Hitotsubashi University
Institute of Innovation Research

Contributions of young scholars in team-based scientific research

Masatsura Igami
Sadao Nagaoka
John P. Walsh

IIR Working Paper WP#13-02

Jan 2013

Institute of Innovation Research
Hitotsubashi University
Tokyo, Japan
http://www.iir.hit-u.ac.jp

Contributions of young scholars in team-based scientific research

Masatsura IGAMI[1, *], Sadao NAGAOKA[2] and John P. WALSH[3]

[1] *National Institute of Science and Technology Policy, 3-2-2 Kasumigaseki, Chiyoda-ku, Tokyo, 100-0013 (Japan)*

[2] *Institute of Innovation Research, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo, 186-8603 (Japan)*

[3] *School of Public Policy, Georgia Institute of Technology, 685 Cherry Street, Atlanta, GA 30332-0345 (USA)*

January 2013

Abstract

How to design and manage a research team has become an increasingly important issue in knowledge creation in science. This paper aims to understand how young scholars contribute to scientific research in the context of a research team. We have found that the likelihood of the involvement of postdoctoral fellows is high in research teams working on a rapidly advancing research theme, facing competitive threat, and in those research teams with foreign-born scholars and when exploiting advanced research equipment. Focusing on the papers where the order of the authors follows contribution, the probability of postdoctoral fellows becoming the first authors is more likely to exceed that due to a random assignment in the research teams facing competitive threat, with foreign-born postdoctoral scholars and when exploiting advanced research equipment. Finally, we have found that the involvement of postdoctoral fellows is positively associated with research performance in terms of citation counts and with shorter time to the publication even if the size of project is controlled for. The finding that postdoctoral fellows accelerate the speed of research is consistent with the fact that they become more likely to be the first authors in the face of strong scientific competition. These findings illuminate how young scholars contribute to the team-based research.

JEL codes: O30, D23

Keywords

Team; Young scholars; Science; First Author; Competition

* Address for correspondence:
National Institute of Science and Technology Policy
3-2-2 Kasumigaseki, Chiyoda-ku, Tokyo 100-0031, Japan
E-mail: igami@nistep.go.jp
Tel: +81-3-6733-4910, Fax: +81-3-3503-3996

1. Background and purposes

Teams are increasingly the fundamental unit of knowledge creation in science. Recent studies show that the number of authors in scientific publications has been increasing over time (Adams et al., 2005; Wuchty, Jones & Uzzi, 2007). These trends indicate that how to design and manage a research team has become an increasingly important issue in knowledge creation in science. In this context, it is important for us to deepen our understanding of how young scholars, especially postdoctoral fellows, contribute to the team-based research. In both the US and Japan, the number of postdoctoral fellows increased sharply in recent years.[1] There are several studies that shed light on the importance of young scholars in scientific teams. Larivière (2010) found that about 30% of Quebec university papers have PhD students as authors. Black & Stephan (2010) pointed out that PhD students and postdoctoral fellows appear disproportionately more as first authors in US articles in the journal *Science*.[2] Although this latter result indicates the significance of demand side reason for young scholars in science, it does not explain the sources

In this paper we aim at exploring the sources of the contributions of young scholars, in particular, postdoctoral fellows, to team-based scientific research, based on large-scale original survey data. We can identify several potential reasons why they are important for enhancing the performance of research under certain conditions. First, postdoctoral fellows who have recently completed their PhD training are likely to have the most advanced knowledge and technique to conduct competitive research. Their participation may become especially important in those areas where science moves very fast. Second, they can contribute to accelerating the speed of research

[1] In the US, the number of postdoctoral fellows, especially in life sciences, increased sharply with the doubling of the NIH budget. Many of them, however, have had difficulty to find tenure-track positions in higher education institutions (Stephan, 2012). The situation is similar in Japan. In the first S&T basic plan (FY1996-2000), the Japanese government introduced a program to increase postdoctoral fellows up to 10,000 in order to strengthen Japan's research capability.

[2] It is important to note, however, that the order of the authors do not necessarily follow the order of contributions. One novelty of our study is that we recognized the first author as the largest contributor only when the authors are ordered according to their contributions.

since they can devote their time fully to the research. Thus, they may play an important role when speed maters in obtaining priority in research (Merton, 1973). Third, they are more mobile internationally (see section 4 for direct evidence), so that hiring a postdoctoral fellow in a research team is an efficient way to engage foreign talent (Empirica, 2005; Franzoni, Scellato & Stephan, 2012). The contribution of foreign-born scholars in knowledge creation in science is especially large in the US (Stephan & Levin, 2001), where high quality research universities attract students from across the world. Fourth, research facilities/equipment play very important roles for some scientific research (see Stephan, 2012 for its increasing role). For examples, the inventions and the progress of a particle accelerator, a scanning tunneling microscope, and a DNA sequencer have been major sources for advancing of research in physics, materials science and life sciences. Young scholars may be vital in exploiting these cutting-edge facilities/equipment.

This paper aims to understand how these potential channels of the contributions of young scholars affect the involvement of young scholars as well as their (relative) contributions in the team measured by how frequently they appear as the first authors when the order represents relative contributions. Furthermore, to the extent that young scholars contribute fundamentally in advancing research, we expect that their presence is associated with high research performance. We will verify such prediction.

The results of a Japan-US scientists' survey on the knowledge creation process in science are used for the analysis, in addition to bibliographic information. The latter alone provides only limited information on who are the authors, so that their academic positions, country of origin, and disciplinary diversities as well as whether the order of the authors follow their contributions are unknown. The survey in Japan was jointly conducted by the Institute of Innovation Research (IIR) of Hitotsubashi University and the National Institute of Science and Technology Policy (NISTEP) from the end of 2009 to the summer 2010 (Nagaoka et al., 2010). The survey in the US was implemented by the Georgia Institute of Technology, in collaboration with IIR and NISTEP,

from autumn 2010 to early 2011 (Nagaoka et al., 2011). It collected around 2,100 responses from scientists in Japan and 2,300 responses from scientists in the US. The survey collected information on the research projects that generated the papers that were the target of the surveys. One advantage of the scientist survey is that it is a comprehensive and standardized micro-data set, covering the characteristics of research projects, the composition of the research team, research funding source used in the research projects, external knowledge sources that inspired the research project, serendipities in the research projects, outputs yielded by the research projects among others. In addition to coverage of both the US and Japan, the scientist survey also collected data on research projects that yielded top 1% highly cited papers and the research projects that yielded other papers (normal papers). The rich data regarding research projects enable us to analyze the relationship between research team characteristics and performance of the research teams.

Based upon the dataset obtained by the scientist survey, the ultimate goal of our study is to address the following questions regarding the research teams.

1. To what extent are young scholars involved in the research team? What determines their involvement?

2. What are the sources of the contributions of young scholars to team-based science research measured by the frequency with which they appear as the first authors when the order of the authors represents their relative contribution?

3. Is the presence of young scholars associated with higher research performance, controlling for project size and research budget?

This paper is organized as follows: Section 2 presents an overview of the Hitotsubashi-NISTEP-Georgia Tech scientist survey; Section 3 shows the descriptive statistics regarding the involvement of young scholars in research teams; Section 4 discusses hypotheses on the determinants of the participation of young scholars in research teams and their contributions.

Section 5 shows the results of the econometric analysis; Section 6 concludes and summarizes some remaining research issues.

2. Overview of the surveys

2.1. Identification of focal papers

The population of the survey is articles and letters in the Web of Science database of Thomson Reuters. Reviews were excluded from the population. The time window of the papers for the survey is from 2001 to 2006 (in the database year). Database year refers to the year when the documents are recorded into the database of Thomson Reuters, which generally corresponds to the publication year. The bibliographic information and the number of citations as of the end of December 2006 were used in the identification of the focal papers. Two sets of focal papers were selected from the population as below:

Highly cited papers (approximately 3,000 in each survey): Top 1% highly cited papers in each journal field and in each database year; with at least one organization of authors should be located in Japan for the Japanese survey and in the US for the US survey. All highly cited papers in the time window were selected for the Japanese survey and approximately 3,000 highly cited papers were randomly selected from the highly cited papers in the US survey.

Normal papers (approximately 7,000 in each survey): Randomly selected papers in each journal field and in each database year from the population of the survey, excluding the above highly cited Top 1% papers; with at least one organization of authors should be located in Japan for the Japanese survey and in the US for the US survey.

We covered all science fields, including the social science, although the coverage of social science journals by the database is not comprehensive and we have a relatively small number of the publications in Japan in this field. Social sciences were dropped from the rest of the

analysis, given the small sample in Japan and the differences of research activities with natural sciences.

2.2. Identification of survey targets and research projects for the survey

Corresponding authors or equivalents of approximately 20,000 possible focal papers were searched and identified as survey targets. If multiple papers were assigned to a single corresponding (or equivalent) author, one paper was randomly selected as a focal paper while the priority was given to the highly cited papers in the selection process. As a result, 7,652 survey targets were identified for Japanese survey. Of those, there are 1,932 scientists whose focal paper is the highly cited paper; and there are 5,720 scientists whose focal paper is the normal paper. In the US, 8,864 survey targets were identified. Of those, there are 2,882 scientists whose focal paper is the highly cited paper; and there are 5,982 scientists whose focal paper is the normal paper.

2.3. Response rate

Out of 7,652 survey targets, we got 2,081 responses in the Japanese survey. The total response rate was 27.2%. The response rate was 29.3% for the highly cited papers and 26.5% for the normal papers. The total response rate in the US survey was 26.3%. We got 2,329 responses out of 8,864 survey targets. The response rate was 27.7% for the highly cited papers and 25.6% for the normal papers.

2.4. Field classification for the analysis

Results of the survey to be presented in this paper are based on 22 ESI journal fields. Some results are based on 3 broad fields obtained by further aggregation, i.e. physical sciences, life sciences, and medicine. Natural sciences aggregate physical sciences, life sciences, and medicine. Papers of multidisciplinary fields, those published in the journals like *Nature* and

Science, are reclassified into one of 21 fields based on the backward citations of the multidisciplinary papers.

2.5. Data available in the scientist survey

We have constructed a comprehensive and standardized micro-data set from the two surveys, covering the characteristics of research projects, the composition of the research team, research funding used in the research projects, external knowledge sources that inspired the research project, serendipities in the research projects, outputs yielded by the research projects among others.

As for the composition of the research team, our surveys asked a respondent to identify the authors' academic positions in their organizations, academic field of expertise, the country of birth, specialized skill, and the sector of the organization with which the author was affiliated. This question on the author profile was asked for all authors when the number of authors is 6 or less and for up to 6 authors, the first, last and corresponding authors plus randomly selected authors, when the number of authors is 7 or more. In addition to the information obtained by the survey, information on the number of authors and affiliations of authors were retrieved from the Web of Science. This paper focuses on research teams in higher education institutions (HEIs) and in natural sciences.

3. Participation of young scholars in the research teams

A research team can consist of a wide variety of human resources, such as graduate and undergraduate students, postdoctoral fellows, technicians, and professors or other scientists. There are several studies that show the importance of young scholars in the knowledge creation process (Larivière, 2010; Black & Stephan, 2010), however these analyses are limited to a specific journal or country and they do not examine the reasons for the participation of young scholars. In this section we first extend their analysis by covering all Web of Science journals

and by focusing on the publications. The order of the authors when this is linked to their contributions is also considered in the analysis of the first authors.

Figures 1 (a) and (b) show the composition of authors, in HEIs, in Japan and the US respectively. As for Japan, the share of professors is the largest, followed by associate professor and assistant professors. Professors account for around 38% of all authors in the highly cited and 35% in the normal papers. On the other hand, young scholars, who are undergraduates, graduate students, or postdoctoral fellows, account for 27% and 26% of the authors of highly cited papers and normal papers, respectively. In the US, professors also account for the largest share in both the highly cited and normal papers. Young scholars (MA and/or undergraduate students; PhD students; and postdoctoral fellows) account for 34% of the authors of the highly cited papers and 33% of the normal papers in the US. The composition of MA and/or undergraduate student; PhD student; and postdoctoral fellow differs between Japan and the US. The share of MA and/or undergraduate students is greater in Japan than in the US.

Table 1 summarizes the percentage of young scholars as the first authors of the focal papers, limiting to those papers in which authors are listed in order of their degree of contributions. In other word, it indicates the types of scientists who made the most contributions to the focal papers.

We first look at the contribution of young scholars in the normal papers. As shown in Tables 1 (a) and (b), young scholars account for 35% of the first authors in Japan and for 49% of the first authors in the US. Young scholars appear as the first authors more frequently in the US than in Japan. The share of young scholars as the first authors is much larger than for all authors in both countries (compare Table 1 and Figure 1). The contribution of young scholars as the first authors is especially large in life sciences in both countries. In life sciences, 45% and 61% of the first authors are young scholars in Japan and the US, respectively. In physical sciences, the share of young scholar first authors is around 30% in Japan and more than 50% in the US.

The rate of young scholar first authors is even greater for highly cited papers, especially in Japan: increasing from 35% of normal papers to 39% of highly cited papers in natural sciences in Japan and from 49% to 51% in natural sciences in the US. The share of postdoctoral fellows in the highly cited papers is very large in life sciences both in Japan and the US. As we can see from Table 1, the composition of postdoctoral fellows and students as the first authors is different by the type of papers. The participation of students as first author is more common in the normal papers compared to the highly cited papers in both countries.

4. Hypotheses and descriptive statistics

First, we focus on the determinants of the participation of postdoctoral fellows in a research team and of their contributions. Although we do not have data directly measuring the contribution of a postdoctoral fellow, we assess his/her contribution by examining whether a particular factor can account for the probability of a postdoctoral fellow being a member of the research team and becoming the first author of the paper respectively. Given that a postdoctoral fellow is more likely to be engaged in a research project if his/her contribution is high, we expect that the same set of factors will serve as the determinants for both their participation and their contribution.

4.1. Sources explaining the participation of postdoctoral fellows in a research team and their contributions to research

4.1.1. Novelty and competitive threat

Given that postdoctoral fellows are likely to have the most recent knowledge and technique to conduct a research, it is likely that they can contribute most when the novelty of the research theme is high. Thus, we propose the following hypothesis regarding the relationship

between the participation and the contribution of postdoctoral fellows and the novelty of the research theme.

*Hypothesis 1: P*ostdoctoral fellows are more likely to participate in the research team and their contributions in the research team tend to be high, when the novelty of research theme is high.

We use the average citation time-lags between the publication year of the focal paper and the publication years of the references therein as a proxy to measure the novelty of the research theme. The focal papers with long citation time-lags are considered the outputs from the research on established research theme, while short citation time-lags can be considered to indicate a research area in rapid development. Figure 2 shows box-plots of the citation time-lags by four types of teams: only senior scholars; student(s) involved; postdoctoral fellow(s) involved; and both student(s) and postdoctoral fellow(s) involved. The median citation time-lag (6.77 years) is the longest in research teams consisting of only senior scholars. This median citation time-lag is comparable to that in research teams consisting of senior scholars and student(s) (6.75 years). The median citation time-lag is shorter by about 1 year in the research teams in which at least one postdoctoral fellow participated (5.74 years for research teams in which only postdoctoral fellow(s) participated, 5.58 years for those where both student(s) and postdoctoral fellow(s) participated). In the regression analyses, citation time-lag is normalized by field average in order to eliminate the field dependency of the lag.

The level of competition is also considered. Scientific research is characterized as a competitive process of gaining reputation by establishing priority (Merton, 1973). Postdoctoral fellows can contribute to accelerating the speed of research since they can devote their time fully to the research. Professors spend a significant amount of time for teaching, university administration, developing project proposals and getting funds, and research management (Decker et al., 2007). Students spend significant time for learning and for their dissertations. Thus, postdoctoral fellows may play an important role when speed matters in obtaining priority in research.

Hypothesis 2: Postdoctoral fellows are more likely to participate in the research team and their contributions in the research team tend to be high, when the research team faces a significant competitive threat of being scooped.

4.1.2. Engaging foreign talent and expertise

Young scholars are more mobile internationally (see the evidence provided below), so that hiring a postdoctoral fellow in a research team is an efficient way to engage a foreign talent. The contribution of foreign-born scholars in knowledge creation in science is especially large in the US (Stephan, 2012). We propose the following hypothesis on the relation between internationalization of the research team and the participation and contribution of young scholars.

Hypothesis 3: Postdoctoral fellows are more likely to participate in the research team and their contributions in to the research team tend to be high, when the research project is international.

Our scientist survey asked the respondents the country of birth of up to 6 authors. The combination of the information of international co-authorship and of country of birth of authors enables us to differentiate the impact of two aspects of internationalization on research activities: engaging foreign institution and engaging directly a foreign born scholar. Table 2 summarizes the results. In Japan, around 85% of the domestic papers involve no foreign-born scientists in the authors. In contrast, more than 90% of the internationally co-authored papers involve foreign-born scientists. These observations indicate that the majority of foreign-born scientists in the Japanese survey were affiliated with an organization outside of Japan. In contrast, about 70% of domestic papers involve at least one foreign-born scientist in the US.

Figure 3 shows the distribution of the country of birth of the authors by their academic positions. Only domestic papers are analyzed in order to exclude the influence of the foreign-born scholars affiliated with foreign institutions outside of Japan or the US. The percentage of the foreign-born scholars varies across the academic position in both countries. The share is the highest in the postdoctoral fellows, followed by PhD students. These results indicate that

postdoctoral fellows and PhD students are significantly more mobile than those with other academic positions both in Japan and the US.

4.1.3. Utilization of cutting-edge facilities/equipment

Research facilities/equipment plays a very important role for scientific research. Young scholars typically play an important role for operating these machines for research, since the operation often requires investing time in acquiring new skills. Thus, the adoption of advanced research equipment would be an important motivation for hiring young scholars.

Hypothesis 4: Postdoctoral fellows are more likely to participate in the research team and their contributions in to the research team tend to be high, when the state-of-the-art facilities/equipment is owned by a research team.

In the Japanese survey, we asked respondents whether they used the state-of-the-art research facilities/equipment such as a supercomputer, a telescope, and a particle accelerator. We also differentiated between those owned by a research team and the external ones.

Table 3 summarizes the level of the use of facilities/equipment in Japanese survey. The percentage of the research teams that utilized the internal or external state-of-the-art facilities/equipment is the smallest in research teams consist of only senior scholars (33% for internal, 15% for external). The likelihood of the utilization of the internal state-of-the-art facilities/equipment increases by around 20% points in research teams with postdoctoral fellow(s). The probabilities of the utilization of the external state-of-the-art facilities/equipment also increase; however, the degree of increase is not so significant compared with that in the internal facilities/equipment, perhaps because there are specialized staffs operating the external facilities/equipment.

4.2. The participation of postdoctoral fellows in a research and research performance

If a postdoctoral fellow contributes to research performance significantly as suggested by a high frequency of them becoming the first author, we expect that the performance of the research project is high, when it involves a postdoctoral fellow. This will be the case even if we control for the endogeneity that high quality research tends to be large so that it tends to hire a postdoctoral fellow.

Hypothesis 5: The research project with a postdoctoral fellow performs well, even if we control for the size of the project.

5. Estimation of the involvement of young scholars and their impact on research outputs

This section presents the results of econometric estimations, which test the above hypotheses. First we analyze how the involvement of postdoctoral fellows and that of the students can be (separately) accounted for by the characteristics of the research project and the teams (section 5.1). Second we analyze how the likelihood of the young scholars becoming the first authors can be accounted for by those characteristics (section 5.2). In this context, one strong advantage of our survey is that it enables us to analyze the focal papers in which the first authors were the most important contributor. Finally, we assess the impact of young scholars' participation on research outputs using two output indicators (citations and the duration to the publication, section 5.3). Table 4 summarizes the variables used in the economic models. Table 5 provides the summary statistics of the variables. In the following estimations we eliminate the solo-authored papers from our samples, so that our analysis focuses on team- based research.

5.1. Participation of young scholars in the research team

The likelihood of the participation of young scholars (postdoctoral fellows and students separately) into a research team is estimated based on a logit regression by incorporating the variables representing the above hypotheses as independent variables. In this estimation, we pool the sample of "normal papers" and "highly cited papers", given that there is no obvious difference between these two types of the papers on the likelihood of the participation of young scholars, given our set of independent variables (in particular, project size).[3] The model for estimation is given by the following:

$$\begin{aligned} \textit{Indicator variable} = {} & \beta_1(\textit{citation\_lag}) + \beta_2(\textit{threat}) + \beta_3(\textit{foreign\_talent}) \\ & + \beta_4(\textit{foreign\_inst}) + \beta_5(\textit{equipment\_owned}) + \beta_6(\textit{equipment\_ext}) + \delta(\textit{controls}) + \varepsilon \end{aligned} \quad (1)$$

Two indicators, or dependent variables, are the following: One is the dummy variable regarding the participation of postdoctoral fellows ("POSTDOC_PARTICIPATION") and another is the dummy variable regarding the participation of students ("STUDENT_PARTICIPATION"). Each variable stands at 1 for participation and at 0 otherwise.

The main explanatory variables are the following. Citation lag ("LG_CITATION_LAG") is the log of the average citation time-lags between the publication year of the focal paper and the publication year of the references therein, normalized by science field average (that is, the percent difference from the field average). Threat (two dummies, "Moderately concerned" and "Concerned") are based on a subjective evaluation of the level of concern of the respondents about the possibility of competitors would have priority over research results (Hagstrom, 1974). The variable stands at 2 ("Concerned") if the level of concerned is high, at 1 ("Moderately concerned") if the level of concern is moderate, at 0 if a respondent was not at all concerned about the priority loss.[4] Foreign talent ("Domestic paper with foreign-born scholars") is a

[3] The results using only normal papers are similar to those presented here. They are available from the authors on request.

[4] In the actual estimation, the variable is recoded as two 1/0 dummies.

dummy representing the case where all authors belong to the domestic institutions but foreign-born scholars exit in the authors. Foreign institution ("International paper") is a dummy representing the case where both domestic and foreign institutions exist in the paper as the affiliations of the authors.

We also introduce the following control variables. The first is the diversity of the field of expertise of the research team. The combination of knowledge across the fields of science is getting crucial in the knowledge creation (Saka, Igami & Kuwahara, 2010). Stephan (2012 and references therein) point out that increasing importance of interdisciplinary research and emerging disciplines encourages PIs (principal investigators) to collaborate with scientists with a different knowledge and skill set. Thus, there is a possibility that young scholars are hired in order to diversity the field of expertise of the research team, so that we control for such possibility by introducing the index of the diversity of expertise ("DIV_EXPERTISE"). In our survey, we collected information about the field of expertise of authors. It was found that the degree of diversity is different by the type of involvement of young scholars. The percentage of research teams that cover two or more fields of expertise is about 42% in research teams consisting of only senior scholars. The percentage increases by 10 percentage points when postdoctoral fellows are involved in the research team. The percentage is the lowest in research teams consisting of senior scholars and students.

We include the number of authors ("LG_AUTHORS") as a control variable in the regression analysis. The possibility of the involvement of young scholars would increase with an increasing number of authors on the focal paper. Research expenditure per man-month ("LG_EXP_MANMONTH") is also considered. As for the personnel expenditures, the surveyed amount includes only those for employing scientists and technicians specifically for the research projects, which are typically defrayed by extramural funds. For this reason, the survey mainly measures personnel cost on young scholars employed for the project. The likelihood of the involvement of young scholars would increase in the research projects that use more research

expenditure per man-month since more financial resources become available for hiring them for a given time period, and the effect would vary depending on positions of young scholars, since they have different financial needs.

Dummy variables for the fields of science ("FIELD21") are included to measure field dependency of the participation of young scholars. As we have seen in the previous section, the degree of participation of postdoctoral fellows and PhD students strongly depends on the field of science. Finally, a country dummy ("COUNTRY", 1 for the US) is also included.

Tables 6 and 7 show the results of estimation about the involvement of postdoctoral fellows and students, respectively. The coefficients presented are the marginal effects. We show the estimation results for the sample pooling the US and Japan and that of only Japan, for both robustness check and for the fact that only the survey in Japan had information on the utilization of facilities/equipment. In all models, the number of authors in the focal paper ("LG_AUTHORS") has a positive and statistically significant coefficient. The amount of research expenditure per man-month ("LG_EXP_MANMONTH") has a positive impact only on the participation of a postdoctoral fellow, which makes sense since the entire cost of a postdoctoral fellow often comes from the project research budget.

Controlling for these, citation time-lag ("LG_CIATION_LAG") has a negative and statistically significant coefficient in the Models 1-1, 1-3 with $p < 0.01$, and Model 1-4 with $p < 0.05$. Given that citation time-lag is the average differential between the publication year of the focal paper and of the references, the likelihood of the participation of postdoctoral fellows increases in research teams working in rapidly advancing research theme in each research field. The estimated coefficient is smaller but significant in the Japanese sample (Model 1-1 vs. Model 1-4), although it becomes insignificant when type of authorship is introduced (Model 1-7). The result supports the first part (on participation) of hypothesis 1, especially for the US. In contrast, citation time-lag has a *positive* and statistically significant coefficient in the Models 2-1, 2-3, and 2-4 for students. This may indicate that students are more involved in a research

project based on more established scientific base, since they are in the process of learning, while postdoctoral students are involved in an emerging science field.

Competitive threat on research priority ("THREAT") also increases significantly the probability of the participation of postdoctoral fellows ($p < 0.01$ for Models 1-1 and 1-3, and $p < 0.05$ for Model 1-4), but has no impact on the participation of students. The likelihood of the participation of postdoctoral fellows becomes higher when the level of concern to the priority loss is high in both the sample pooling the US and Japan and that of the Japanese sample, although the significance is lost in Model 1-8 where all independent variables are introduced. The result is also consistent with the first part (on participation) of hypothesis 2.

As for internationalization of the research team, the involvement of the foreign-born scholars in domestic paper (that is, the paper involving only domestic universities or research institutions) has a positive and statistically significant ($p < 0.01$) coefficient in all models. The results strongly support the first part (on participation) of hypothesis 3 for postdoctoral fellows. When foreign-born scholars participate in domestic research, they are often postdoctoral fellows (but not students) and they tend to have excellent research capability, due in part to the selection process (e.g., costs of going abroad). International co-authorship also increases the possibility of the participation of postdoctoral fellows in all models, while it has negative impact on the participation of students. International institutional collaboration facilitates the foreign postdoctoral fellows to participate in a research project.

The likelihood of the participation of postdoctoral fellows increases significantly if the research team owns the state-of-the-art facilities/equipment (Models 1-6 and 1-8, $p < 0.01$), consistent with hypothesis 4. Exploiting state-of-the-art facilities/equipment is an important reason for the participation of postdoctoral fellows who have specialized skill to operate them.

The number of fields of expertise ("DIV_EXPERTISE") has no significant correlation with the involvement of postdoctoral fellows or students. Diversity of field of expertise is more likely to be achieved through collaboration among senior scholars having different fields of

expertise rather than through participation of recruiting young scholars. Given that one of the crucial functions of research projects is training of young scholars, it would be quite likely that the field of expertise of young scholars coincides with that of the PIs. From the view point of postdoctoral fellows, it may be too risky for them to join in the research team led by a PI whose field expertise is different from that of the postdoctoral fellows.

5.2. Contribution of foreign young scholars as the first authors

The analyses in the last section showed that the participation of the postdoctoral fellows is high in the research team with foreign-born scholars, in the research team working on the competitive and rapidly advancing research theme, and in the research project exploiting advanced research equipment. This section analyzes more directly their contributions by analyzing whether the probability of the postdoctoral fellows becoming the first author is high if these characteristics are satisfied, when the order of the authors follows the levels of their contributions. We estimate the following model by OLS (Ordinal Least Squares Method):

$$\begin{aligned} \text{POSTDOC\_FIRST\_PROB} = {} & \beta_1(citation\_lag) + \beta_2(threat) + \beta_3(postdoc\_foreign\_\%) \\ & + \beta_4(equipment\_owned\_eff) + \delta(controls) + \varepsilon \end{aligned} \quad (2)$$

We introduce the dependent variable which shows whether the probability of postdoctoral fellows becoming the first authors exceed that when the first authorship is randomly assigned (the latter probability is simply equal to the share of postdoctoral fellows in the entire authors). Variable $\text{POSTDOC\_FIRST\_PROB} = \text{POSTDOC\_FIRST} - \text{POSTDOC\_PERCENTAGE}$ is the difference between POSDOC_FIRST, a dummy variable for whether a postdoctoral fellow is the first author or not and POSTDOC_PERCENTAGE, which is the percentage of postdoctoral fellows among the authors of the focal paper. If the postdoctoral fellow becomes the first author disproportionately, the average of POSTDOC_FIRST_PROB will be positive.

To assess the sources of the contributions of postdoctoral fellows, four independent variables are considered. They are log citation lag ("LG_CITATION_LAG"), competitive

threat ("THREAT"), the percentage of foreign-born among the postdoctoral fellows in the authors ("POSTDOC_FOREIGN_PERCENTAGE"), and a level of effectiveness of the owned facilities/equipment for the research ("ADV_EQUIP_OWNED_EFFECTIVE"). The POSTDOC_FOREIGN_PERCENTAGE is supposed to measure the importance of foreign young talent in increasing research productivity, and the ADV_EQUIP_OWNED_EFFECTIVE represents the effectiveness of the state-of-the-art facilities/equipment for the research. This latter variable stands at 2 if the facilities/equipment was used and very effective, at 1 if the facilities/equipment was used, but not effective, and at 0 if the facilities/equipment was not used.

OLS regression is used to estimate the above equation, using the pooled sample of "normal" and "highly cited papers".[5] In the estimations, we limited our samples to the focal papers in which the authors were ordered accordingly with their contributions. Furthermore, we had to limit our samples only to those with 2 – 6 authors and with one or more postdoctoral fellows, since we have information on the authors' characteristics only up to 6 authors. Due to a small number of the resulting samples (N = 394 for the US and Japan and N = 153 for Japan), we controlled for the effects of the field of science by FIELD3 instead of FIELD21. Note that we do not need to control for the project size and the number of the postdoctoral fellows, since we estimate whether the probability of the postdoctoral fellows becoming the first authors exceeds the probability due to a random assignment of the first author which is equal to the share of the postdoctoral fellows in the authors.

Results are shown in Table 8. The LG_CITATION_LAG has no significant correlation with the contribution of postdoctoral fellows. As shown in the last section, the probability of the participation of postdoctoral fellows is higher in a research team working on the rapidly advancing theme, but the novelty of research theme per se does not make a postdoctoral fellow the first author of the paper. On the other hand, the probability of a postdoctoral fellow becoming

[5] The results using only normal papers are similar to those presented here. They are available from the authors on request.

the first author is significantly high when competitive threat is high (significant in Models 3-1 and 3-4 with $p < 0.01$ and Model 3-7 with $p < 0.1$). It becomes insignificant only if all the other independent variables are introduced. The result supports the second part of hypothesis 2, but not the second part of hypothesis 1

The percentage of foreign-born among the postdoctoral fellows ("POSTDOC_FOREIGN_PERCENTAGE") has positive and statistically significant coefficients in Models 3-2 and 3-5 with $p < 0.01$ and in Model 3-3 $p < 0.1$. The likelihood of a postdoctoral fellow becoming the first author increases if he/she is foreign-born. The foreign-born postdoctoral fellows are not just mobile workforces; they are an important contributor of knowledge creation in science. The result supports the second part of hypothesis 3

The possibility of the postdoctoral fellows become the first author is higher than average if the facilities/equipment is effective in research (Model 3-6 with $p < 0.01$). This result supports the second part of hypothesis 4. One of the direct contributions of the postdoctoral fellows on the research may be their skill in operating the state-of-the-art facilities/equipment effectively to conduct the research.

5.3. Impact of young scholars on research performance

Finally we will assess the impact of young scholars' participation on research outputs. We use ordered citation counts of the focal papers as proxies to measure the quality of the focal paper and the time to the publication from a research project as a performance of the research project. We use only the "normal papers" as our sample for these estimations, since "highly cited papers" are selected by citation performance. The model for estimations is as follows:

$$Research\_performance = \beta_1(postdoctoral\_\%) + \beta_2(student\_\%) + \beta_3(foreign\_talent) + \beta_4(foreign\_inst) + \delta(controls) + \varepsilon \quad (3)$$

As for the ordered citation counts ("ORDERED_CITATION_COUNTS"), we first classified the focal papers by whether they are within the top 10%, 20%, 30%, and 40% in

forward citations as of the end of 2011. Base on the classification, 5, 4, 3, 2, 1 point were given to the top 10%, 20%, 30%, 40%, and other papers, respectively. The ordered citation counts were determined by each ESI's field and each database year in order to eliminate the difference in citation behavior by the field of science and the effects of the truncation.

The second performance measure is the duration of the research project to the publication of the focal paper ("LG_DURATION_PUBLICATION"). As we have seen in the previous section, the postdoctoral fellows are more likely to be involved in the research project focusing on the research theme in rapid advancement and facing competitive threat. Our survey identified how many years it took from the launch of the research project to the publication of the focal paper, by asking the scientists the year they actually started their research projects.

Ordered logit regression is used to estimate the impact of the involvement of young scholars on the ordered citation counts. Ordinary least square regression is used to estimate the duration from the launch of the project to the publication of the focal paper. In order to estimate the contribution of young scholars, the percentage of postdoctoral fellows among the authors ("POSTDOC_PERCENTAGE") and the percentage of students among the authors ("STUDENT_PERCENTAGE") are adopted as independent variables. We limited our samples only to the normal papers in the analysis. We introduce the same variables as controls as for the model explaining the participation of young scholars, in addition to the type of authorship. In particular, project size (the number of authors and the level of funding) controls for the quality of the research and the endogeneity of the share of the postdoctoral fellows.

Table 9 summarizes the results of the regression analysis. Even after the project size and the field of science are controlled for, the percentage of postdoctoral fellows among the authors of the focal papers ("POSTDOC_PERCENTAGE") has positive and statistically significant coefficients in the ordered citation counts ($p < 0.1$). The estimation also shows that the time to the publication declines with increasing the percentage of postdoctoral fellows in the team. These results support hypothesis 5.

The number of authors also has a positive impact on the ordered citation counts, as expected. International collaboration as well as the involvement of foreign-born scholars in domestic papers also have positive impacts on ordered citation counts. Positive impact of international co-authorship (Glänzel, 2001; Saka & Kuwahara, 2011) and the number of authors on citations (Glänzel el al., 2006; Leimu & Koricheva, 2005) is consistent with the results of previous literatures.

6. Conclusions and discussions

This paper aimed to analyze how young scholars contribute to team-based scientific research, based on a new original large scale survey of scientists in the US and Japan. Our study has confirmed that young scholars are important contributors for research efforts both in Japan and the US. The percentage of young scholars among the first authors is remarkably high in those papers in which the authors are ordered by their contribution. The contribution of young scholars varies across the fields of science. The likelihood of the young authors becoming the first author is high in life sciences. In the US, postdoctoral fellows are the first author on half of highly cited papers in life sciences. The postdoctoral fellows are more often the first authors of the top 1% highly cited papers than normal papers.

We have found that the likelihood of the involvement of postdoctoral fellows is high in research teams working on the rapidly advancing research theme, facing competitive threat in priority, those with foreign-born scholars, and in research projects exploiting advanced research equipment. Focusing on the papers where the order of the authors follows contribution, the probability of postdoctoral fellows becoming the first authors is more likely to exceeds that due to a random assignment in research teams facing competitive threat, with foreign-born postdoctoral scholars and when exploiting advanced research equipment. Thus, our analysis identified four significant sources of the contributions of postdoctoral fellows to the research and three of them are also significant for making them the first authors of the papers where the order

matters. Finally, we have found that the involvement of postdoctoral fellows is positively associated with research performance in terms of citation counts, as well as for shortening the time to the publication even if the size of project and internationalization of research team is controlled for. The finding that postdoctoral fellows accelerate the speed of research is consistent with the fact that they are more likely to be the first authors when the team faces competitive threat. These findings underpin importance of young scholars in the research team and illuminate how they contribute.

On the other hand, the number of fields of expertise has no significant correlation with the involvement of postdoctoral fellows and students. Diversity of field of expertise is more likely to be achieved through collaboration among senior scholars having different fields of expertise rather than through recruiting young scholars. The mechanism for collaboration across fields is another issue which we would like to address in a future study.

The recent comprehensive analysis on the mobility of scientist shows that postdoctoral fellows are the most mobile and the percentage of foreign postdoctoral fellows are high (more than 40%) in major developed countries (Franzoni, Scellato & Stephan, 2012; Van Noorden, 2012), consistent with our findings. Our analysis showed that foreign born postdoctoral fellows do not simply provide labor force for operating research equipment, even though this is also an essential part of the research. They are also important contributors of knowledge creation in science. Given increasing mobility of the postdoctoral fellows, how to attract young scholars will become a very important policy issue for enhancing the national capability in knowledge creation in science.

Acknowledgements

This research has been being conducted in collaboration between Institute of Innovation Research (IIR) of Hitotsubashi University and National Institute of Science and Technology Policy (NISTEP) as a part of "Industry-university-government joint research on the innovation process" with support of Special Funds for Education and

Research of The Ministry of Education, Culture, Sports, Science and Technology. This work has also been supported by JSPS KAKENHI 21243020. Georgia Institute of Technology implemented the survey in the US in collaboration with IIR and NISTEP. Paula Stephan significantly contributed to the design of our survey. We have benefitted significantly from the participants in conference on "The Organisation, Economics and Policy of Scientific Research" in February 2011 (Torino); EUSPRI Conference on "Path - breaking innovation. Understanding, managing and providing support for continuous radical change in science and innovation" in June 2012 (Milano); and 17th International Conference on Science and Technology Indicators in September 2012 (Montreal).

References

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

Black, G. & Stephan, P.E. (2010). The Economics of University Science and the Role of Foreign Graduate Students and Postdoctoral Scholars. In Charles T. Clotfelter (Eds.), *American Universities in a Global Market* (pp. 129-161). Chicago: University of Chicago Press

Decker, R.S., Wimsatt, L., Trice, A.G. & Konstan, J.A. (2007), A PROFILE OF FEDERAL-GRANT ADMINISTRATIVE BURDEN AMONG FEDERAL DEMONSTRATION PARTNERSHIP FACULTY, *A Report of the Faculty Standing Committee of the Federal Demonstration Partnership*

Empirica GmbH (2005), Post-Docs in the Life Science, *Paper prepared for the NetReAct project*. Dowloadable at: www.netreact-eu.org/documents/NetreactDeliverable1.3.pdf

Franzoni, C., Scellato, G. & Stephan, P. (2012), FOREIGN BORN SCIENTISTS: MOBILITY PATTERNS FOR SIXTEEN COUNTRIES, *NBER WORKING PAPER SERIES*, Working Paper 18067

Franzoni, C., Scellato, G. & Stephan, P. (2012), Patterns of international mobility of researchers: evidence from the GlobSci survey, *Paper prepared for the International Schumpeter Society Conference.* Dowloadable at: http://www.aomevents.com/media/files/ISS%202012/ISS%20SESSION%207/Scellato.pdf

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

Glänzel, W., Debackere, K., Thijs, B. & Schubert, A. (2006). A concise review on the role of author self-citations in information science, bibliometrics and science policy, *Scientometrics*, 67, 263–277.

Hagstrom, W.O. (1974), Competition in science. *American Sociological Review*, 39, 1-18.

Jones, B. F. (2009), The burden of knowledge and the death of the Renaissance man. *Review of Economic Studies*, **76**, 283–317.

Merton, R.K. (1973), *The Sociology of Science: Theoretical and Empirical Investigations*. University of Chicago Press, Chicago, II

Nagaoka, S., Igami, M., Eto, M. & Ijichi, T. (2010). Knowledge Creation Process in Science: Basic findings from a large-scale survey of researchers in Japan. *IIR Working Paper*, WP#10-08

Nagaoka, S., Igami, M., Walsh, J.P. & Ijichi, T. (2011). Knowledge Creation Process in Science: Key Comparative Findings from the Hitotsubashi-NISTEP-Georgia Tech Scientists' Survey in Japan and the US. *IIR Working Paper*, WP#11-09

National Institute of Science and Technology Policy (2011). Japanese Science and Technology Indicators 2011

Lariviere, V. (2010). On the shoulders of students? A bibliometric study of PhD students' contribution to the advancement of knowledge, *Abstracts of 11th International Conference on Science and Technology Indicators* (pp. 155-157).

Leimu, R. & Koricheva, J. (2005). What determines the citation frequency of ecological papers?, *Trends in Ecology & Evolution*, 20, 28–32.

Saka A., Igami M. & Kuwahara T. (2010). Science Map 2008. National Institute of Science and Technology Policy, NISTEP REPORT No.139.

Saka, A., Kuwahara, T. (2011). Benchmarking Scientific Research 2011, National Institute of Science and Technology Policy, Research Material No. 204.

Stephan, P.E. (2012) *How Economics Shapes Science*. Harvard University Press

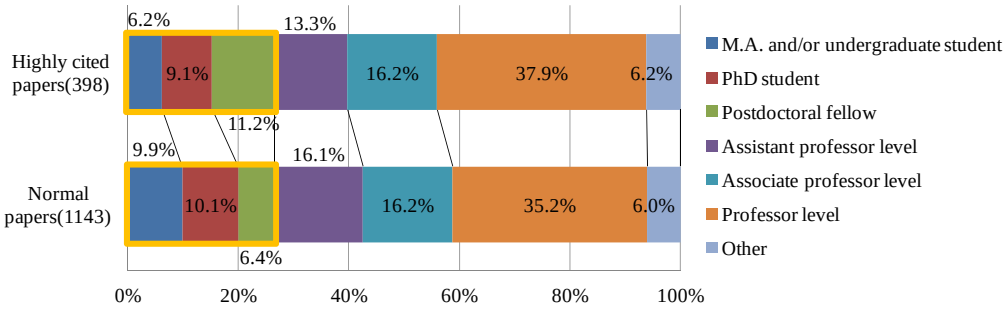
Stephan, P. E., & Levin, S. G. (2001). Exceptional contribution to US science by the foreign-born and foreign-educated. *Population Research and Policy Review*, 20, 59–79.

Van Noorden, R. (2012). Global mobility: Science on the move. *Nature* 490, 326–329.

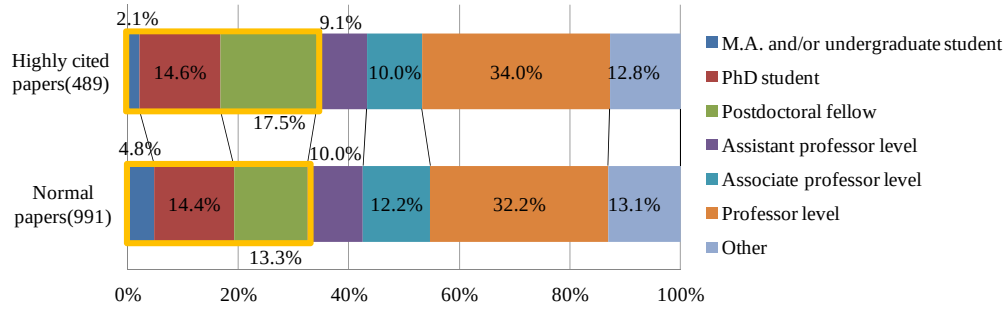
Wuchty, S., Jones, B. & Uzzi, B. (2007). The increasing dominance of teams in the production of knowledge. *Science*, 316, 1030–1036.

Figure 1. Compositions of authors in academic position (a paper basis, natural sciences)

(a) Higher education institutions, Japan (natural sciences)



(b) Higher education institutions, US (natural sciences)



Note: “Other” includes technician, the others and unknown.

Figure 2. Citation time-lags by the type of the involvement of young scholars

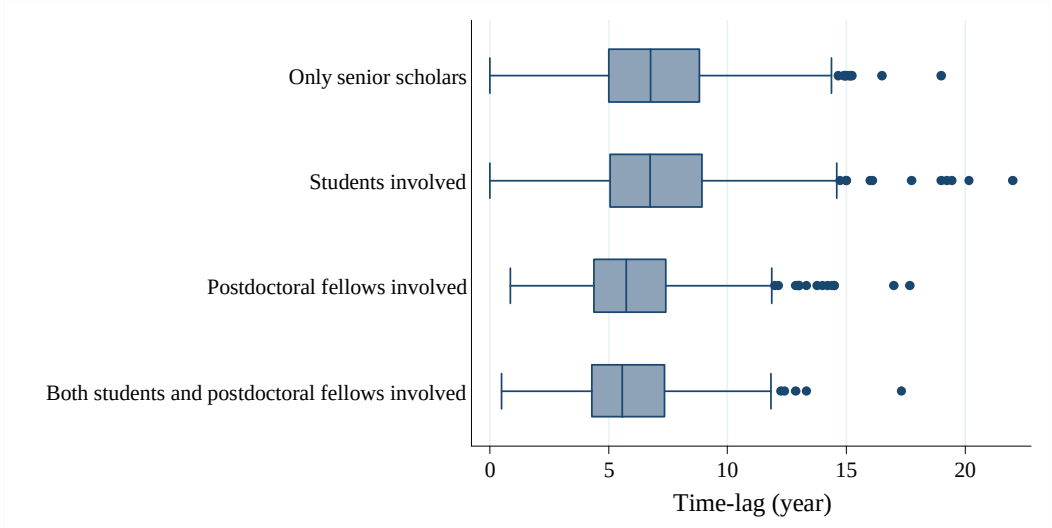
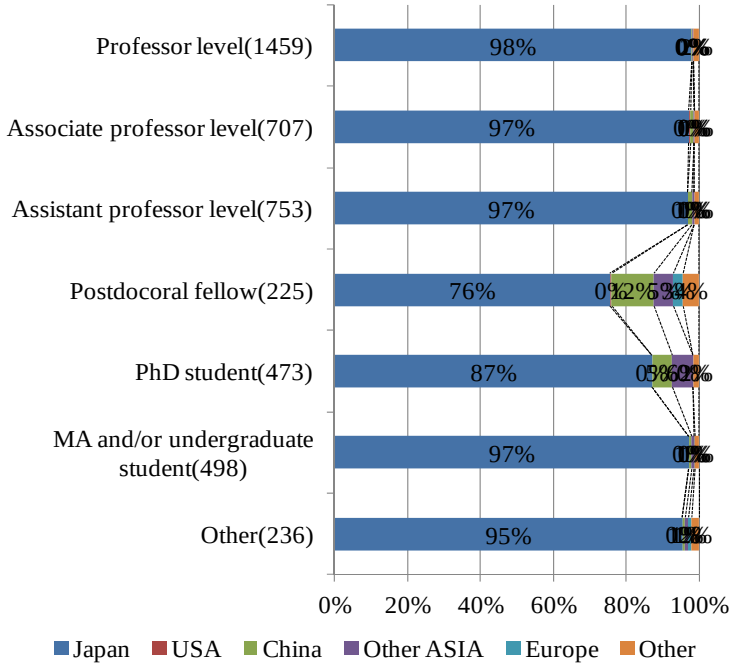
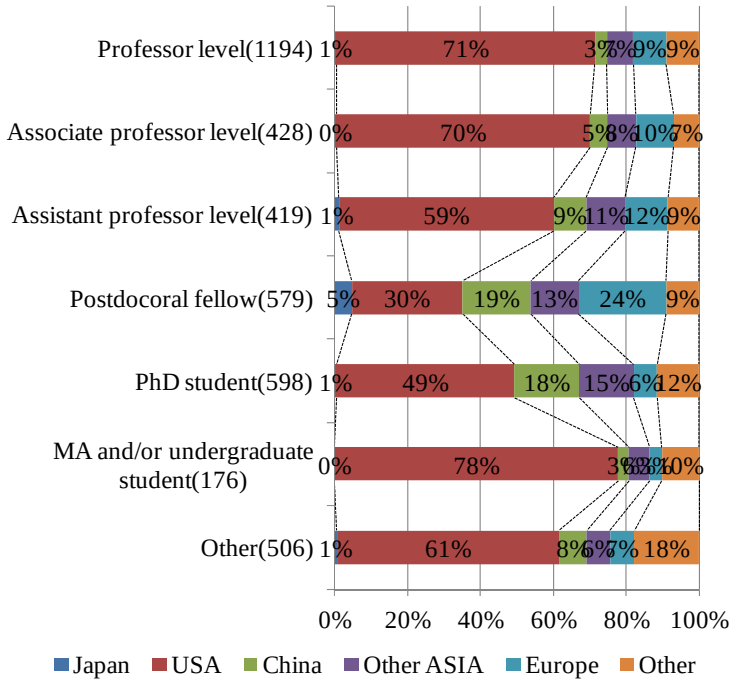


Figure 3. Country of birth of authors by their academic positions

(a) Japan



(b) US



Note: “Other” includes technician, the others and unknown.

Table 1 Percentage of young scholars among first authors of the focal papers in (a) Japan and (b) the US when the authors are listed in order of the contributions

(a) Japan

JPN	Normal papers				Highly cited papers			
	N	Young scholars	Students	Postdoctoral fellows	N	Young scholars	Students	Postdoctoral fellows
Natural sciences	849	35.0%	25.5%	9.5%	274	39.1%	19.0%	20.1%
Physical sciences	448	31.0%	22.3%	8.7%	158	32.9%	18.4%	14.6%
Life sciences	270	45.2%	34.1%	11.1%	66	51.5%	19.7%	31.8%
Medicine	131	27.5%	18.3%	9.2%	50	42.0%	20.0%	22.0%

(b) US

USA	Normal papers				Highly cited papers			
	N	Young scholars	Students	Postdoctoral fellows	N	Young scholars	Students	Postdoctoral fellows
Natural sciences	606	49.3%	30.7%	18.6%	261	50.6%	22.6%	28.0%
Physical sciences	298	53.4%	37.9%	15.4%	129	56.6%	38.0%	18.6%
Life sciences	177	60.5%	33.3%	27.1%	59	64.4%	13.6%	50.8%
Medicine	131	25.2%	10.7%	14.5%	73	28.8%	2.7%	26.0%

Table 2 Type of authorship (affiliated institutions) vs. Composition of the countries of birth of the authors

JPN	Country of birth			
	N	Only Japan	Only other country	Japan + other country
Domestic authorship	1,099	86.3%	0.1%	13.6%
International authorship	441	7.0%	1.8%	91.2%

USA	Country of birth			
	N	Only USA	Only other country	USA + other country
Domestic authorship	1,074	33.1%	16.9%	49.9%
International authorship	393	3.3%	28.0%	68.7%

Table 3 Percentage of the research teams that used state-of-the-art research facilities/equipment owned by the teams (results of Japanese survey)

	N	Internal facilities/equipment	External facilities/equipment
Only senior scholar	437	33.4%	14.6%
Student involved	652	40.3%	15.6%
Postdoctoral fellow involved	215	50.7%	21.4%
Both student and postdoctoral fellow involved	155	51.6%	25.2%

Table 4 Summary of variables used in the estimation models

Variable	Description
ADV_EQUIP_EXT	1: Team used the state-of-the-art external research facilities/equipment 0: Otherwise
ADV_EQUIP_OWNED	1: Team owned the state-of-the-art research facilities/equipment 0: Otherwise
ADV_EQUIP_OWNED_EFFECTIVE	2: Facilities/equipment was used and very effective 1: Facilities/equipment was used, but not effective 0: Facilities/equipment was not used
COUNTRY	County dummy (0: JPN, 1: USA)
DIV_EXPERTISE	The number of the field of expertise covered by a team
FIELD3	Field dummies (3 large fields)
FIELD21	Field dummies (21 fields in the Essential Science Indicators)
LG_AUTHORS	The number of authors (in logarithm)
LG_CITATION_LAG	Normalized average citation time-lags between the publication year of the focal paper and the publication year of the references therein (in logarithm)
LG_DURATION_PUBLICATION	How many years it took from the launch of research projects to the publication of the focal paper (in logarithm)
LG_EXP_MANMONTH	Research expenditure per man-month (in logarithm)
ORDERED_CITATION_COUNTS	Ordered citation counts of the focal paper
POSTDOC_FIRST_PROB	The probability of a postdoctoral fellow becoming the first author
POSTDOC_FOREIGN_PERCENTAGE	The percentage of foreign-born in postdoctoral fellows
POSTDOC_PARTICIPATION	1: One or more postdoctoral fellows are in the authors of the focal paper 0: Otherwise
POSTDOC_PERCENTAGE	Percentage of postdoctoral fellows in the authors of the focal paper
STUDENT_PARTICIPATION	1: One or more students are in the authors of the focal paper 0: Otherwise
STUDENT_PERCENTAGE	Percentage of students in authors of the focal paper
THREAT	2: Level of concerned is high 1: Level of concern is moderate 0: Not at all concerned about the priority loss.
TYPE_AUTHORSHIP	2: International paper 1: Domestic paper with foreign-born scholars 0: Domestic paper without foreign-born scholars

Table 5 Summary statistics of the variables used in the estimation models

Variable	Observations	Mean	Std. Dev.	Min	Max
ADV_EQUIP_EXT	1,342	0.417	0.493	0	1
ADV_EQUIP_OWNED	1,342	0.173	0.378	0	1
ADV_EQUIP_OWNED_EFFECTIVE	1,342	0.794	0.958	0	2
DIV_EXPERTISE	2,568	1.630	0.830	1	6
LG_AUTHORS	2,568	1.497	0.601	0.693	5.521
LG_CITATION_LAG	2,568	-0.079	0.422	-2.550	1.108
LG_DURATION_PUBLICATION	2,417	1.392	0.547	0.000	3.611
LG_EXP_MANMONTH	2,568	7.577	1.802	-0.924	16.300
ORDERED_CITATION_COUNTS	2,561	3.098	1.772	1	5
POSTDOC_FIRST_PROB	560	0.215	0.461	-0.667	0.857
POSTDOC_FOREIGN_PERCENTAG	887	0.631	0.455	0	1
POSTDOC_PARTICIPATION	2,568	0.346	0.476	0	1
POSTDOC_PERCENTAGE	2,568	0.111	0.180	0	1
STUDENT_PARTICIPATION	2,568	0.544	0.498	0	1
STUDENT_PERCENTAGE	2,568	0.196	0.214	0	1
THREAT	2,568	1.005	0.739	0	2

Table 6 Prediction of the involvement of postdoctoral fellows (Marginal effects)

		Model 1-1 JPN + USA Logit Marg eff.	Model 1-2 JPN + USA POSTDOC_PARTICIPATION Logit Marg eff.	Model 1-3 JPN + USA Logit Marg eff.	Model 1-4 JPN Logit Marg eff.	Model 1-5 JPN Logit Marg eff.	Model 1-6 JPN POSTDOC_PARTICIPATION Logit Marg eff.	Model 1-7 JPN Logit Marg eff.	Model 1-8 JPN Logit Marg eff.
LG_CITATION_LAG		-0.116*** [0.024]	-	-0.105*** [0.025]	-0.065** [0.030]	-	-	-0.046 [0.031]	-0.044 [0.031]
THREAT (Base: not concerned)	Moderately concerned	0.089*** [0.025]	-	0.080*** [0.025]	0.079** [0.038]	-	-	0.062 [0.038]	0.055 [0.038]
	Concerned	0.112*** [0.031]	-	0.107*** [0.031]	0.081** [0.041]			0.081** [0.041]	0.067 [0.041]
TYPE_AUTHORSHIP (Base: Domestic paper without foreign-born scholars)	Domestic paper with foreign-born scholars	-	0.283*** [0.029]	0.273*** [0.029]	-	0.295*** [0.049]	-	0.290*** [0.049]	0.284*** [0.049]
	International paper	-	0.248*** [0.028]	0.245*** [0.028]	-	0.226*** [0.033]	-	0.223*** [0.033]	0.224*** [0.033]
ADV_EQUIP_OWNED		-	-	-	-	-	0.078*** [0.026]	-	0.071*** [0.026]
ADV_EQUIP_EXT		-	-	-	-	-	0.026 [0.033]	-	0.014 [0.032]
DIV_EXPERTISE		0.021* [0.013]	0.015 [0.013]	0.017 [0.013]	0.003 [0.016]	-0.005 [0.017]	-0.002 [0.017]	-0.006 [0.016]	-0.01 [0.017]
LG_AUTHORS		0.135*** [0.020]	0.113*** [0.020]	0.103*** [0.020]	0.118*** [0.024]	0.084*** [0.024]	0.112*** [0.024]	0.079*** [0.024]	0.069*** [0.024]
LG_EXP_MANMONTH		0.032*** [0.007]	0.035*** [0.006]	0.029*** [0.006]	0.029*** [0.008]	0.029*** [0.007]	0.030*** [0.008]	0.025*** [0.007]	0.023*** [0.007]
COUNTRY		0.187*** [0.023]	0.059** [0.024]	0.091*** [0.026]	-	-	-	-	-
FIELD21		YES	YES	YES	YES	YES	YES	YES	YES
Chi-squared		309.097***	374.166***	397.956***	114.125***	181.505***	114.824***	189.621***	195.060***
Log-likelihood		-1454.021	-1415.507	-1396.445	-692.391	-658.696	-692.086	-654.858	-650.503
Psuedo-R2		0.122	0.145	0.157	0.096	0.140	0.096	0.145	0.151
Observations		2568	2568	2568	1342	1342	1342	1342	1342

Robust standard errors in parentheses. Significant at *, **, ***: 90%, 95%, 99%.

Table 7 Prediction of the involvement of students (Marginal effects)

		Model 2-1	Model 2-2	Model 2-3	Model 2-4	Model 2-5	Model 2-6	Model 2-7	Model 2-8
		JPN + USA	JPN + USA	JPN + USA	JPN	JPN	JPN	JPN	JPN
			STUDENT_PARTICIPATION				STUDENT_PARTICIPATION		
		Logit	Logit	Logit	Logit	Logit	Logit	Logit	Logit
		Marg eff.	Marg eff.	Marg eff.	Marg eff.	Marg eff.	Marg eff.	Marg eff.	Marg eff.
LG_CITATION_LAG		0.058**	-	0.053**	0.068*	-	-	0.046	0.047
		[0.025]		[0.026]	[0.035]			[0.037]	[0.037]
THREAT (Base: not concerned)	Moderately concerned	0.010	-	0.013	-0.034	-	-	-0.024	-0.024
		[0.026]		[0.026]	[0.041]			[0.043]	[0.043]
	Concerned	0.010	-	0.010	-0.064	-	-	-0.072	-0.073
		[0.030]		[0.031]	[0.043]			[0.045]	[0.045]
TYPE_AUTHORSHIP (Base: Domestic paper without foreign-born scholars)	Domestic paper with foreign-born scholars	-	-0.093***	-0.090***	-	-0.063	-	-0.061	-0.061
			[0.029]	[0.029]		[0.050]		[0.050]	[0.050]
	International paper	-	-0.218***	-0.217***	-	-0.288***	-	-0.288***	-0.288***
			[0.027]	[0.027]		[0.035]		[0.035]	[0.035]
ADV_EQUIP_OWNED		-	-	-	-	-	0.010	-	0.014
							[0.031]		[0.032]
ADV_EQUIP_EXT		-	-	-	-	-	-0.025	-	-0.016
							[0.041]		[0.043]
DIV_EXPERTISE		-0.001	0.000	0.000	0.005	0.014	0.005	0.014	0.014
		[0.014]	[0.014]	[0.014]	[0.020]	[0.021]	[0.020]	[0.021]	[0.022]
LG_AUTHORS		0.094***	0.135***	0.139***	0.116***	0.173***	0.110***	0.179***	0.179***
		[0.021]	[0.022]	[0.023]	[0.030]	[0.032]	[0.030]	[0.033]	[0.033]
LG_EXP_MANMONTH		-0.014**	-0.013**	-0.011*	0.000	0.000	-0.006	0.005	0.004
		[0.007]	[0.007]	[0.007]	[0.009]	[0.009]	[0.009]	[0.009]	[0.009]
COUNTRY		0.053**	0.092***	0.092***	-	-	-	-	-
		[0.025]	[0.026]	[0.027]					
FIELD21		YES	YES	YES	YES	YES	YES	YES	YES
Chi-squared		226.338***	270.026***	274.097***	121.877***	164.365***	117.720***	171.048***	171.334***
Log-likelihood		-1642.014	-1613.074	-1610.891	-844.325	-814.197	-847.388	-811.408	-811.268
Psuedo-R2		0.072	0.089	0.090	0.084	0.117	0.081	0.120	0.120
Observations		2568	2568	2568	1342	1342	1342	1342	1342

Robust standard errors in parentheses. Significant at *, **, ***: 90%, 95%, 99%.

Table 8 Prediction of the probability of postdoctoral fellows becoming the first authors

		Model 3-1 JPN + USA OLS Coeff	Model 3-2 JPN + USA POSTDOC_FIRST_PROB OLS Coeff	Model 3-3 JPN + USA OLS Coeff	Model 3-4 JPN OLS Coeff	Model 3-5 JPN OLS Coeff	Model 3-6 JPN POSTDOC_FIRST_PROB OLS Coeff	Model 3-7 JPN OLS Coeff	Model 3-8 JPN OLS Coeff
LG_CITATION_LAG		0.003 [0.055]	-	0.006 [0.054]	-0.020 [0.089]	-	-	-0.005 [0.087]	0.001 [0.091]
THREAT (Base: not concerned)	Moderately concerned	0.099** [0.044]	-	0.07 [0.047]	0.176*** [0.065]	-	-	0.125* [0.073]	0.078 [0.079]
	Concerned	0.139*** [0.051]	-	0.103* [0.054]	0.199*** [0.069]	-	-	0.151* [0.077]	0.104 [0.081]
POSTDOC_FOREIGN_PERCENTAGE		-	0.145*** [0.047]	0.099* [0.053]	-	0.213*** [0.064]	-	0.117 [0.079]	0.106 [0.080]
ADV_EQUIP_OWNED_EFFECTIVE		-	-	-	-	-	0.112*** [0.032]	-	0.064 [0.039]
COUNTRY		0.105** [0.042]	0.047 [0.050]	0.049 [0.051]	-	-	-	-	-
FIELD3		YES	YES	YES	YES	YES	YES	YES	YES
F		11.788***	18.174***	10.850***	4.211***	6.286***	6.638***	3.872***	3.909***
R-squared		0.155	0.153	0.163	0.121	0.105	0.111	0.133	0.149
Adj-R-squared		0.142	0.145	0.148	0.091	0.087	0.093	0.098	0.108
Observations		394	394	394	153	153	153	153	153

Robust standard errors in parentheses. Significant at *, **, ***: 90%, 95%, 99%.

Table 9 Prediction of the impact of the involvement of young scholars

		Model 4-1	Model 4-2
		JPN + USA	
		ORDERED_ CITATION_COUNTS	LG_DURATION_ PUBLICATION
		Ordered Logit	OLS
		Marg eff.	Coeff
POSTDOC_PERCENTAGE		0.059*	-0.198**
		[0.034]	[0.087]
STUDENT_PERCENTAGE		0.022	-0.016
		[0.024]	[0.069]
TYPE_AUTHORSHIP (Base: Domestic paper without foreign-born scholars)	Domestic paper with foreign-born scholars	0.039***	-0.034
		[0.015]	[0.037]
	International paper	0.041***	-0.051
		[0.015]	[0.036]
DIV_EXPERTISE		-0.005	0.002
		[0.006]	[0.018]
LG_AUTHORS		0.057***	0.117***
		[0.011]	[0.031]
LG_EXP_MANMONTH		0.021***	0.007
		[0.003]	[0.009]
COUNTRY		0.042***	-0.194***
		[0.013]	[0.035]
FIELD21		YES	YES
Constant		-	1.336***
			[0.084]
Chi-squared/F		228.197***	7.824***
Log-likelihood/R-squared		-2357.618	0.093
Psuedo-R2/Adj-R-squared		0.051	0.079
Observations		1824	1720

Robust standard errors in parentheses. Significant at *, **, ***: 90%, 95%, 99%. The estimated marginal effect shows that the likelihood of the focal paper becoming top 10% both for the model4-1.