEPII data: geo-cepii
	Dist	Distance between the capitals of two countries	CEPII data: dist-cepii
	EU	A dummy in which two countries are both EU members (1: two countries are both EU members, 0: Otherwise)	Ministry of Foreign Affairs' Japan website

Note: WDI stands for World Development Indicators. UIS stands for UNESCO Institute for Statistics, and CEPII stands for Centre d'Etudes Prospectives et d'Informations Internationales

[8] "Integer counting" adds up the number of papers irrespective of the co-authors' country affiliation. Details will be discussed in the next section.

Table2. Target period and the number of target countries for the three models

Model name	Period	Time points	Number of countries	Number of data	Number of 0 on data
Model of national production of articles	1985～2008	24	146	3,504	1578 (Number of paper)
Model of international co-authorship of articles	1999～2008	10	64	20,160	15,968 (Number of paper)
Model of international mobility of researchers	1989～2009	1	64	4,032	3,880 (Number of researchers)

Methods of counting papers

There are three ways in which credit can be shared among authors of a paper: all coauthors are given full credit as the most common approach (integer counting), credit is shared equally among authors (fractional counting), and credit is shared according to the order that authors are listed (Tol 2011). Each method has its advantages and disadvantages because there is no detailed information about the author contributions. Because the objective of the study is to observe the international network, we use the first method of integer counting, which gives all countries full-credit without considering the number of authors affiliated to each country.

Integer counting represents involvement in global research activities, whereas fractional counting represents the contribution to knowledge creation. When integer counting is employed, countries frequently involved in international co-authorship receive more counts, and this does not necessarily translate into a measure of the creation of knowledge. Therefore, the present study uses independent variables counted by fractional counting as alternatives in the national knowledge production model to determine the difference between counting methods.

Previous studies including Choi (2012) and Glänzel (2001) have used Salton's measure as an indicator of international collaboration strength[9]. The measure is defined as the number of joint publications divided by the square root of the product of the total number of publications for each country (Glänzel and Schuberet 2004). Because this study uses the number of papers as count data, it seems appropriate to use the actual count instead of a portion of the collaboration; however, we use Salton's measure to determine the appropriateness for estimation of the count data model.

General Analysis

Before answering the specific research questions, we provide a descriptive analysis of the data. The number of countries participating per paper is given in Table 3. There are 24,797 (68.5%) domestic papers and 11,411 (31.5%) international papers. Among international papers,

[9] It is shown by ($\frac{P_{i,j}}{\sqrt{P_i \cdot P_j}}$), when $P_{i,j}$ is the number of internationally co-authored papers between countries i and j, and P_i or P_i is the number of total papers for the country i or j.

69.6% have authors from two countries and 19.3% have authors from three countries. The classification of country by number of papers is listed in

Table 4. The United States is listed first regardless of the counting method. All countries in the top 10 are high-income and OECD member countries. The ratio of the fractional count to integer count is high in Australia (71.9%), Japan (69.4%), and Canada (69.1%), and low in Switzerland (45.2%), Italy (50.0%), and the Netherlands (52.2%). The former countries contribute more to world knowledge production, and the latter countries contribute more to participate into research activities because of higher international collaboration ratio.

Table 3 Papers classified by numbers of participating countries

Types	Number of participating countries	Number of articles	Ratio of international co-authorship
Domestic	1	24,797	-
International	2	7,945	69.6%
	3	2,205	19.3%
	4	646	5.7%
	5	268	2.3%
	6	121	1.1%
	7	67	0.6%
	8	47	0.4%
	9	32	0.3%
	10 or more	80	0.7%
Total	-	36,208	100.0%

Table 4 Rankings of county by number of papers

Number	Countries	Number of articles(Integer counting)	Number of articles (Fractional counting)	Articles by fractional counting/Articles by integer counting
1	United States	25,597	17,378	67.9%
2	United Kingdom	5,372	2,746	51.1%
3	Germany	3,787	2,379	62.8%
4	France	2,695	1,762	65.4%
5	Japan	2,444	1,696	69.4%
6	Canada	1,942	1,341	69.1%
7	Switzerland	1,381	624	45.2%
8	Netherlands	1,188	620	52.2%
9	Australia	1,049	755	71.9%
10	Italy	1,019	510	50.0%

The data obtained from WoS contained author information including the authors' names, attributed institutions, and countries where they are located, as well as reprint address information. Whether countries participated in internationally co-authored papers was determined from both the authors' and reprinted authors' addresses. International connections were counted according to internationally co-authored papers. Because authors' names uniquely corresponded with countries only after 2008 in the author column, international mobility is limited to the movement of researchers in the reprint address column where the author is uniquely identified.

Descriptive analysis of the author is presented in Table 5[10]. A total of 15,743 authors (75.5% of the total) had only one paper published during the target period. Meanwhile, 20,185 researchers (96.8%) never moved internationally. The overwhelming majority (95.2%) of moves took place between two countries. It should be noted that among researchers who did not move, about 57.9% belonged to institutions in the United States.

Table 5. Breakdown of movement and the number of researchers (left: number of papers published, middle: number of moves, right: number of countries to which researchers moved)

Number. of papers published	Number. of researchers
1	15,743
2	3,011
3	1,032
4	488
5 or more	574

Number. of moves	Number. of researchers
0	20,185
1	552
2	101
3	9
4	1

Number. of countries to which researchers moved to	Number. of researchers
0	20,185
1	631
2	31
3	1

The international relationship of internationally co-authored papers is shown in Figure 1. The lines in the figure present the number of co-authored papers. The figure shows that countries have co-authoring links with each other and that the connections with the United States are the strongest, followed by those with the United Kingdom.

The volume of international flow (total inflow and outflow) of researchers is shown in Figure 2. The thicknesses of the arrows in the figure represent the number of moves by researchers. The figure indicates two-way movement between countries. A remarkable trend is a large volume of flow between developed countries, such as between the United States and the United Kingdom, Germany, and Japan.

[10] The names in the Reprint Address are expressed as the family name and the initial of the first name. This indicates the possibility of mistaking two or more different people as the same person. However, as a result of checking sample names against research fields, it was decided that different people were not treated as the same single person.

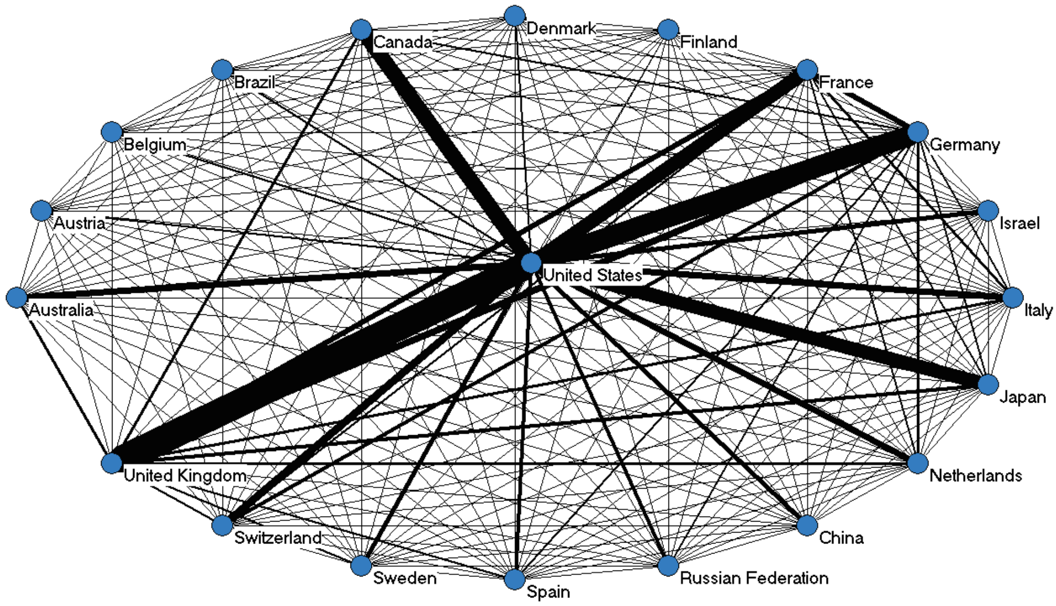
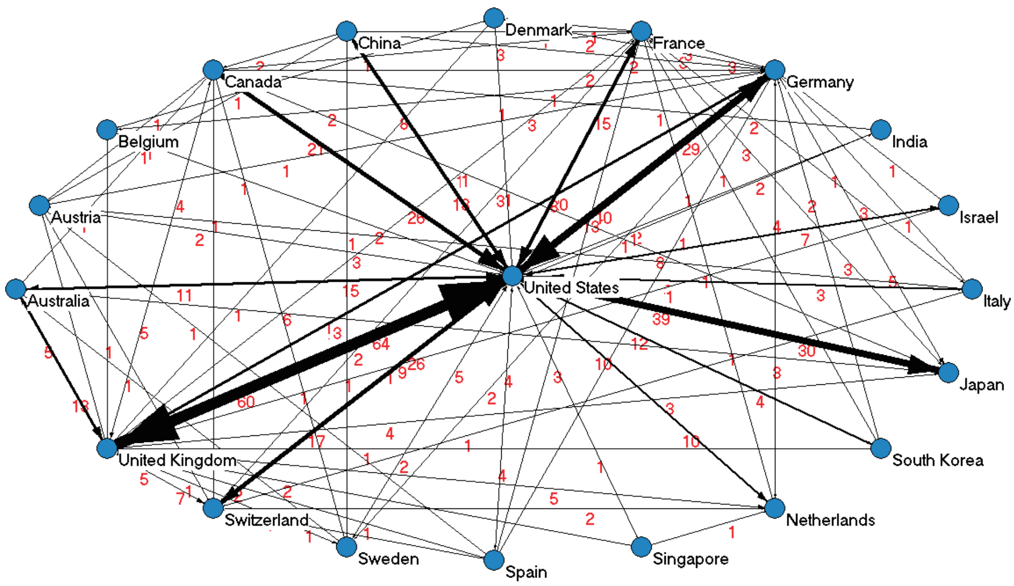


Figure 1. Number of papers co-authored between two countries and published in *Nature* and *Science* (top 20 countries)



Note: When there are two figures along an arrow, closer to the node means move in, and farer means move out.

Figure 2 International mobility of researchers having published papers in *Nature* and *Science* (top 20 countries)

1.4. Estimation methods

There are two types of count data model: one has a maximum value such as a model used in targeting a particular number (N) of personnel (i.e., the maximum number is N) while the other does not have a maximum. A representative distribution model of the former type is a binominal model while that of the latter type is a Poisson model. Considering the characteristics of our data, the Poisson model would seem appropriate. However, a Poisson model can underestimate the amount of dispersion, and a negative binominal model is more appropriate in that case (Long and Freese 2006). To detect over-dispersion and decide which model to use, deviance and Pearson goodness-of-fit statistics are used.

Regarding national or international scientific output, we assume that all countries or all pairs of countries have a possibility of publishing although the number of publications varies. This would not be realistic, however. For instance, some countries might not have a suitable environment in which to conduct quality research, and some pairs of countries may have severed diplomatic relations. In that case, we need to use zero-inflated models. This assumes that there are two groups: one always has zero publishing probability, and the other can have zero and non-zero probability. The formulas of Poisson regression and Zero-inflated Poisson regression are described as follows, as according to StataCorp(2011), when y is defined as a given number of events occurring in a fixed interval time or space Country j ;

In Poisson regression, the log likelihood with weights w is given by

$$\ln L = \sum_{j\in(s\cup\bar{s})}^{n} w_j\{-e^{\varphi_j} + \varphi_j y_j - \ln(y_j!)\} \cdot \cdot \cdot (4)$$

$$\varphi_j = x_j\beta$$

On the other hand, the log likelihood of Zero-inflated Poisson regression with weights w_j is defined by

$$\ln L = \sum_{j\in S} w_j \ln\left[F(\varphi_j^{\gamma}) + \{1 - F(\varphi_j^{\gamma})\}\exp(-\tau_j)\right] +$$

$$\sum_{j\notin S} w_j[\ln\{1 - F(\varphi_j^{\gamma})\} - \tau_j + \varphi_j^{\beta} y_j - \ln(y_j!)] \cdot \cdot \cdot (5)$$

$$\varphi_j^{\beta} = x_j\beta,\ \varphi_j^{\gamma} = z_j\gamma$$

Where F is the inverse of the logit or probit function, and S is the set of observations for which the outcome y_i =0.

To decide whether to use a zero-inflated model, we carry out a Vuong test as commonly done. The Vuong test to consider two models including a non-nested model such as a zero-inflated Poisson model is defined as

$$m_i = ln\left\{\frac{\widehat{P_1}(y_i|x_i)}{\widehat{P_2}(y_i|x_i)}\right\}.$$

$\widehat{P_1}(y_i|x_i)$ is the probability for the first model and $\widehat{P_2}(y_i|x_i)$ is that for the second model. The Vuong statistic to test the hypothesis (E(m) = 0) is

$$V = \frac{\sqrt{N}\bar{m}}{s_m},$$

where $\bar{m}$ is the mean and s_m is the standard deviation of m_i. V has an asymptotic normal distribution. If V > 1.96, the first model is favored; if v < –1.96, the second model is favored. When we carry out a Vuong test, we set the zero-inflated model as the first model and the usual Poisson or negative binominal model as the second model.

1.5. Results

Model of the national production of papers

First, Model (1) was estimated with a Poisson model using the number of papers counted employing either the integer method or fractional method as the dependent variable. The performance of model is evaluated by Pseudo R^2 [11]. The results are presented in Table 6. Deviance and Pearson goodness-of-fit statistics show that over-dispersion does not need to be considered and a negative binominal model does not need to be used (integer model: deviance $\chi^2(296) = 167.9$, not significant (n.s.), Peason $\chi^2(296) = 130.2$, n.s. ; fractional model: deviance $\chi^2(296) = 139.6$, n.s., Peason $\chi^2(296) = 122.0$, n.s.). Regarding (a) in Table 6, the coefficients of variables of expenditure on R&D and the number of foreign students are positive and significant at the 1% level. The level of expenditure on R&D might represent the level of the research environment, and the number of foreign students expresses the level of higher education carrying out basic research as well as the contribution of graduate students to research. The number of researchers is positive and significant at the 10% level only in the model with the dependent variable counted using the integer method. The coefficient of the variable for competition among researchers is not significant. This indicates that competition does not relate to research output in our data. Possible reasons for this are that the way that competition is expressed is not appropriate or the level of the journal targeted is not suitable for competition as a variable. The coefficient of the variable for English as an official language is not significant, which is also interpreted as relating to the level of the targeted journals because there are no alternatives to prestigious distributed journals in other languages. Because the tendency of the results was similar between the two counting methods ((a) and (b) in Table 6) other than a better Pseudo R^2 in fractional counting, we show only the results of integer counting hereafter.

---

[11] McFadden's Adjusted R^2 is expressed like $R^2 = 1 - \frac{lnL(\widehat{M_{Full}}) - K}{lnL(\widehat{M_{Intercept}})}$ when M_{full} is model with predictors, $M_{intercept}$ is model without predictors, and $\hat{L}$ is estimated likelihood.

Publishing a paper usually takes time after funds are allocated by government or after students graduate to become researchers. Therefore, an estimated lag of 1 year or 3 years was added to the independent variables. The result is presented as (c) and (d) in Table 6. The tendency of the result is similar to that of the model without lags apart from competition being significant at the 10% level for the model with lag of 1 year.

Some countries had not published papers in *Nature* or *Science* in our data for almost the whole targeted period and a zero-inflated model was thus also estimated. Here we assume that expenditure on R&D and the number of researchers can be used to express the group that does not publish papers. The result of the Vuong test suggests the use of a zero-inflated model (z = 2.15, P > 0.05). The result shown in (e) in Table 6 is similar to that of the estimation for the integer counting model. The model with fractional counting did not converge in our trial as the zero-inflated model does not to converge in contrast to other counting models (Long and Freese 2006).

Table 6 Estimation results of the model of the national publication of papers

Model	(a)	(b)	(c)	(d)	(e)
Counting Method	Integer count	Fractional count	Integer count		
Estimation method	Poisson regression				zero-inflated Poisson regression
Lag / Independent variable	0		1 year	3 year	0
Expenditure on R&D RFund($RFund_{(i)}$)	0.697*** (3.94)	0.832*** (3.91)	0.730*** (4.22)	0.702*** (3.67)	0.566*** (3.03)
Number of researchers (RPeople	0.163* (1.65)	0.175 (1.44)	0.133 (1.40)	0.071 (0.71)	0.189* (1.91)
Number of foreign students (RExchange)	0.251*** (6.10)	0.327*** (6.54)	0.233*** (5.82)	0.229*** (5.36)	0.258*** (6.23)
Competition among researchers (RComp)	0.076 (1.48)	0.060 (0.97)	0.086* (1.71)	0.075 (1.41)	0.063 (1.21)
English as official language (Eng)	-0.090 (-0.95)	-0.077 (-0.71)	-0.095 (-1.01)	-0.077 (-0.76)	-0.088 (-0.93)
Constant	-3.701*** (-6.22)	-4.975*** (-6.74)	-3.297*** (-5.78)	-2.661*** (-4.53)	-3.805*** (-6.37)
Expenditure on R&D /zero-inflated($RFund_{(i)}$)					-52.089 (-0.07)
Number of researchers /zero-inflated(RPeople					16.256 (0.86)
Constant /zero-inflated					-113.107 (-0.86)
Pseudo R^2 (McFadden's Adjusted R^2)	0.287	0.340	0.274	0.263	(0.260)
Number of observations	302	302	302	232	302 (Number of 0 is 57)

Note: *** 1%, **5%, *10% significance levels; z values are given in parentheses

Model of international co-authored papers

The result of regression for Model (2) is presented in Table 7. Deviance and Pearson goodness-of-fit statistics show that over-dispersion does not need to be considered (deviance: $\chi^2(2679) = 54.6$, n.s., Peason: $\chi^2(2679) = 60.2$, n.s.). The results show that in (a) Table 7, as expected, the coefficients of all dependent variables were positive and significant. The coefficients of expenditure on R&D, the number of foreign students, the distance between capital cities, and membership in the EU for the two countries were positive and significant at the 1% level. The coefficients of both the number of researchers and the concurrence of the official language were positive and significant at the 10% level. The results indicate that a pair of countries with a better research environment such as countries with high income and OECD member countries collaborate and publish more in high-level journals. The result shows a negative relationship between international collaboration between two countries and the distance between capital cities, as has been shown in previous studies. Even today with the development of information and communication technologies, face-to-face meetings appear to be necessary for scientific collaboration, or cultural closeness is preferred; however, this result could change once middle- or lower-income countries increase their R&D, as these countries might prefer high-income and OECD countries to collaborate with rather than neighboring countries with a similar research level. The issue could be access instead of geographical distance owing to better international transportation. Therefore, the effect of distance would need to be explored in more depth in future work to be confirmed.

A model with a dependent variable was estimated using Salton's measure, and the result showed in (b) Table 7 that only the coefficient of the number of foreign students is positive and significant at the 1% level. When the dependent variable with a lag of 1 year or 3 years was used alternatively, all coefficients of variables except the concurrence of official language were positive and significant at the 1% level as shown in (c) and (d) in Table 7.

The combination of two countries in our dataset includes the combination that does not usually publish papers in *Nature* and *Science* in the target period; e.g., a combination of least-developed countries. Therefore, regression was carried out with the zero-inflated Poisson model, taking the number of foreign students as an indicator to categorize a pair of countries into the always-zero group or not-always-zero group. The result is presented in (e) Table 7. The result of the Vuong test does not support the use of the zero-inflated model ($z = 0.83$, n.s.), although the tendency of the sign and significance level of the coefficient are similar to the result estimated by Poisson regression.

Table 7 International co-authored papers

Model	(a)	(b)	(c)	(d)	(e)
Counting Method	integer count	Salton's measure	integer count		
Estimation method	Poisson regression				zero-inflated Poisson regression
Lag / Independent variable	0		1 year	3 year	0
Quadratic Expenditure on R&D ($RFund_{(i)} \cdot RFund_{(j)}$)	0.799*** (12.64)	0.759 (1.61)	0.672*** (11.28)	0.632*** (10.12)	0.801*** (12.66)
Quadratic Number of researchers ($RPeople_{(i)} \cdot \mathrm{RPeople}_{(j)}$)	0.060* (1.79)	-0.051 (-0.19)	0.094*** (2.98)	0.067*** (2.14)	0.058* (1.73)
Number of foreign students ($IStud$)	0.275*** (28.38)	2.694*** (3.06)	0.259*** (28.37)	0.245*** (25.72)	0.272*** (27.30)
Concurrence of official language ($Lang$)	0.108* (1.74)	0.017 (0.03)	0.082 (1.36)	0.040 (0.63)	0.111* (1.80)
Distance between capital cities ($Dist$)	0.149*** (6.95)	-0.034 (-0.23)	0.148*** (7.38)	0.137*** (6.52)	0.148*** (6.94)
Membership in EU (EU)	0.565*** (10.53)	0.390 (1.07)	0.535*** (10.52)	0.463*** (8.51)	0.564*** (10.51)
Constant	-5.266*** (-9.94)	-4.343 (-1.05)	-5.400*** (-10.88)	-4.567*** (-9.26)	-5.206*** (-9.81)
Number of foreign students ($IStud$)/zero-inflated					-32.247 (-0.03)
Constant /zero-inflated					21.449 (0.03)
Pseudo R^2 (McFadden's Adjusted R^2)	0.255	0.055	0.236	0.0558	(0.249)
Number of observations	4,543	2,686	4,543	2,686	4,543(Number of 0 is 2,736)

Note 1: *** 1%, **5%, *10% significance levels; z values are given in parentheses

Note 2: Foreign student is also calculated by Salton's measure when dependent variable is calculated by Salton's measure. Lag is added to variables of expenditure on R&D, number of researchers, and foreign students

Model of the international mobility of researchers

The Model (3) was estimated by Poisson regression. The result is presented in Table 8. Again, deviance and Pearson goodness-of-fit statistics showed that there is no significant over-dispersion (deviance: $\chi^2(1406) = 485.82$, n.s., Peason: $\chi^2(1406) = 1999.14$, n.s.). As shown in (a) Table 8, the coefficients of the number of researchers, concurrence of the official language, and membership in the EU for two countries are positive and significant at the 1% level. Researchers move to countries that have more researchers than their origin countries. Researchers seem to move internationally with a professional reason to work or train in a better R&D environment that has already attracted researchers. At the same time, it seems that the concurrence of the official language and the comfort of living are factors of international mobility. The distance between two countries is not significant, which might be due to a good system for international travel.

Assuming that the choice of country that a researcher chooses as a destination is limited, regression using a zero-inflated Poisson model was carried out using the number of foreign students as an indicator to categorize countries into the always-zero and not-always-zero groups. The result of the Vuong test supports the use of that model ($z = 8.89$, $p < 0.01$). As shown in (b) Table 8, the sign and significance level of the coefficient are similar to those obtained without the zero-inflated model.

Table 8 International mobility of researchers

Model	(a)	(b)
Estimation method / Independent variable	Poisson regression	zero-inflated Poisson regression
Expenditure on R&D (expenditure if the one in destination is more than that of originally from) ($\max\{RFund_{(j,t)} - RFund_{(i,t)}, 0\}$)	-0.198 (-1.42)	-0.144 (-0.69)
Number of researchers (number if the one in destination is more than that of originally from) ($\max\{RPeople_{(j,t)} - RPeople_{(i,t)}, 0\}$)	0.152*** (3.63)	0.351*** (6.19)
Concurrence of official language ($Lang$)	1.779*** (7.48)	1.627*** (5.87)
Distance between capital cities ($Dist$)	-0.182 (-1.56)	-0.073 (-0.60)
Membership in EU (EU)	1.130*** (3.79)	1.027*** (3.02)
Constant	-1.653 (-1.51)	-2.916** (-2.18)
N Number of researchers (number if the one in destination is more than that of originally from) /zero-inflated ($\max\{RPeople_{(j,t)} - RPeople_{(i,t)}, 0\}$)		0.836** (2.30)
Constant /zero-inflated		-5.266* (-1.81)
Pseudo R^2 (McFadden's Adjusted R^2)	0.117	(0.078)
Number of observations	1412	1412 (Number of 0 is 1340)

Note: *** 1%, **5% *10% significance levels; z values are given in parentheses

1.6. Conclusion and discussion

The number of scientific papers as a form of knowledge production has been increasing in recent years. Internationally co-authored papers have become increasingly common and are more cited than domestic papers. This study performed empirical analysis to examine factors relating to the national production of papers, international co-authorship of papers, and international mobility of researchers using data on papers published in the past 20 years in *Nature* and *Science*. A count data model was applied considering the distribution pattern and characteristics of the data. Dependent variables directly related to R&D activities were also used.

The estimated results for the national production of knowledge have a positive relationship with expenditure on R&D and the number of foreign students. On an international scale, greater spending on research means greater outputs in high-level journals. The positive relationship between the national performance in R&D and the number of foreign students indicates that the international attractiveness of higher education includes the research reputation and that an environment affirming cultural diversity is suitable for scientifically creative activities. R&D activities seem to be enhanced through collaboration and competition but competition among researchers was not a factor in this work.

In the estimation of internationally co-authored papers, the following variables have positive relationships: expenditure on R&D, number of researchers, number of foreign students, membership in the EU for the two countries, and concurrence of the official language. The result indicates that greater research resources in each country of a pair results in research collaboration. In this sense, countries with already established performance in R&D collaborate more. The result is rational if we consider that the objective of researchers is to make scientific findings and be recognized by peers. To enhance collaboration elsewhere, such as collaboration with or among developing countries, science policy could encourage researchers to engage in varied networks, even though research network is generated by researchers themselves and not on a country or systematic level as Wagner (2008) pointed out. The results show that the geographical distance between capital cities has a negative relationship with international collaboration; however, this could change if middle or lower-income countries increase their R&D activities and prefer high-income and OECD membership countries to collaborate with rather than neighboring countries having a similar R&D level.

The international mobility of the researcher has a positive relationship with the number of researchers in the destination country, concurrence of the official language, and membership in the EU for the two countries. Researchers move internationally towards countries with more researchers than the origin country instead of avoiding competition; this seems like the Matthew effect in action. Researchers might consider the comfort of living because the concurrence of the official language has a positive relationship with mobility. Because EU membership also has a positive relationship with international mobility as well as international collaboration, a regional framework is effective in encouraging international collaboration and the mobility of researchers who publish papers in leading journals.

This study employed count data regression as an estimation method. The obtained results appear acceptable, and the methodology does not seem to have drawbacks from our analysis. Therefore, the count data model should be used in the analysis of international research collaboration. Future research topics should include the expansion of standards and areas of papers as well as analysis of the researchers' curriculum vitae to confirm the international mobility more precisely.

References

Ackers, L., and Gill, B. (2008). *Moving people and knowledge: Scientific mobility in an enlarging European Union*. Cheltenham: Edward Elgar.

Beaver, D.D. (2001). Reflections on scientific collaboration (and its study): past, present, and future. *Scientometrics*, 52 (3), 365–377.

Bhagwati, J. (2009). Overview of issues. In Bhagwati, J., & Hanson, G (Eds.), *Skilled immigration today: Prospect, problems, and Policies* (pp.3-11). New York: Oxford University Press.

BIS. (2011). *International comparative performance of the UK research base – 2011.* http://www.bis.gov.uk/assets/biscore/science/docs/i/11-p123-international-comparative-performance-uk-research-base-2011. Accessed 28 August 2012.

Cardillo, A., Scellato, S., & Latora, V. (2006). A topological analysis of scientific coauthorship networks. *Physica A: Statistical mechanics and its applications*, 372 (2), 333–339.

Choi, S. (2012). Core-periphery, new clusters, or rising stars?: International scientific collaboration among 'advanced' countries in the era of globalization. *Scientometrics,* 90(1), 25–41.

Ding, Y. (2011). Scientific collaboration and endorsement: Network analysis of coauthorship and citation networks. *Journal of Informetrics,* 5(1), 187–203.

Franzoni, C., Scellato, G., & Stephan, P. (2012). Foreign born scientists: mobility patterns for sixteen countries. *NBER working paper series*. http://www.nber.org/papers/w18067. Accessed 29 August 2012.

Funatsu, H. (2008). International business and human capital mobility. *Shogaku Kenkyu*, 58(4), 37-63.

Glänzel, W. (2001). National characteristics in international scientific co-authorship relations. *Scientometrics,* 51(1), 69–115.

Glänzel, W., & Schubert, A. (2004). Analysing scientific networks through co-authorship. In Moed, H.F., Glänzel, W., & Schmoch, U. (Eds.), *Handbook of qualitative science and technology research: The use of publication and patent statistics in studies on S&T systems* (pp. 277-298). Dordrecht: Kluwer Academic Publishers.

Glänzel, W., & Schubert, A. (2001). Double effort = double impact? A critical view at international co-authorship in chemistry. *Scientometrics*, 50 (2), 199-214.

Ioannidis, J.P.A. (2004). Global estimates of higher-level brain drain and deficit. *The Journal of the Federation of American Societies for Experimental Biology*, 18(9), 936-939.

Kato, M., & Chayama, H. (2010). *Analysis on research activities in developing countries and international networking of researchers*. Research Material 178. NISTEP

Katz, J. S., & Martin, B. R. (1997). What is research collaboration?. *Research Policy*, 26(1), 1-18.

Laudel, G. (2002). Collaboration and reward: What do we measure by co-authorships?. *Research Evaluation*, 11(1), 3–15.

Leydesdorff, L., & Wagner, C. (2008). International collaboration in science and the formation of a core group. *Journal of Informetrics*, 2(4), 317-325.

Long S. J. (1997). *Regression models for categorical and limited dependent variables*. Thousand Oaks: Sage publishers.

Long S. J., & Freese, J. (2006). *Regression models for categorical dependent variables using STATA*. 2nd Edition. Texas: A Stata Press Publication

Maggioni, M. A., & Uberti, T. E. (2009). Knowledge networks across Europe: Which distance matters? . *Annals of Regional Science*, 43(3), 691-720.

Nagpaul, P. S. (2003). Exploring a pseudo-regression model of transnational cooperation in science. *Scientometrics*, 56(3), 403–416.

NISTEP. (2011). *Japanese science and technology indicators 2011*. NISTEP

OECD. (2008). *The global competition for talent: Mobility of the highly skilled. Paris*: OECD Publishing

Saka, A., & Kuwahara, T., (2011). *Benchmarking scientific research 2011: Bibliometric analysis on dynamic alteration of research activity in the world and Japan*. NISTEP

Schmoch, U., & Schubert, T. (2008). Are international co-publication an indicator for quality of

scientific research?. *Scientometrics*, 74(3), 361–377.
StataCorp. (2011). *Stata 12 Base Reference Manual*. College Station, Texas: Stata Press.
Thorn, K., & Holm-Nielsen, L.B. (2008). International mobility of researchers and scientists: Policy options for turning a drain into a gain. In Solimano, A (Ed.), *The international mobility of talent: Types, causes, and development impact* (pp.145-167). New York: Oxford University Press.
Tinbergen, J. (1962). Shaping the world economy: Suggestions for an international economic policy, *Twentieth century fund*, New York: 264-265.
Tol, R.S.J. (2011). Credit where credit's due: Accounting for co-authorship in citation counts. *Scientometrics,* 89(1), 291-299.
Wagner, C.S. (2008). *The new invisible college: Science for development*. Washington: Brookings Institution Press.
Wuchty, S., Benjamin, J. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science,* 316 (5827), 1036-1039.
Zitt, M., Bassecoulard, E., & Okubo, Y. (2000). Shadows of the past in international cooperation: Collaboration profiles of the top five producers of science. *Scientometrics*, 47(3), 627-657.

Appendix

Because *Nature* and *Science* are top-level scientific journals, we needed to corroborate the validity of the results using more typical journals. Therefore, we estimated Model (2), the core of the regression analysis, using the chemistry paper dataset described later in section 2.2

The result of regression analysis for Model (2) in 1.2 is presented in Table 9. Deviance and Pearson's Goodness of Fit statistics show that over-dispersion does not need to be considered (deviance: χ^2 (2679) = 54.6, n.s., Pearson: χ^2 (2679) = 60.2, n.s.). The results show that in (a) Table 9, as expected, the coefficients of all dependent variables were positive and significant. The coefficients of expenditure on R&D, the number of foreign students, the distance between capital cities, and membership in the EU for the two countries were positive and significant at the 1% level. The coefficients of both the number of researchers and the concurrence of the official language were positive and significant at the 10% level. The results indicate that pairs of countries with a good research environment, such as countries with high income, or OECD countries collaborate and publish more papers in high-level journals. The results also show a negative relationship between international collaboration between two countries and the distance between capital cities, as has been shown in previous studies. Even today with the development of information and communication technologies, face-to-face meetings appear to be necessary for scientific collaboration, and cultural closeness is beneficial; however, this could change once middle- or lower-income countries increase their R&D, as these countries might prefer to collaborate with high-income and OECD countries rather than neighboring countries with a similar level of research. Access due to better international transportation, rather than geographical distance could be the issue. Therefore, the effect of geographical distance needs to be explored in more depth and confirmed in future work.

A model with a dependent variable was estimated using Salton's measure, and the result showed in (b) Table 9 that only the coefficient of the number of foreign students is positive and significant at the 1% level. When the dependent variable with a lag of 1 year or 3 years was used alternatively, all coefficients of variables except the concurrence of official language were positive and significant at the 1% level as shown in (c) and (d) in Table 9.

The combinations of countries in our dataset includes combinations of countries that do not usually publish papers in *Nature* and *Science* in the target period; e.g., combinations of the least-developed countries. Therefore, regression was carried out with the zero-inflated Poisson model, taking the number of foreign students as an indicator to categorize a pair of countries into the always-zero group or not-always-zero group. The result is presented in (e) Table 9. The result of the Vuong test does not support the use of the zero-inflated model (z = 0.83, n.s.), although the tendency of the sign and significance level of the coefficient are similar to the result estimated by Poisson regression.

Table 9 Estimation results of the model of international co-authored papers with chemistry papers

Dependent variable / Estimation method / Independent variable	(a)	(b)	(c)	(d)
	integer count	Fractional count	integer count	Fractional count
	Poisson regression		zero-inflated Poisson regression	
Quadratic Expenditure on R&D ($RFund_{(i)}$ ・ $RFund_{(j)}$)	0.368*** (6.72)	0.339*** (6.78)	0.371*** (6.79)	0.356*** (7.10)
Quadratic Number of researchers ($RPeople_{(i)}$ ・ $R\text{People}_{(j)}$)	0.176*** (6.33)	0.183*** (7.26)	0.172*** (6.20)	0.168*** (6.66)
Number of foreign students ($IStud$)	0.390*** (46.95)	0.360*** (47.75)	0.385*** (44.28)	0.330*** (35.32)
Concurrence of official language ($Lang$)	-0.155*** (-2.89)	-0.179*** (-3.54)	-0.152*** (-2.83)	-0.162*** (-3.20)
Distance between capital cities ($Dist$)	0.055*** (3.14)	0.046*** (2.86)	0.056*** (3.19)	0.050*** (3.10)
Membership in EU (EU)	0.365*** (7.59)	0.369*** (8.37)	0.056*** (3.19)	0.369*** (8.35)
Constant	-6.368*** (-14.61)	-5.989*** (-15.08)	0.366*** (7.61)	-5.575*** (-13.85)
Number of foreign students ($IStud$)/zero-inflated			-5.777 (-0.99)	-1.373*** (-8.35)
Constant /zero-inflated			4.341 (1.03)	2.065*** (-13.85)
Pseudo R^2 (McFadden's Adjusted R^2)	0.361	0.345	(0.353)	(0.312)
Number of observations	6,053	6,053	6,053(Number of 0 is 3,677)	6,053 (Number of 0 is 3,677)

2. The relationship between research performance and international collaboration in chemistry

2.1. Introduction

Increasing team work involves a common pursuit in knowledge production (Wuchty et al. 2007). The reasons researchers indicate for collaboration include access to expertise or equipment, or for education/mentoring (Beaver 2001). In summary, advantages of collaboration include an increase in scientific productivity, research quality, innovative capacity, and accumulation of science and technology human capital (S&T HC), although some disadvantages have also been pointed out (Ordonez-Matamoros 2008). Research collaboration and co-authorship are not always consistent because some authorship is not based on collaborative contributions and research collaboration does not always result in publication (Laudel 2002). However, using co-authorship as an indicator of research collaboration has four advantages. These include verifiability, stability, ease of measurement, and data availability, which have been cited as the best documented indicators for research collaboration (Katz & Martin 1997).

Among co-authored papers, internationally co-authored papers (hereafter, international papers) are on average, with a few exceptions, more highly cited than domestic papers (Glanzel 2001; Glanzel & Shubert 2001). Over time, international papers have increased in amount and accounted for 21.6% of the world's publications in 2010, indicating transcendence in knowledge production of the framework of institutes and countries, although it is apparent that the international co-authorship rate varies among countries and disciplines (NISTEP 2011). This phenomenon might be correlated with scientific policy implementation favoring international research collaboration, such as the EU Research Framework Network.

International co-authorship may reflect individual researcher's interests and motivations as stated above, or may be influenced by macro reasons such as international knowledge diffusion, including that due to researchers in China (Bell 2007) or the influence of large and special equipment such as CERN in Switzerland. Links among countries are influenced by proximity of economic, geographical, historical, linguistic, and people's mobility among countries (Zitt et al. 2000; Nagpaul 2003; Choi 2012).

The higher citation rate of international papers on the macro level may also be due to an increase in diversity in research teams (Adams et al. 2005), severe bilateral selection processes to compensate expenses for international liaison, or increases in readers who would like to reduce search costs and prefer papers written by familiar authors (Schmoch & Schubert 2008). The reasons are still under discussion because these have not yet been confirmed. Recently, however, Abramo et al. (2011) presented the results of an empirical analysis that showed, based on an analysis of Italian faculties, a positive relationship between researchers' productivity and the international co-authorship rate. Their study reached conclusions by taking the individual researcher as the unit of analysis. Previous literature lacked this individual perspective.

Researchers' productivity and international networks could be linked through research grants. Many studies around the world have investigated this relationship; Bozeman & Corley (2004) in the United States, Defazio (2009) in the EU, and Ubfal & Maffiolo (2010) in Argentina. Bozeman & Corley (ibid) confirmed that researchers having more grants have bigger networks than those who do not, as well as that professors have bigger networks than do post-doctoral fellows. Other studies indicated a positive relationship between the size of an author's research network and its quality (Ding 2011; Kretschmer 1994). Regarding international collaborations, researchers with higher productivity and more grants increase the number of international students because usually research grants make it possible to attend or hold international conferences. There is a positive relationship between international students and international collaboration (Choi 2011; Regets 2007). Therefore, the amount of international collaboration can easily be assumed to have a positive relationship with the amount of research grants and publication productivity.

Considering the importance for scientific policy as well as for a theoretical framework, the robustness of the Abramo et al. (2011) results should be confirmed (Research Question 1). Although the results could be interpreted that higher productivity of researchers who engage in international collaboration might explain the reason why international papers are more highly cited than domestic papers, other factors stated above, such as greater diversity of teams, may also be involved (Research Question 2). If the higher quality of international papers is only because of researchers' higher productivity, there is little motivation to encourage international collaboration to enhance the quality of research output.

S&T HC encompasses the productive social capital network that enables researchers to create and transform knowledge and ideas, and add to individual human capital endowments and tacit knowledge. This is because knowledge creation is neither a solitary nor singular event (Bozeman et al. 2001). How factors in S&T HC are related each other and lead to scientific output is not yet clear. However, because the experience through international collaboration (social capital network) is embedded into researchers' mindsets, it enhances knowledge or skills (S&T HC) of researchers and results in higher productivity. Higher productivity then leads to grant acquisition and international collaboration, with international collaboration resulting in better output with more S&T HC. Thus, the relationship between international collaboration and researchers' productivity is mutually reinforcing. In this sense, engagement in international collaboration might increase researchers' productivity (Research Question 3).

International paper has better quality. If it is because of researchers' better productivity, productivity will be matter, not international collaboration. Many studies have investigated factors underlying researcher productivity. These factors include grants, age (Levin & Stephan 1991) (science is a young person's game?), position (Abramo et al. 2011), gender (Sandström 2009), education (selectivity), individual talent (taste for "puzzle solving"), size of the laboratory to which researchers belonged (Carayol & Matt 2006), type of employment (Stephan 2005), and collaboration (Lee & Bozeman 2005). David (1994) pointed out cumulative advantage as a reason underlying the grossly unequal distribution of scientific productivity. That is to say, renowned researchers receive more grants and success because of the Matthew effect. Although the factors that initiate favorable cycles have not yet been examined, they might include having new, trans-disciplinary ideas (Burt 2004). Scientists in all fields are also interested in international collaboration simply because of a search for new ideas to be gained beyond their usual neighbors (Wagner 2008)

The way researchers gain new ideas or perspectives is related to inter-institutional / inter-sectorial / international movement. The positive relationship between inter-institutional movement and the productivity of researchers is not yet clearly confirmed (Dietz & Bozeman 2005). Regarding international mobility, however, the rate of stay in foreign countries by Japanese most-highly-cited researchers was higher (73.4%) than that of average researchers (8.9%) (Kato 2011). In the United Kingdom, researchers who experienced international stays were more productive than who did not (BIS 2011). Therefore, it is possible that internationally mobile researchers are more productive, with international networks leading to more internationally co-authored papers compared with researchers who are not as internationally mobile (Research Question 4).

Based on the literature analysis stated above, we examined the following questions:

Research Question 1: Is the positive relationship between researcher productivity and international co-authorship confirmed by a data set different from that of Abramo et al. (2011)?

Research Question 2: Does researcher productivity entirely explain the higher quality of international papers? [Is the quality of international papers higher than that of domestic papers, controlling for productivity (among researchers who author both types of papers)?]

Research Question 3: Does international collaboration enhance researcher productivity? (Are the number and quality of domestic papers authored by researchers with both international and domestic papers higher than that of domestic papers authored by researchers with only domestic papers?)

Research Question 4: Are researchers with international mobility more productive than researchers without international mobility?

2.2. Data set, indicators, methodology, and model

Data set

To minimize field-specific biases, we selected only chemistry. The first reason for this selection was that internationally refereed journals serve an important role in the chemistry research community, making bibliometric analysis applicable (Van Raan 2004). The second reason was the potential linkage with industry (Defazio et al. 2009). Chemistry had a lower international collaboration rate than the average of all fields from 1995 on, and the rate has increased slower than that of the others (NISTEP 2011). Thus, this field has more room for policy support for international collaboration.

Then we selected the top 16 journals in chemistry based on Impact Factors (IF) from the Journal of Citation Records (JCR) in 2002, 2003, 2005, 2007. We obtained meta-information of articles yearly from 1985 to 2005 from Thomson Reuters Web of Science (WoS).

We retrieved 245,246 articles, which became 188,081 after excluding those lacking the name of the reprinted author and the number of times cited [12]. The number of researchers (reprinted authors only) was 49,599, but we excluded one for empirical analysis since count data models' regression failed to converge with it. The issue here was identification of researchers, especially those with short or popular names, or very large numbers of publications. We checked the year and affiliation of 32 researchers with 100 or more papers and found that the average number of affiliated institutions was 3.3 [13]. We considered the possibility of unclear name identification but bracketed this issue for future studies.

Table 1 shows a comparison of the data between the current study and that of Abramo et al. (2011). The main differences are the countries included and inclusion of IFs of the targeted journals. Because the relationship between research collaboration and co-authorship in developing countries is different from those in developed countries (Duque et al. 2005) we categorized countries based on OECD membership if necessary [14].

---

[12] We used author information that only appears in the reprinted authors' column because only in recent years did the names in the author column uniquely correspond with countries.
[13] Six authors affiliated with only one institute. Most researchers published almost all papers in only one institute.
[14] OECD member countries include the following, which became members before 1990: Austria, Belgium, Canada, Denmark, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States, Japan, Finland, Australia, New Zealand.

Table 10 Data set comparison with Abramo et al. (2011)

	This study	Abramo et al. (2011)
Country	87 countries	Italy
Time	1985-2005	2001-2005
People	Researcher in Reprint Address of paper	Faculty in Italian university (stable and publish one paper or more in the period)
Subject field	Chemisty	9 areas in natural science
Journals	Journals with high Impact Factor	no mention
Original data source	Web of Science	Web of Science

Indicators

To examine the link between researchers' productivity and internationalization, we used six indicators that Abramo et al. (2011) created; three for research performance and three for internationalization. However, we slightly amended these based on characteristics of our data, as follows.

Performance Indicators:

- Productivity (P): total publications by a reprinted author in the period under observation;
- Fractional Productivity (FP): total contributions to publications authored by a reprinted author;
- Average Quality (AQ): the quality of each publication as proxied by number of citations (times each publication was cited divided by the average number of citations of all publications in the same year).

Internationalization Indicators:

- International Collaboration Intensity (ICI): total publications with at least one researcher from countries different than that of a reprinted author;
- International Collaboration Rate (ICR): ratio of ICI to P;
- International Collaboration Amplitude (ICA): total foreign countries represented in a cross-national publication.

Methodology

We referred to the methodology used in Abramo et al. (2011). The existence of international papers and the degree of productivity were regressed by a binary logistic regression model (logit model). Abramo et al. (2011) regarded the number of papers as count data and used binary logit, Poisson, and negative binominal, then showed only the results regressed by binary logit because of quite similar results. We also mainly showed the results regressed by binary logit for the comparison. We followed the method of Abramo et al. (2011) of using an ordered logistic model for analysis of degree of internationalization.

Model

Predicted probability of logistic model is shown as follows

$$p_j(0) = \frac{1}{1+\exp\big(f(x_j)\big)}, p_j(1) = 1 - p_j(0) \quad ・・・(1)$$

The probability for ordered logistic model is shown with $K_{(i)}$ representing the threshold between (i)th and (i+1)th categories.

$$p_j(1) = \frac{1}{1+\exp\big(f(x_j) - K_{(1)}\big)}$$

$$p_j(i) = \frac{1}{1+\exp\big(f(x_j) - K_{(i)}\big)} - \frac{1}{1+\exp\big(f(x_j) - K_{(i-1)}\big)} \quad (i = 2, \ldots, I-1) \ldots (2)$$

$$p_j(I) = 1 - \frac{1}{1+\exp\big(f(x_j) - K_{(I-1)}\big)}$$

Category i is selected by $\underset{i}{\text{argmax}}\, p_j(i)$ with K_0 and K_I being defined as -∞ and +∞, respectively. Based on a logistic or ordered logistic models stated above (1) and (2), we formulated our model as follows.

Research Question 1:

Dummy of international collaboration intensity (ICI); DICI=0 if ICI = 0, and DICI=1 if 1 ≤ ICI) is formulated using logistic model (1) as

$$ICI(P): f(x_j) = \alpha_0 + \alpha_1 P_j + \alpha_2 AQ_j \quad ・・・(3)$$

$$ICI(FP): f(x_j) = \beta_0 + \beta_1 FP_j + \beta_2 AQ_j \quad ・・・(4)$$

Dummy of international collaboration rate (ICR); DICR = 0 if ICR =0, DICR = 1 if 0.01 ≤ ICR ≤ 0.25, DICR= 2 if 0.251 ≤ ICR ≤ 0.5, DICR = 3 if 0.501 ≤ ICR ≤ 0.75, DICR = 4 if 0.751 ≤ ICR ≤ 1) is formulated using ordered logistic model (2) as

$$ICR(P):\ f(x_j) = \gamma_0 + \gamma_1 P_j + \gamma_2 AQ_j \quad ・・・(5)$$

$$ICR(FP):\ f(x_j) = \delta_0 + \delta_1 FP_j + \delta_2 AQ_j \quad ・・・(6)$$

Research Question 3:

Model is same as Research Question 1 but targets domestic papers only.

Research Question 4:

Dummy of international mobility (Imove; Imove=0 if number of international move is 0, and Imove is 1 if number of international move is 1 or more) is formulated using logistic model (1) as

$$IMove(P):\ f(x_j) = \theta_0 + \theta_1 DICI_j + \theta_2 P_j + \theta_3 AQ_j \quad ・・・(7)$$

$$IMove(FP): f(x_j) = \mu_0 + \mu_1 DICI_j + \mu_2 FP_j + \mu_3 AQ_j \quad ・・・(8)$$

2.3. General analysis

Papers

Table 2 shows the number of foreign countries involved in international publications[15]. Domestic papers accounted for 82.7% of the papers and international papers accounted for 17.3%. The percentage of cross-national publications involving single foreign countries was 86.8%. Papers with involvement of three or more foreign countries represented only 1.6% of the total of cross-national publications. The Appendix shows titles of the journals surveyed and the chemistry subfields involved.

Table 11 Number of foreign countries involved in international publications

Number of foreign countries involved	Number of publications	% on Total in Cross-national publications	% on Total in publications
0	155,613	-	82.7%
1	28,176	86.8%	15.0%
2	3,785	11.7%	2.0%
3	444	1.4%	0.2%
More than 3	63	0.2%	0.0%
Total	188,081	100.0%	100.0%

Table 3 classifies countries by number of publications. We included 87 countries in our data set. The United States topped the list in numbers of both domestic and international papers. European countries held a higher share of international papers, with countries such as Japan, China, and India holding higher shares in domestic papers among the top 15 countries containing data[16]. When we examined publications per 1,000 researchers, the number of domestic papers was relatively smaller in Asian countries than in other countries. However, publications per 1,000 researchers for international papers were smaller in countries with large populations, including the United States.

[15] The scale effect of countries should be considered. For instance, smaller countries in terms of population have higher rates of internationally co-authored papers compared with more populated countries (Kato & Chayama 2010).

[16] These Asian countries have large populations, and their mother tongue is not English; therefore they might publish papers in domestic journals using their own languages. However, considering that almost 80% of Japanese doctoral dissertations in physics are written in English but only 25% in engineering (Matsuoka et al. 2003), papers in chemistry are usually published in English even for domestic papers. This could be included in our data.

Table 12 Classification of countries by number of publications

Domestic papers				International papers			
Country	Total Publications	% on Total Publications	Publication per 1,000 researchers	Country	Total Publications	% on Total Publications	Publication per 1,000 researchers
United States	60,348	38.8%	47.1	United States	6,675	20.5%	5.2
Japan	17,976	11.6%	27.3	Germany	3,019	9.3%	11.8
Germany	10,662	6.9%	41.6	United Kingdom	2,914	8.9%	15.7
United Kingdom	10,172	6.5%	54.8	France	2,206	6.8%	12.6
France	6,714	4.3%	38.3	Spain	2,073	6.4%	27.0
China	6,222	4.0%	8.6	Italy	1,955	6.0%	28.0
Spain	5,504	3.5%	71.7	Japan	1,482	4.5%	2.2
Canada	5,344	3.4%	48.3	Canada	1,252	3.8%	11.3
Italy	5,302	3.4%	75.9	China	962	3.0%	1.3
India	2,533	1.6%	19.1	Switzerland	879	2.7%	35.8
Netherlands	2,404	1.5%	58.6	Netherlands	777	2.4%	18.9
Korea, Rep.	2,374	1.5%	18.5	Sweden	709	2.2%	15.6
Australia	2,252	1.4%	32.6	Australia	619	1.9%	9.0
Switzerland	2,170	1.4%	88.4	Belgium	552	1.7%	18.4
Taiwan	2,094	1.3%	-	India	550	1.7%	4.1
Total	142,071	91.4%	-	Total	26,624	81.7%	-

Figure 3 shows the change in quantity and quality of papers from 1985 to 2005. The numbers of both domestic and international papers increased. The international rate more than doubled over 20 years: 7.9% in 1985 and 19.4% in 2005. Paper quality as proxied by number of citations was smaller in recent years, and roughly similar between domestic and international papers. Mean of AQ in domestic papers was 1.00, and one in domestic papers was 0.99.

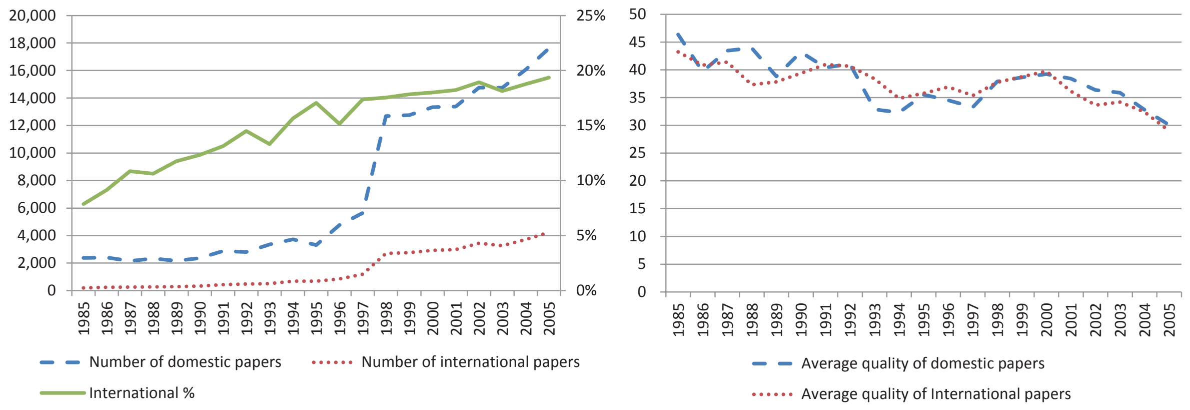


Figure 3 Change in quantity and quality of papers from 1985 to 2005

The result is different from general conception that regards internationally co-authored papers with higher citation rates. However, as there is variety in citation impact between counties or within one individual country between fields as Glanzel (2001) mentioned, the level of journal might relate to difference in the citation rate between domestic and international papers. Seeing change of average citations using data of articles published in either Nature or Science from1990 to2009, same tendency, little difference between average quality of domestic and international papers, is confirmed.[17]

Researchers

Researchers with one paper authored consisted of 53.6% of the sample (Table 13). Researchers affiliated with only one country were the most represented (93.1%) (Table 14), and among 23,029 researchers authoring two or more papers—in other words, researchers in our data set having the possibility to move—researchers affiliated with more than two countries represented 14.3% of the total. Researchers who stayed in OECD-member countries represented 83.2% of the total, with the rest staying in both OECD and non-OECD countries or non-OECD countries only (Table 15).

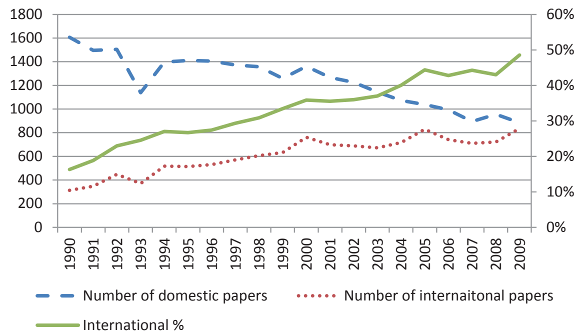
Table 13 Classification of researchers by number of papers authored

Number of papers	Number of authors	%
1	26,570	53.6%
2	7,808	15.7%
3	3,781	7.6%
4	2,358	4.8%
5 to 9	5,038	10.2%
10 to 19	2,607	5.3%
20 to 99	1,405	2.8%
100 or more	32	0.1%
Total	49,599	100.0%

Table 14 Classification of researchers by number of countries researcher was affiliated with

Number of countries	Number of authors	%
1	46,160	93.1%
2	3,027	6.1%
3	355	0.7%
4	48	0.1%
5	7	0.0%
6	2	0.0%
Total	49,599	100.0%

[17] Following is the result analyzed on data from Nature or Science from1990 to2009.



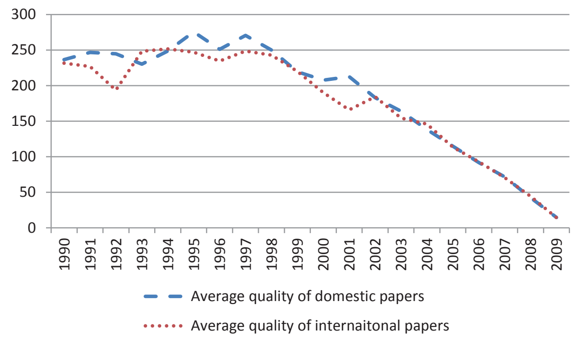


Table 15 Classification of researchers who stayed in OECD or non-OECD countries

Category	Observations	%
OECD only	41,266	83.2%
Non-OECD only	6,977	14.1%
Both OECD and Non-OECD	1,356	2.7%
Total	49,599	100.0%

Research performance and international collaboration

Table 7 shows descriptive statistics of 49,599 researchers. Compared with the data shown in Abramo et al. (2011), the mean AQ in our sample was higher (0.86 vs. 0.69 in Abramo et al. (2011)) but means of other indicators were lower. Table 7 also shows statistics with and without international mobility. Means of all six indicators were higher for researchers with international mobility than those without [18].

Table 16 Descriptive statistics of researcher indicators

Categories / Variable	Total (49,599 observations)				Author without international mobility (46,160 observations)				Author with international mobility (3,439 observations)			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
P	3.67	7.52	1	297	3.19	6.36	1	229	10.06	15.16	1	297
FP	1.18	2.49	0.04	95.35	1.04	2.14	0.04	89.28	3.18	4.84	0.1	95.35
AQ	0.86	1.20	0.02	80.8	0.85	1.22	0.02	80.82	1.03	0.87	0.02	14.41
ICI	0.65	1.87	0	91	0.49	1.35	0	63	2.86	4.52	0	91
ICR	0.19	0.34	0	1	0.17	0.34	0	1	0.34	0.31	0	1
ICA	0.55	1.06	0	18	0.43	0.87	0	16	2.25	1.64	1	18

Researchers were categorized into three groups: (i) those with domestic papers only, (ii) those with international papers only, and (iii) those with both domestic and international papers. Table 17 presents descriptive statistics by group. Researchers with domestic papers only numbered 34,434 (69.4%), and researchers with international papers numbered 15,165 (30.6%).

Researchers with both domestic and international papers had the highest productivity in quantity and quality among the three groups, and researchers with only international papers were least productive. If international papers have some kind of advantage because of internationality, researchers who only authored papers with such a "bonus" had a lower performance than those without. Considering the level of journals we chose, these researchers could author domestic papers in journals with smaller IFs.

[18] Age differences should be noted between the two groups; for instance, researchers without international mobility were younger and might not have had enough time to stay in foreign countries at that point in his/her career; however, this type of information was not available in our data set.

Table 17 Descriptive statistics of researcher indicators categorized by internationality

Category	Number of authors	P	AQ
All Papers	49,599	3.67	0.86
Domestic paper only	34,434	2.49	0.84
Include international Paper	15,165	6.75	0.91
(International paper only)	5,864	1.37	0.82
(Both domestic and international paper)	9,301	10.15	0.97

2.4. Results

This section attempts to provide answers to the research questions posed earlier.

Relationship between research productivity and international collaboration (Research Question 1)

Table 9 shows results of Spearman correlations between indicators (H_0: No correlation between two indicators; coefficient = 0). As Abramo et al. (2011) showed, the correlation analysis indicates a strong link between productivity and international collaboration although the coefficient is smaller, except for the relationship between ICR and ICI. The Spearman correlation coefficient between P and ICI was 0.4170. We found similar results for FP. The correlation between ICI and AQ was also significant and positive (0.1273).

The correlation between productivity and ICR, while again significant, was quite weak compared with the others (0.0835 for FP and 0.0840 for AQ). The degree of propensity for international collaboration was weakly correlated with the contribution to papers and average quality of papers.

Table 18 Spearman correlations between indicators

	P	FP	AQ	ICI	ICR	ICA
P	1	0.7696***	0.2246***	0.4170***	0.2553***	0.4745***
FP		1	0.1340***	0.2439***	0.0835***	0.3439***
AQ			1	0.1273***	0.0840***	0.1854***
ICI				1	0.9574***	0.8079***
ICR					1	0.7549***
ICA						1

Number of observations: 49,599; Statistical significance: ***p <0.01

We applied binary logistic regression analysis to assess the relationship between research performance and intensity of international collaboration. The dependent variable ICI was assumed to be 1 if researchers had one or more paper, otherwise it was nil. Table 10 presents the results. The coefficient of P was positive and significant, but not that of AQ. The results differ from those of Abramo et al. (2011) which showed positive and significant coefficients for both variables.

Assuming a difference in the relationship between developed and developing countries, we performed a regression separating countries by OECD membership (Table 19). The coefficient of AQ was positive and significant at the 10% level for OECD countries, and at 1% for non-OECD countries. Therefore, the relationship between research performance and intensity of international collaboration could be slightly different depending on the level of economic development of the country involved.

When FP replaced P, the coefficient of performance indicators (FQ and AQ) showed positive and significant results, though the coefficient of AQ (0.026) was quite low compared with the 0.889 found by Abramo et al. (2011).

Table 19 Binary logistic regression of international collaboration versus performance indicators (P and AQ)

Categories / Variables	Total (Number of obs: 49,598, Pseudo R2: 0.082)				OECD only (Number of obs: 41,265, Pseudo R2: 0.0823)				Non-OECD included (Number of obs: 8,333, Pseudo R2: 0.0827)			
	Coef.	Std. Err.	z	Pr>z	Coef.	Std. Err.	z	Pr>z	Coef.	Std. Err.	z	Pr>z
P	0.146	0.003	54.34	0	0.143	0.003	49.26	0	0.160	0.007	22.32	0
AQ	0.012	0.008	1.44	0.149	0.014	0.008	1.65	0.099	0.080	0.031	2.62	0.009
Cons	-1.343	0.015	-90.29	0	-1.427	0.016	-86.84	0	-1.019	0.037	-27.4	0

Table 20 Binary logistic regression of international collaboration versus performance indicators (FP and AQ)

	Coef.	Std. Err.	z	Pr>z
FP	0.269	0.006	42.81	0
AQ	0.026	0.008	3.19	0.001
Cons	-1.161	0.014	-82.94	0

Number of observations: 49,598; Pseudo R2: 0.0441

To examine the relationship between the international collaboration rate and performance indicators, we used an ordered logistic regression with the ICR as the dependent variable, categorized into four as stated earlier.

Table 12 and Table 13 show the results of the ordered logistic regression. The international collaboration rate shows positive and significant dependence for both P and AQ as Abramo et al. (2011) presented. Similar results entail when FP was used instead of P as the independent variable.

Table 21 Ordered logistic regression of international collaboration rate to performance indicators (P and AQ)

	Coef.	Std. Err.	z	Pr>z
P	0.037	0.001	31.39	0
AQ	0.015	0.008	1.96	0.05
/cut1	1.010	0.013		
/cut2	1.402	0.014		
/cut3	1.969	0.016		
/cut4	2.160	0.016		

Number of observations: 49,598; Pseudo R2: 0.0119

Table 22 Ordered logistic regression of international collaboration rate to performance indicators (FP and AQ)

	Coef.	Std. Err.	z	Pr>z
FP	0.076	0.003	22.94	0
AQ	0.023	0.008	2.83	0.005
/cut1	0.957	0.013		
/cut2	1.344	0.014		
/cut3	1.908	0.015		
/cut4	2.100	0.016		

Number of observations: 49,598; Pseudo R2: 0.0061

Quality difference between domestic papers and international papers presented by researchers with both types of papers (Research Question 2)

To examine the possibility that factors other than researchers' productivity influenced the higher quality of international papers, we conducted Student's *t* tests and F test to compare the quality between international and domestic papers among researchers who authored both types of papers. The results of *t* tests showed that both means were statistically different from each other at the 1% level (t (9300) = -36.239, p <.01) as well as that the mean value representing quality of domestic papers was less than that of international papers at the 1% level (p <.01).

When we compared the standard deviations (variances) between the quality of international and domestic papers using F test, we could reject the hypothesis that the standard deviations were the same at the 10% significance level (F(9300, 9300) = 0.962, p <.01) as well as that the variance of quality in domestic papers was less than that of international papers at the 5% level. Therefore, the quality of international papers was different and probably higher than that of domestic papers among researchers who authored both international and domestic papers.

Relationship between accumulation of S&T HC and international collaboration (Research Question 3)

To examine relationship between accumulation of S&T HC and international collaboration we conducted the regression exactly as with Research Question 1 except we used only domestic papers targeting researchers with domestic papers only and those with both international and domestic papers. We assume that if international collaboration accumulates S&T HC, the quality and quantity of domestic papers authored by researcher with international collaboration is higher than that of colleagues without the collaboration. Table 14 and Table 15 present the results of the regression. The international collaboration rate showed a positive and significant relationship with P, FP, and AQ at the level of 1%. Researchers who collaborated internationally had higher performances even in productivity of domestic papers.

Table 23 Binary logistic regression of international collaboration versus domestic performance indicators (P and AQ)

	Coef.	Std. Err.	z	Pr>z
P(Domestic)	0.134	0.003	50.62	0
AQ(Domestic)	0.043	0.010	4.41	0
Cons	-1.866	0.018	-104.49	0

Number of observations: 43,734; Pseudo R2: 0.0887

Table 24 Binary logistic regression of international collaboration versus domestic performance indicators (FP and AQ)

	Coef.	Std. Err.	z	Pr>z
FP(Domestic)	0.343	0.007	47.65	0
AQ(Domestic)	0.053	0.010	5.17	0
Cons	-1.799	0.018	-102.05	0

Number of observations: 43,734; Pseudo R2: 0.0753

Using ICR, categorized into five categories as stated above, as the dependent variable in an ordered logistic regression, we examined the relationship between the international collaboration rate and domestic performance indicators. Table 16 and Table 17 present the results. The international collaboration rate showed a positive and significant relationship with P, FP, and AQ at the 1% level. Therefore, researchers who collaborated internationally more heavily than those who did not had a higher performance even in productivity of domestic papers.

Table 25 Ordered logistic regression of international collaboration rate to performance indicators (P and AQ)

	Coef.	Std. Err.	z	Pr>z
P(Domestic)	0.054	0.001	36.5	0
AQ(Domestic)	0.049	0.010	4.97	0
/cut1	1.604	0.016		
/cut2	2.237	0.019		
/cut3	3.748	0.030		
/cut4	5.441	0.063		

Number of observations: 43,734; Pseudo R2: 0.0258

Table 26 Ordered logistic regression of international collaboration rate to performance indicators (FP and AQ)

	Coef.	Std. Err.	z	Pr>z
FP(Domestic)	0.145	0.004	34.23	0
AQ(Domestic)	0.055	0.010	5.43	0
/cut1	1.583	0.016		
/cut2	2.212	0.019		
/cut3	3.722	0.030		
/cut4	5.415	0.063		

Number of observations: 43,734; Pseudo R2: 0.0227

Research productivity and international mobility (Research Question 4)

We applied binary logistic regression analysis to assess the relationship between research productivity and international mobility. We assumed the dependent variable to be 1 if researchers moved internationally (i.e., they were affiliated with two or more countries); otherwise it was nil. Independent variables were an international collaboration dummy, which was 1 if researchers authored one or more international paper; otherwise it was nil, P, and AQ. Table 18 presents the results. Coefficients of all three variables were positive and significant; the international collaboration dummy was 5% and both P and AQ were 1%. Only researchers with two or more papers had the possibility of international movement in our data set. We examine the data restricted to those researchers and confirmed that the tendency was similar to the results shown in Table 28. Therefore, researchers with international movement had higher productivity than researchers without such movement.

Table 27 Binary logistic regression of international movement versus performance indicators (ICI (Dummy), FP, and AQ)

	Coef.	Std. Err.	z	Pr>z
ICI(Dummy)	1.255	0.023	54.03	0
P	0.023	0.002	13.91	0
AQ	0.053	0.011	4.9	0
Cons	-5.184	0.056	-92.25	0

Number of observations: 49,599; Pseudo R2: 0.1898

Table 28 Binary logistic regression of international movement versus performance indicators among researchers with two or more papers (ICI (Dummy), FP, and AQ)

	Coef.	Std. Err.	z	Pr>z
ICI(Dummy)	0.789	0.027	29.28	0
P	0.011	0.002	6.82	0
AQ	0.134	0.026	5.25	0
Cons	-3.598	0.076	-47.57	0

Number of observations: 15,221; Pseudo R2: 0.0886

2.5. Conclusion and discussion

This study used a data set in chemistry to examine the robustness of the results presented by Abramo et al. (2011) and the possible impact of international collaboration and mobility among researchers. A summary of answers to the four research questions follows.

One of the results confirmed the positive relationship found by Abramo et al. (2011) between researchers' performance and international collaboration (Research Question 1). However, the higher quality of international papers was not solely explained by the higher performance of researchers because the quality of international papers was higher than that of domestic papers controlling researchers' productivity (Research Question 2). Therefore, international research collaboration seems to exert some kind of "bonus" effect due to internationalization. The results also showed that the quantity and quality of domestic papers by researchers with both international and domestic papers was higher than that of researchers with only domestic papers (Research Question 3). This could indicate that researchers who collaborate internationally accumulate S&T HC by acquiring diverse or new ideas from colleagues, resulting in higher quality of domestic papers. An alternative explanation posits the existence of selectivity for internationalization among researchers with domestic papers. This explanation would have to be explored in more depth to be confirmed. Finally, our results show a positive relationship between international mobility and researchers' performance. The direction of cause and effect is not yet clear but indicates the possibility of an impact of international mobility on researchers' performance.

Future research should include methodological improvement and additional themes. Methodological improvements could be made to researcher identification and data collection. Data in the future should include a wider variety of journal levels based on IF, field, and countries. Including researchers' curricula vitae could enable more detailed analyses regarding international movement and collaborations. Additional themes include examining the direction of cause and effect between researchers' performance and international collaboration, considering how international co-authorship impacts S&T HC accumulation. Since the present definition of international research is affected when the country merges or becomes independent, an alternative measure invariant of country scales should also be developed.

References

Abramo, G., D'Angelo, C. A., & Di Costa, F. (2011). Research productivity: Are higher academic ranks more productive than lower ones?. *Scientometrics*, 88(3), 915-928.

Abramo, G., D'angelo, C. A., & Solazzi, M. (2011). The relationship between scientists' research performance and the degree of internationalization of their research. *Scientometorics*, 86(3), 629-643.

Adams, J. D., Black, G. C., Clemmons, J. R., & Stephan, P.E. (2005). Scientific teams and institutional collaboration: Evidence from U.S. universities, 1981-1999. *Research Policy*, 34(3), 259–285.

Beaver, D. D. (2001). Reflections on scientific collaboration (and its study): past, present, and future. *Scientometrics*, 52 (3), 365–377.

Bell, R. K., Hill, D., & Lehming, R. F. (2007). The changing research and publication environment in American research universities. NSF Working Paper. SRS 07-204 . http://www.nsf.gov/statistics/srs07204/pdf/srs07204.pdf. Accessed 13 September 2010.

BIS. (2011). *International comparative performance of the UK research base – 2011.* http://www.bis.gov.uk/assets/biscore/science/docs/i/11-p123-international-comparative-performance-uk-research-base-2011. Accessed 28 August 2012.

Bozeman, B., & Corley, E. A. (2004). Scientists' collaboration strategies: Implications for scientific and technical human capital. *Research Policy*, 33(4), 599-616.

Bozeman, B., Dietz, J.S., & Gaughan, M. (2001). Scientific and technical human capital: An alternative model for research evaluation. *International Journal of Technology Management*, 22 (7), 716-740

Burt, R. S. (2004). Structural holes and good ideas. *American Journal of Sociology*, 110(2), 349–399.

Carayol, N., & Matt, M. (2006). Individual and collective determinants of academic scientists' productivity. *Information Economics and Policy,* 18(1), 55-72.

Choi, S. (2012).Core-periphery, new clusters, or rising stars?: International scientific collaboration among 'advanced' countries in the era of globalization. *Scientometrics,* 90(1), 25–41.

David, P. A. (1994). Positive feedbacks and research productivity in science: reopening another black box. In O.Grandstrand (Ed.), *Economics of technology* (pp.65-89). Amsterdam: North-Holland.

Defazio, D., Lockett, A., & Wright, M. (2009). Funding incentives, collaborative dynamics and scientific productivity: Evidence from the EU framework program. *Research Policy,* 38(2), 293–305.

Dietz, J.S., & Bozeman, B. (2005). Academic careers, patents, and productivity: industry experience as scientific and technical human capital. *Research Policy*, 34(3), 349–367.

Ding, Y. (2011). Scientific collaboration and endorsement: Network analysis of coauthorship and citation networks. *Journal of Informetrics,* 5(1), 187–203.

Duque, R B.,Ynalvez, M., Sooryamoorthy, R., Mbatia, P., Dzorgbo, D. S., & Shrum, W. (2005). Collaboration paradox: Scientific productivity, the internet, and problems of research in developing areas. *Social Studies of Science*, 35(5), 755-785.

Glänzel, W., & Schubert, A. (2001). Double effort = double impact? A critical view at international co-authorship in chemistry. *Scientometrics*, 50 (2), 199-214.

Glänzel, W. (2001). National characteristics in international scientific co-authorship relations. *Scientometrics,* 51(1), 69–115.

Kato, M. (2011). Analysis on career paths of the most highly cited scientists in Japan through international comparison. Discussion Paper no. 78. NISTEP

Kato, M., & Chayama, H. (2010). Analysis on research activities in developing countries and international networking of researchers. Research Material 178. NISTEP

Katz, J. S., & Martin, B. R. (1997). What is research collaboration?. *Research Policy,* 26(1), 1-18.

Kretschmer, H. (1994). Coauthorship networks of invisible colleges and institutionalized communities. *Scientometrics*, 30(1). 363-369.

Laudel, G. (2002). Collaboration and reward: What do we measure by co-authorships?. *Research Evaluation*, 11(1), 3–15.

Lee, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science*, 35(5), 673-702.

Levin, S. G.,& Stephan, P. E. (1991). Research productivity over the life cycle: Evidence for academic scientists. *The American Economic Review*, 81(1), 114 - 132.

Muraoka, T., Nishina, K., Fukao, Y., Chinami, K., & Otani, S. (2003). The choice of language for international graduate students in scientific fields. *Journal of Technical Japanese Education*, 5, 55-60

NISTEP. (2011). Japanese science and technology indicators 2011. Research Material No.198. NISTEP.

Ordóñez-Matamoros, G. (2008). International research collaboration, research team performance, and scientific & technological capabilities in Colombia –A bottom-up perspective. Paper presented in the IV Globelics Conference at Mexico City, September 22-24.

Regets, M. C. (2007). Research issues in the international migration of highly skilled workers: A perspective with data from the United States. NSF Working Paper SRS 07-203. http://www.nsf.gov/statistics/srs07203/. Accessed 28 August 2012.

Schmoch, U., & Schubert, T. (2008). Are international co-publication an indicator for quality of scientific research? . *Scientometrics*, 74(3), 361–377.

Stephan, P. E. (2008). Job market effects on scientific productivity. http://www.cso.edu/upload/PDF_rencontres/SEM_ES_Paula-Stephan.pdf. Accessed 28 August 2012.

Ubfal, D., & Maffioli, A. (2011). The impact of funding on research collaboration: Evidence from developing country. *Research Policy*, 40(9), 1269-1279.

Sandström, U (2009). Combining curriculum vitae and bibliometric analysis: mobility, gender and research performance. *Research Evaluation*, 18 (2), 135-142.

Van Raan. A. F. J. (2004) Measuring Science: Capita selecta of current main issues. In Moed, H.F., Glänzel, W., & Schmoch, U. (eds.) *Handbook of qualitative science and technology research: The use of publication and patent statistics in studies on S&T systems* (pp. 19-50). Dordrecht: Kluwer Academic Publishers.

Wagner, C.S. (2008). *The new invisible college: Science for Development*. Washington: Brookings Institution Press.

Wuchty, S., Benjamin, J. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science,* 316(5827), 1036-1039.

Appendix

Journal titles included

Journal Title	Total Publications	Shares
Journal of The American Chemical Society	27,103	14.41%
Journal of Organic Chemistry	19,567	10.40%
Journal Of Physical Chemistry B	15,329	8.15%
Langmuir	13,784	7.33%
Inorganic Chemistry	12,284	6.53%
Chemical Communications	11,522	6.13%
Journal of Chromatography A	10,708	5.69%
Analytical Chemistry	10,249	5.45%
Angewandte Chemie-International Edition	9,608	5.11%
Organometallics	9,123	4.85%
Journal of Medicinal Chemistry	7,979	4.24%
Organic Letters	7,902	4.20%
Journal of The Chemical Society-Dalton Transactions	6,517	3.46%
Chemistry of Materials	6,349	3.38%
Journal of Materials Chemistry	5,290	2.81%
Journal of Catalysis	5,196	2.76%
Molecular Sieves: From Basic Research To Industrial Applications, PTS A and B	4,960	2.64%
Electrophoresis	4,611	2.45%
Total	188,081	100.00%

Chemistry subfields

Category	Total Publications	Shares
Chemistry	134,107	71.35%
Chemistry; Materials Science	25,416	13.52%
Biochemistry & Molecular Biology; Chemistry	15,271	8.12%
Pharmacology & Pharmacy	7,978	4.24%
Chemistry; Engineering	5,195	2.76%
Total	187,967	100.00%

3. Status of Japan in international scientific collaboration in the field of Chemistry

In this section, we analyze the status of Japan in international scientific collaboration in the field of chemistry using the dataset described in section 2.2 by comparing with the United Kingdom (U.K.) and the United States (U.S.) as leading countries and China as an emerging power in the development of scientific knowledge. Here, we define papers as being associated with a particular country by considering the affiliated institution of the author. For instance, if the author of the published paper (author appeared in reprint address in the paper) is affiliated with an institute in Japan, we classify the paper as being a Japanese paper, regardless of the nationality of the author. We also define Japanese authors as those who have had at least one paper published, and who are affiliated with an institution located in Japan.

3.1 Number of papers and average citations

Number of papers and average citations from 1985 to 2005

The number of Japanese papers published from 1985 to 2005 was 19,458, which was more than that of the U.K. and China (almost 1.5 times that of the U.K. and 2.7 times that of China), but less than that of the U.S. (0.3 times). The ratio of citations from international papers in Japanese papers was 7.62%, with almost one third of this from the U.K. (22.27%), similar to that of the U.S. (9.96%).

The average number of standardized citations (AQ) in Japanese papers was 0.95, which was more than that of the U.K. (0.88) and less than that of the U.S. (1.15) and China (1.17). In Japan, the U.K. and China, the AQ of international papers was greater than those of domestic ones; 0.07 in Japan, 0.10 in the U.K. and 0.08 in China. On the other hand, this is smaller in the U.S; the AQ of international papers is 0.02 smaller than those of domestic papers.

Table 29 Number of papers and average citations among four countries

	All papers			Japan		
	Number of papers	Ratio	AQ	Number of papers	Ratio	AQ
Domestic Papers	155,613	82.74%	1.00	17,976	92.38%	0.95
International Papers	32,468	17.26%	0.99	1,482	7.62%	1.02
Total	188,081	100.00%	1.00	19,458	100.00%	0.95

	United Kingdom			United States		
	Number of papers	Ratio	AQ	Number of papers	Ratio	AQ
Domestic Papers	10,172	77.73%	0.86	60,348	90.04%	1.16
International Papers	2,914	22.27%	0.96	6,675	9.96%	1.14
Total	13,086	100.00%	0.88	67,023	100.00%	1.15

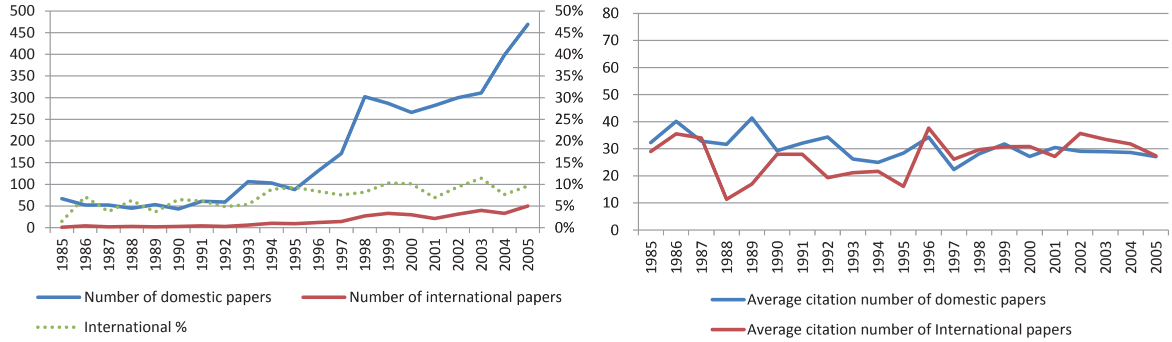
	China		
	Number of papers	Ratio	AQ
Domestic Papers	6,222	86.61%	1.16
International Papers	962	13.39%	1.24
Total	7,184	100.00%	1.17

Change in quantity and quality of papers from 1985 to 2005

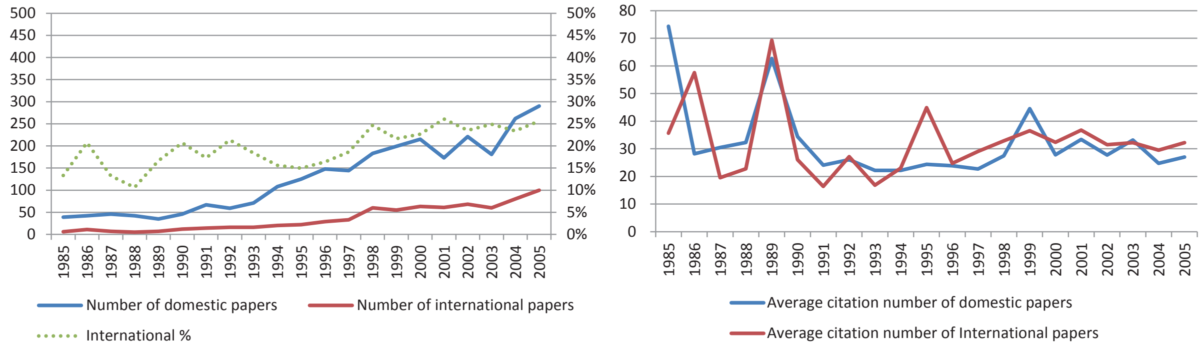
Figure 4 shows the change in quantity and quality of papers in Japan, the U.K. the U.S., and China from 1985 to 2005. Until 2000, the number of international papers published in China was around zero, therefore the international ratio and average number of citations should be disregarded for this period in China.

The number of both domestic and international papers published in these four countries has increased. The international rate has also increased in all four countries. The quality of the papers as determined by the number of citations has been similar in recent years between domestic and international papers, although in the citation of international papers was less in the late 80s and early 90s in Japan.

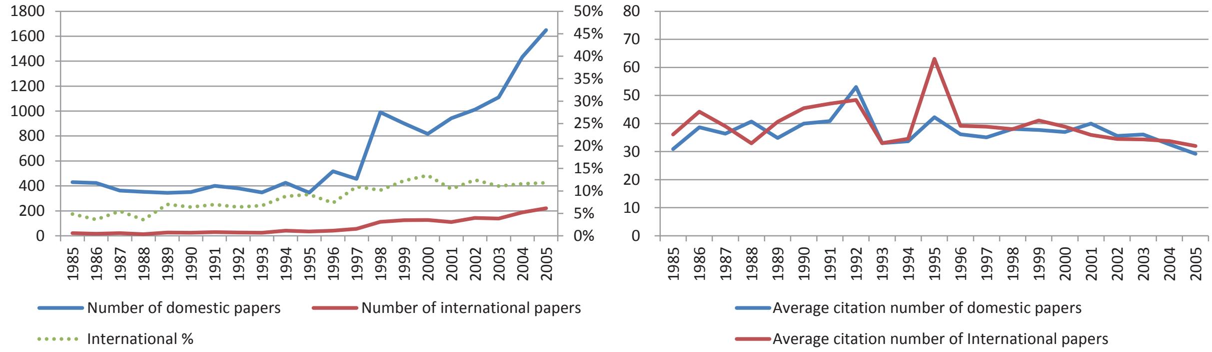
【Japan】



【United Kingdom】



【United States】



【China】

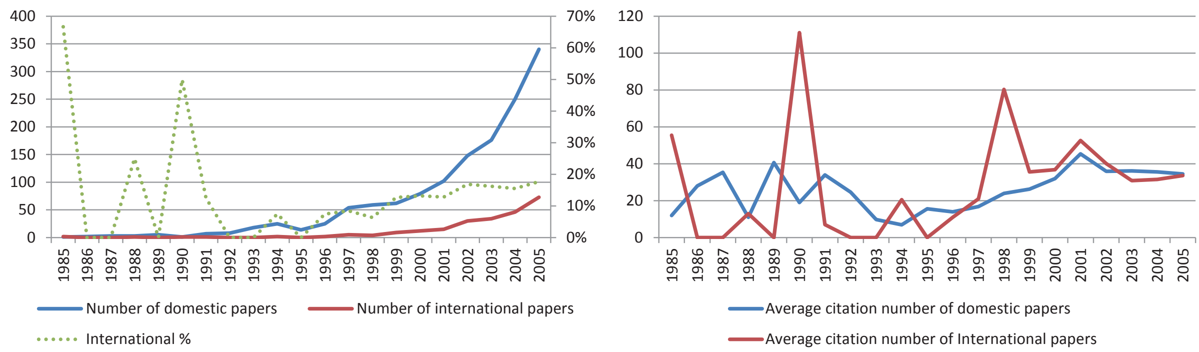


Figure 4 Change in quantity and quality of papers among four countries from 1985 to 2005

3.2. Researchers’ performance and internationalization of papers

General analysis

Ratio of researchers who published international papers

Table 30 shows the number of authors among the four countries. The number authors in Japan was 4,170（8.41%） among 49,599 researchers (reprinted authors only). 4,939 authors were in the U.K. and 16,426 were in the U.S. On the other hand, the number of authors in China was relatively small at only 1,892.

The ratio of authors who collaborated only domestically and were affiliated with institutes in Japan was 77.70%, which is more than the ratio of all authors in the dataset (69.42%). On the other hand, the ratio of authors who only collaborated internationally and were affiliated with institutes in Japan was 4.53%, which is less than the ratio of authors in the dataset who did the same (11.82%). This was interpreted as meaning that Japanese researchers tend to publish papers in internationally circulated journals without international collaboration.

The ratio of international collaboration is similar between Japan and the United States. Authors in the U.K. are more likely to internationally collaborate than colleagues in these two countries. For instance, the ratio of international collaboration in the U.K. (37.92%) is more than that of Japan (22.30%) the U.S. (22.78%) and China (12.81%).

Among Japanese researchers, authors who published papers through both domestic and international collaboration had the highest AQ (0.93), which was 10% more than the others, followed by authors who published papers through only domestic collaboration (0.79). The lowest score was by authors who published papers only through international collaboration (0.77). Authors who published papers through both domestic collaboration and international collaboration tended to have the highest average number of citations, which is the same in all four countries. However, the order of other two was the opposite between both Japan and China and both the U.K. and the U.S.

Table 30 Number and ratio of authors who publish papers internationally

	All papers			Japan		
	Number	Ratio	AQ	Number	Ratio	AQ
Domestic papers only	34,434	69.42%	0.84	3,240	77.70%	0.79
International papers only	5,864	11.82%	0.82	189	4.53%	0.77
Both international and domestic papers	9,301	18.75%	0.97	741	17.77%	0.93
Total	49,599	100.00%	0.86	4,170	100.00%	0.82

	United Kingdom			United States		
	Number	Ratio	AQ	Number	Ratio	AQ
Domestic papers only	3,066	62.08%	0.83	12,684	77.22%	0.98
International papers only	651	13.18%	0.89	949	5.78%	1.01
Both international and domestic papers	1,222	24.74%	0.94	2,793	17.00%	1.12
Total	4,939	100.00%	0.86	16,426	100.00%	1.01

	China		
	Number	Ratio	AQ
Domestic papers only	1,259	25.49%	0.93
International papers only	138	2.79%	0.84
Both international and domestic papers	495	10.02%	1.21
Total	1,892	38.31%	1.00

Table 31 shows the indicators of performance and internationalization of researchers. We used indicators described in section 2.2 in order to evaluate researchers' performance and internationalization among the four countries. The average number of papers (P) per Japanese author was 5.21, whereas the average P of all countries as a whole was 3.67. Japanese P was also higher than that of the U.K. and the U.S., and almost same as that of China. Total contribution to publications (FP) by Japanese authors was 1.50, which was similar to that of the U.S. and China, and bigger than that of the U.K.

The standardized average number of citations (AQ) per Japanese author was 0.82, whereas the average of the total as a whole was 0.86. Japanese AQ was little bit smaller than that of the U.K. (0.86), and smaller than that of the U.S. (1.01) and China (1.00). Indicators of internationalization such as the ICI, ICR, and ICA of Japanese papers were quite similar to those of the U.S. but less than those of the U.K. and China, and less than those of all papers as a whole; therefore the internationalization of Japanese authors was lower than the world average, that of the U.K., and that of China in the field of chemistry.

Table 31 Descriptive statistics of researcher indicators in performance and internationalization

	All papers（N=49,599)				Japan(N=4,170)			
Variable	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
P	3.67	7.52	1.00	297.00	5.21	9.70	1.00	237.00
FP	1.18	2.49	0.04	95.35	1.50	3.01	0.04	80.12
AQ	0.86	1.20	0.02	80.80	0.82	0.84	0.02	14.62
ICI	0.65	1.87	0.00	91.00	0.51	1.75	0.00	48.00
ICR	0.19	0.34	0.00	1.00	0.09	0.24	0.00	1.00
ICA	0.55	1.06	0.00	18.00	0.58	1.15	0.00	12.00

	United Kingdom(N=4,939)				United States(N=16,426)			
Variable	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
P	3.83	7.28	1.00	115.00	4.43	9.86	1.00	297.00
FP	1.20	2.35	0.05	50.58	1.55	3.39	0.04	95.35
AQ	0.86	1.52	0.02	80.82	1.01	1.37	0.02	80.32
ICI	0.94	2.21	0.00	41.00	0.54	1.98	0.00	91.00
ICR	0.23	0.36	0.00	1.00	0.11	0.26	0.00	1.00
ICA	0.67	1.17	0.00	12.00	0.57	1.19	0.00	16.00

	China(N=1,892)			
Variable	Mean	Std. Dev.	Min	Max
P	5.25	10.38	1.00	171.00
FP	1.48	3.23	0.08	63.48
AQ	1.00	0.92	0.03	13.62
ICI	0.83	2.47	0.00	48.00
ICR	0.16	0.29	0.00	1.00
ICA	0.82	1.29	0.00	18.00

Relationship between researchers' performance and international papers

Table 32 shows the results of Spearman correlations between indicators (H_0: No correlation between two indicators; coefficient = 0) of the four countries. As shown earlier in Table 18, the correlation analysis in Table 32 indicates a link between productivity and international collaboration for all four countries. Although there are small differences between the countries, for instance, the relationship between ICR and FP in the U.K is almost half that of Japan and the U.S., overall tendencies, such as the relationship between ICR and P being around 0.3 is similar among the four countries.

Table 32 Spearman correlations between indicators

【Japan】

	P	FP	AQ	ICI	ICR	ICA
P	1	0.8775***	0.2820***	0.4014***	0.3203***	0.5113***
FP		1	0.2229***	0.3310***	0.2484***	0.4632***
AQ			1	0.1504***	0.1245***	0.2653***
ICI				1	0.9816***	0.7266***
ICR					1	0.6961***
ICA						1

【United Kingdom】

	P	FP	AQ	ICI	ICR	ICA
P	1	0.7836***	0.2205***	0.5480***	0.3117***	0.5578***
FP		1	0.1173***	0.3645***	0.1274***	0.4242***
AQ			1	0.1829***	0.1244***	0.1997***
ICI				1	0.9238***	0.8256***
ICR					1	0.7368***
ICA						1

【United States】

	P	FP	AQ	ICI	ICR	ICA
P	1	0.7658***	0.2126***	0.4526***	0.3628***	0.5409***
FP		1	0.1076***	0.3111***	0.2174***	0.4426***
AQ			1	0.1329***	0.1080***	0.2058***
ICI				1	0.9790***	0.7383***
ICR					1	0.7032***
ICA						1

【China】

	P	FP	AQ	ICI	ICR	ICA
P	1	0.8637***	0.2980***	0.4841***	0.3151***	0.5906***
FP		1	0.2382***	0.3849***	0.2144***	0.5460***
AQ			1	0.1697***	0.1205***	0.2476***
ICI				1	0.9501***	0.7504***
ICR					1	0.6878***
ICA						1

3.3 International comparison of international mobility

The ratio of researchers who had experience of staying in multiple counties and conducting research internationally was similar among three countries: 8.99% in Japan, 12.13% in the U.K., and 10.77% in the U.S. On the other hand, the ratio of Chinese researchers who had experience of staying in multiple countries and conducting research internationally was more than double that of these three countries (26.37%).

Table 33 Internationally mobile researchers

Number of countries	All papers	Japan	United Kingdom	United States	China
Single	46,160	3,795	4,340	14,657	1,393
Multiple	3,439	375	599	1,769	499
Ratio of multiple countries	6.93%	8.99%	12.13%	10.77%	26.37%

A common factor in the four countries was that almost sixty to seventy percent of researchers who had experience of staying in multiple counties and conducting research internationally had published papers through both international and domestic collaboration (Japan 72.80%, the U.K. 69.45%, the U.S. 65.57%, and China 61.52%). On the other hand, almost eighty percent of researchers who did not have experience of staying in multiple counties and conducting research internationally had published papers only through domestic collaboration (Japan 83.08%、the U.K. 67.67%, the U.S. 83.03%, and China 77.39%).

Table 34 International mobility and international collaboration

		Domestic papers only		International papers only		Both international and domestic papers	
Country name	Number of Countries	Number	Ratio	Number	Ratio	Number	Ratio
Japan	Single	3,153	83.08%	174	4.58%	468	12.33%
	Multiple	87	23.20%	15	4.00%	273	72.80%
United Kingdom	Single	2,937	67.67%	597	13.76%	806	18.57%
	Multiple	129	21.54%	54	9.02%	416	69.45%
United States	Single	12,169	83.03%	855	5.83%	1,633	11.14%
	Multiple	515	29.11%	94	5.31%	1,160	65.57%
China	Single	1,078	77.39%	127	9.12%	188	13.50%
	Multiple	181	36.27%	11	2.20%	307	72.80%

Table 35 shows the results of binary logistic regression of international movement versus performance indicators in three countries. In Japan, researchers who had experience of staying in multiple counties and conducting research internationally had produced more papers in total, more papers through international collaboration, and had higher AQ scores. On the other hand, in the U.K., U.S. and China, researchers who conducted research internationally had produced more papers, but had not necessarily published papers with higher levels of quality.

Table 35 Binary logistic regression of international movement versus performance indicators

	Japan				United Kingdom			
	Coef.	Std. Err.	z	P>z	Coef.	Std. Err.	z	P>z
ICI(Dummy)	2.574	0.135	19.03	0	1.677	0.110	15.2	0
P	0.027	0.005	5.52	0	0.066	0.006	10.61	0
AQ	0.155	0.067	2.31	0.021	0.026	0.025	1.03	0.304
Cons	-3.831	0.127	-30.08	0	-3.291	0.094	-34.83	0
	Number of observations: 4,170 Pseudo R2: 0.243				Number of observations: 4,939 Pseudo R2: 0.1708			

	United States				China			
	Coef.	Std. Err.	z	P>z	Coef.	Std. Err.	z	P>z
ICI(Dummy)	2.328	0.059	39.29	0	1.492	0.120	12.45	0
P	0.019	0.002	8.87	0	0.057	0.008	6.8	0
AQ	0.009	0.022	0.4	0.689	0.082	0.062	1.33	0.183
Cons	-3.226	0.050	-64.17	0	-2.056	0.104	-19.74	0
	Number of observations: 16,426 Pseudo R2: 0.1988				Number of observations: 1,892 Pseudo R2: 0.1539			

Acknowledgement

This work was supported by Grants-in-Aid for Scientific Research (KAKENHI) (22500238).

DISCUSSION PAPER No.95

Empirical analysis on factors associated with international scientific collaboration

2013 年 7 月

文部科学省　科学技術・学術政策研究所
第 1 調査研究グループ
加藤　真紀
〒100-0013
東京都千代田区霞が関 3-2-2 中央合同庁舎第 7 号館　東館 16 階
TEL:03-3581-2395　FAX:03-3503-3996