NBER WORKING PAPER SERIES

WHY AND WHEREFORE OF INCREASED SCIENTIFIC COLLABORATION

Richard B. Freeman
Ina Ganguli
Raviv Murciano-Goroff

Working Paper 19819
http://www.nber.org/papers/w19819

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2014

We appreciate assistance with the survey from John Trumpbour and input from Paula Stephan and Andrew Wang. We received helpful comments from Adam Jaffe, Ben Jones, Manuel Trajtenberg and participants at the NBER "The Changing Frontier: Rethinking Science and Innovation Policy" Conferences. This research was supported in part by the National Science Foundation's National Nanotechnology Initiative, Award 0531146. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Why and Wherefore of Increased Scientific Collaboration
Richard B. Freeman, Ina Ganguli, and Raviv Murciano-Goroff
NBER Working Paper No. 19819
January 2014
JEL No. J01,J2,J24,J4,J44,J61,J68,O31,O32,O33

ABSTRACT

This paper examines international and domestic collaborations using data from an original survey of corresponding authors and Web of Science data of articles with a US coauthor in Nanoscience and Nanotechnology, Biotechnology and Applied Microbiology, and Particle and Field Physics. The data allows us to investigate the connections among coauthors and the views of corresponding authors about the collaboration. We have four main findings. First, we find that US collaborations have increased across US cities as well as across international borders, with the nature of collaborations across cities resembling that across countries. Second, face-to-face meetings are important in collaborations: most collaborators first met working in the same institution and communicate often through meetings coauthors from distant location. Third, the main reason for most collaborations are to combine the specialized knowledge and skills of coauthors, with however, substantial differences in the mode of collaborations between small lab-based science and big science, where international collaborations are more prevalent. Fourth, we find that citation rates are higher in international collaborations than in domestic collaborations in biotech but not in the other two fields. Moreover, in all three fields, papers with the same number of coauthors had lower citations if they were international collaborations. Overall, our findings suggest that all collaborations are best viewed from a framework of collaborations across space broadly, rather than in terms of international as opposed to domestic collaborative activity.

Richard B. Freeman
NBER
1050 Massachusetts Avenue
Cambridge, MA 02138
freeman@nber.org

Ina Ganguli
SITE
Stockholm School of Economics
Sveavägen 65
PO Box 6501
SE-113 83 Stockholm
Sweden
ina_ganguli@hksphd.harvard.edu

Raviv Murciano-Goroff
Stanford University Economics Department
579 Serra Mall
Stanford, CA 94305-6072
ravivmg@post.harvard.edu

Scientists increasingly collaborate on research with other scientists, producing an upward trend in the numbers of authors on a paper (Jones, Wuchty and Uzzi, 2008; Wuchty, Jones and Uzzi, 2007; Adams, Black, Clemmons, and Stephan, 2005). Papers with larger numbers of authors garner more citations and are more likely to be published in journals with high impact factors than papers with fewer authors (Lawani, 1986, Katz and Hicks, 1997; deB. Beaver, 2004, Wuchty et al., 2007; Freeman and Huang, 2013), which seems to justify increased collaborations in terms of scientific productivity. The trend in coauthorship extends across country lines, with a larger proportion of papers coauthored by scientists from different countries (Indicators, 2012; Adams, 2013). In the US and other advanced economies, the proportion of papers with international coauthors increased from the 1990s through the 2010s, while the proportion of papers with domestic coauthors stabilized. In emerging economies, where collaboration has not yet reached the proportions in the US and other advanced countries, the share of papers with domestic collaborations and the share with international collaborations have both increased.

The spread of scientific workers and research and development activity around the world (Freeman, 2010) has facilitated the increase in international collaborations. The growing number of science and engineering PhDs in developing countries, some of whom are international students and post-docs returning to their country of origins (Scellato, Franzoni, and Stephan, 2012) has expanded the supply of potential collaborators outside the North American and Western European research centers. A rising trend in government and industry R&D spending in developing countries and grant policies by the European Union and other countries favor international cooperation. At the same time, the lower cost of travel and communication has reduced the cost of collaborating with persons across geographic locals. The increased presence of China in scientific research, exemplified by China's move from a modest producer of scientific papers to number two in scientific publications after the US, has been associated with huge increases in collaborations between Chinese scientists and those in other countries.[2]

Finally, the location of scientific equipment and materials, such as the CERN Giant Hadron Collider, huge telescopes located in particular areas, or geological or climatological data

[2] Science and Engineering Indicators 2013, Appendix table 5-27 gives scientific papers for the top five countries in 2009: US 208,601; China 74,019; Japan 49,627; UK 45,649; and Germany 45,002.

available only in certain parts of the world, have also increased collaborations. The US was not a prime funder for CERN, but the largest group of scientists and engineers working at CERN are Americans. China eschewed joining the CERN initiative as an associate member state, but many scientists and engineers born in China work at CERN as members of research teams from other countries.

How successful are collaborations across country lines and across locations in the same country? How do collaborators meet and develop successful research projects? What are the main advantages and challenges in collaborative research?

To answer these questions, we combine data from a 2012 survey that we conducted of corresponding authors on collaborations with bibliometric data from Web of Science (WoS) (Thomson Reuters, 2012) in three growing fields – Nanoscience and Nanotechnology, Biotechnology and Applied Microbiology, and Particle and Field Physics. The survey data allows us to investigate the connections among coauthors in collaborations and the views of corresponding authors about collaborations. The WoS data allows us to examine patterns of collaborations over time and to compare patterns found in our fields to those found in scientific publications broadly. To determine whether borders or space are the primary factor that affects the nature and impact of collaborations, we contrast collaborations across locations in the US, collaborations in the same city in the US, and collaborations with international researchers.

We find that US collaborations increased across US cities as well as internationally and that scientists involved in these collaborations and those who collaborate in the same locale report broad similarities in their experiences. Most collaborators first met while working in the same institution. Most say that face-to-face meetings are important in communicating with coauthors across distances. And most say that specialized knowledge and skills of coauthors drive their collaborations. We find that international collaborations have a statistically significantly higher citation rate than other collaborations in biotech, a modestly higher citation rate in particular physics but not a lower rate in nanotech. The higher rates occur because international collaborations have a greater numbers of authors than other collaborations rather than for any “international magic” operating on collaborations with the same number of researchers.

1. The Growing Trend of International Collaboration

We analyze data from corresponding authors and articles in which researchers collaborate in Nanoscience and Nanotechnology, Biotechnology and Applied Microbiology, and Particle and Field Physics. These three fields cover a wide span of scientific activity, with different research tools and methodologies.

Nanotechnology is a general interdisciplinary applied technology, where engineers often collaborate with material scientists. The electron microscope is a pivotal research tool. The US and other countries made sizable investments in nanotechnology beginning at the turn of the 21^{st} century, when President Clinton called for greater investment in nano-related science and technology. This lead to the 21st Century Nanotechnology Research and Development Act that President Bush signed in 2003. Other countries undertook similar initiatives in the same period.

Biotechnology is lab-based, in which the NIH dominates basic research funding, but where some researchers also have close links to big pharma firms. The most important change in research technology has been the US-sponsored Human Genome Project and new methods of genetic analysis and engineering that allow labs around the world to modify the biological underpinnings of living creatures to advance medicine and improve biological products and processes.

Particle physics has a theoretical part and an experimental part. Leading edge empirical research requires massive investments in accelerators and colilders, of which the Large Hadron Collider is the most striking example. Europe's decision to fund the Hadron Collider while the US rejecting building a large collider in Texas shifted the geographic locus of empirical research from the US to Europe and arguably spurred the greater growth of string theory in the US than in Europe. Particle physics is the most established and mature of the three sciences that we investigate, where highly sophisticated and mathematically rich theories guide empirical work, and where the massive equipment exemplifies big science.

To measure collaboration patterns in the three fields we use publication data from the WoS. We identified all papers in the WoS database from 1990-2010 with at least one US coauthor in journal subject categories *Nanoscience & Nanotechnology, Biotechnology & Applied Microbiology, Physics, Particles & Field*s. From these papers, we identify teams by the names of coauthors and locate the authors by author affiliations. This sample includes 125,808 papers. Using the location of the authors on each paper, we define four types of collaborations:

US-Only Collaborations, divided into *US collocated*, in which all US authors are in the same city; and *US Non-collocated* in which US coauthors are in at least two different cities.

International Collaborations, divided into *International/US collocated* in which US coauthors are in the same city with at least one foreign coauthor; and *International/US Non-collocated*, in which US coauthors are in two or more cities with at least one foreign coauthor.

Distinguishing between these forms of collaborations allows us to identify the impact of international collaborations per se on papers as opposed to the impact of collaborations in different locations, be they in the US or overseas; as well as to identify the effect of collaboration across locations within the US.

Figure 1 displays the proportion of papers in our four categories and the proportion with single authors in the three fields taken together each year over the period of study. The solid top line gives the share of papers in which a US-based author collaborates solely with authors collocated in the same city. It shows a marked decrease in collaborations between persons at the same locale from 1990 through 2000, which then stabilizes at about 40% of papers. The line in the figure labeled *Solo* shows the proportion of papers that are solo-authored. It drops from 20% to about 5% from 1990 to 2010. The line for International/US Collocated papers gives the share of papers for which at least one of the authors is in another country while all US authors are in the same city. It increases by 18 percentage points from 1990 to 2010. The line for International/US Non-collocated increases more modestly. Most of the increase in international collaborations was between US scientists based in one location and persons in another country. The line for US non-collocated collaborations shows an increase of about 5 percentage points from 1990 to 2010. Though the increase is less than for international collaborations, the increased geographic scope of collaborations involved more than crossing national boundaries,

To see whether the trend in collaborations varied noticeably among fields, Figure 2 displays the proportion of papers with by collaboration type for the three fields separately. The data for particle physics in Figure 2a show the highest level of international collaborations, due presumably to the importance of particle accelerators and other equipment that are available at only some sites. Figure 2b and Figure 2c show that in nano and in biotech, the most common form of collaborations are US collocated teams, while international/US collocated collaborations are second most common and US non-collocated collaborations are third in frequency. International collaborations were roughly as common as US non-collocated collaborations in nano and biotech until the late 2000s, when international collaborations rise sharply. In all of the fields, the proportion of papers by sole researchers and by researchers collaborating in the same city falls.

Though our sample covers just three fields, the increase in international collaborations resembles the patterns reported in National Science Board (2012) and in Adams (2013) for science more broadly and are representative of the increased geographic dispersion of coauthorship within the US as well.

2. Survey of Corresponding Authors

To go beyond bibliometric analysis of the pattern of collaborations, in August 2012 we conducted an on-line survey of the corresponding authors of papers published in 2004, 2007, and 2010 in the Web of Science Nano, Biotech, and Particle Physics subject categories. We identified all unique corresponding authors based on email addresses in these categories and selected one paper for each author, randomly choosing the paper from authors who had more than one paper in the database. Using the email address of the corresponding author, we sent a personalized email in English that invited them to complete the survey by clicking a link that connected them to the on-line survey instrument. If there was more than one corresponding author, we selected the one that appeared first. We sent 2 follow-up email reminders in August and September 2012. We used Qualtrics Survey Software and respondents accessed it from the Qualtrics server.

We customized each survey to ask the respondent about the specific collaboration and individual team members. The survey had 25 questions and was designed so that respondents would complete it in 10-15 minutes. The questions sought to discover how the team formed, how it communicated and interacted during the collaboration, the contribution of each coauthor, types of research funding, and the advantages and disadvantages of working with the team. The survey also included an open-ended question for respondents to make comments. Several respondents sent emails with additional thoughts and information about the collaboration.

Between August 13, 2012 and August 20, 2012, we emailed a total of 19,836 individuals, but since some email addresses had expired, changed, or some individuals were deceased, the number of individuals who received the email is lower. We received 3,925 responses, which implies a response rate of 20% – a proportion that is in line with other surveys of scientists

(Sauermann and Roach, 2013). For individuals who published their paper in the most recent year of our survey (2010), the response rate was 26%. The response rate of the emails that reached respondents was necessarily higher – approximately 29%.[3]

The survey asked the respondent which country each coauthor was *"primarily based in during the research and writing"* of the article. This gives us a more accurate measure of whether teams are international than the WoS data, which are based on author affiliations at the time of publication, which can produce errors if affiliations change between the time the research was undertaken and the time of publication, or because some people have affiliations from more than one country.

Table 1 compares the characteristics of collaborations in the papers that we analyze to those in the full sample of WoS papers and to those in the WoS sample in the 2004, 2007, and 2010 publications from which we drew the survey sample. Our final sample includes 3,452 respondents, due in part to the failure of some papers to meet our sample requirements of having at least one author primarily based in the US at the time of the research. Our analysis of the survey data uses the respondents' information to define US collocated, US non-collocated and international teams.[4] The column giving the difference between the distribution of our sample in column 3 and the distribution of the WoS sample in column 2 shows that our survey sample is overrepresented by US collocated teams, the more recent publication year (2010), and publications from biotechnology.

3. Collaborations over Distance

In what ways, if any, do papers with international collaborations differ from collaborations that occur solely in the US?

As Katz and Hicks (1997), Rigby (2009) and Adams (2013) have found, international collaborations tend to produce more highly cited papers than collaborations of persons in a single country. Taking all of our fields together gives a similar pattern where the US is the single country. For papers published in 1990-2000 – dates chosen to allow time for papers to gain

[3] Of those who received the email, 5,744 opened the survey, and 3,925 completed and submitted their answers. While we are unable to precisely count how many emails reached active mailboxes, based on the number of emails which "bounced" back from a sample of the messages sent, we estimate that approximately 32% of emails sent were undeliverable. Given this estimate, we approximate a response rate of 29% from the deliverable messages.

[4] Comparing the 34.01% in row "Int'l Collaboration Survey", which is based on the respondent's answers regarding the location of coauthors, and the 36.35% in "Int'l Collaboration" in Table 1 shows that using only reported author affiliations from publications overestimates the number of international teams by 2.35 percentage points.

substantial numbers of citations -- US papers with foreign authors obtain about 1 more citation in our three fields compared to US collaborations with fellow residents of the US (26.59 citations vs. 25.65 citations). Since US-authored papers average a greater number of citations than papers worldwide, it would have been reasonable to expect the opposite: fewer citations for US-based scientists who collaborate with persons outside the country than for US-based scientists who collaborate with other US scientists.

Does this mean that international collaborations per se produce better science as reflected in numbers of citations?

Our distinction between collaborations among US-based persons in the same city and in different locations in the US provides a natural comparison to answer this question. It allows us to compare citations for papers with international collaborations and citations for papers with collaborations across locales in the US. If the observed international effect is due to something special about international collaborations, the average citations for collaborations with scientists outside the US would exceed average citations for collaborations among non-collocated authors in the US as well as exceed the average citations for collocated authors in the US.

To see if this is the case, we calculated the average number of citations for papers published between 1990 (with 21 years of potential citations) and 2007 (with three years of potential citations) for the three types of collaborations. Figure 3 shows these averages for each year from 1990 to 2007. The number of citations varies over time, from approximately 30 for the older papers to 3-4 citations for the newer papers. In almost all years, papers with international collaborators and papers with non-collocated US collaborators have more citations than those published by collaborators in the same US city. But there is no clear pattern of differences in citations for papers coauthored by people in different cities than for papers coauthored by people in the US and in a foreign location. Among papers published between 1998 and 2007, US non-collocated collaborations obtain a higher number of citations than international papers, but among papers published between 1990 and 1997, there is no clear difference. For our purposes, the key finding is that cites per year between US non-collocated papers and international collaborations are reasonably similar and notably larger than cites to US collocated papers. This suggests that the greater cites of international collaborations reflects multiple locations more than having authors across national borders.

We pursue the comparison of citations across types of collaborations separately for each of our three fields separately. The style of research in the fields differs greatly, with a huge difference between particle physics, where empirical work often involves huge collaborations around particular pieces of equipment, and the more small science collaborations of nanotechnology and biotechnology research. This difference shows itself starkly in our data is in the number of researchers whose names appear on a paper, as given in Appendix Table A1, which summarizes the distribution of number of coauthors in the fields. Particle physics has a much higher average number of authors per paper than the other two fields, with the difference concentrated in the upper tail of the distribution of authors per paper. In physics papers, the upper 95th percentile of the number of authors per paper have 100 authors, while those in the 99th percentile have 523 authors – which far exceed the upper percentile numbers for authors in nano and biotech.

Reflecting the “big science” nature of some of the physics projects, the corresponding author on a physics paper with over 450 coauthors noted in our survey:

> *This research was carried out as part of a very large collaboration in which every member gets authorship and this is listed in alphabetical order on our papers. The collaboration consists of scientists and engineers with a wide range of expertise - many primarily involved in designing building and running instrumentation, and many analyzing data for various kinds of signal. This particular research was primarily carried out by myself, and the majority of the listed coauthors (including three of the selected authors in this survey) had no direct involvement in its preparation other than through collaboration membership.*

We next use a regression analysis to examine the relation between the modes of collaboration and the number of researchers listed as authors, which virtually all studies find to be related to numbers of citations, and the relation between the modes of collaboration and the number of references in a paper, which also tend to be positively related to citations. Table 2 records the regression coefficients and standard errors for regressions of numbers of coauthors and numbers of references on the type of collaboration and a year trend for each of our fields.

For particle physics, the estimated 43.8 coefficient on international collaborations in column 1 shows that the number of authors on papers is much higher for those than for the US collaboration reference group. The more detailed measures of collaborations in column 2 show that this difference is largely driven by international collaborations in which the US scientists are from many locations as well. This reflects the big science nature of empirical particle physics where huge numbers of collaborators work together with massive instruments and machine. Column 3 shows a smaller but still notable difference in the number of references of papers with international collaborations relative to US non-collocated collaborations and a substantial difference in the number of references in those forms of collaboration relative to US collocated collaborations in physics. A potential explanation of the difference in references is that persons in a given location are more likely to cite papers written in their location so that the greater the number of locations, the greater the number of references.

Columns 4 - 6 turn to the relation between types of collaboration and numbers of authors and references in nano, while columns 7 – 9 turn to the same relations in biotech. As both of these fields are dominated by collaborations with relatively small numbers of people, the estimated differences in numbers of authors and references by type of collaboration are much smaller than those in physics. The regression in column 4 shows that international papers have 1.3 more authors than papers written by authors solely in the US. The regression in column 5 shows that the number of authors is largest in papers written by international collaborations with non-collocated US teams. Column 6 shows that references differ markedly among collaborations in nano, and from those found in particle physics. The regressions for biotechnology show similar modest relations between numbers of coauthors and types of collaboration (columns 7 and 8) as in nano, but show relations between numbers of references and types of collaboration similar to those in physics, and the greatest number of references for international/non-collocated US collaborations (column 9).

All told, Table 2 shows that comparisons of papers with international and national collaborations, as in much of the bibliometric literature, can present a misleading picture about the science involved in the various types of collaborations. The collaborations involve drastically different numbers of people and differences in references that may also reflect different research technologies. To see whether there is some “international magic” in collaborations, we examine next the relation between the citations to a paper and the nature of the collaboration that produced the paper using a regression that includes s the number of coauthors and the number of references in the paper. Since citations have a life cycle – with the number of citations increasing sharply in the first 5 to 7 years after publication and then tapering off – we include dummy variables for the year the paper was published as well.

Tables 3a – 3c gives the results of this analysis for each of our fields. Column 1 of each of the tables estimates the difference in citations between international papers and US only collaborations. The estimates show a disparate pattern across the fields: an insignificant positive relation between international collaborations and citations for particle physics; a negative relation in nano; and a positive relation in biotech. Column 2 of each of the tables adds the number of coauthors to the regression. In each of the fields, the addition of numbers of authors reduces the coefficient on international collaborations, turning it from positive to negative for biotech. With the addition of numbers of references in column 3 the estimated relation of international collaborations to citations is significantly negative in all three fields. The disaggregation of types of collaborations in columns 4 in the three tables shows sufficiently weak and different patterns across the fields to suggest that there is nothing universal in the link between international collaborations and good science as reflected in ensuing citations to papers.

All told, the regression analysis in Tables 2 and 3 document the changing patterns of cooperation across locations in the three fields and their disparate relation with citations. By the nature of the bibliometric data, such analysis cannot provide much insight into the ways collaborating scientists work together to conduct the research that leads to published papers. To gain insight into what goes on in collaborations and the link to scientific outcomes, we turn to the survey of corresponding authors described in Section 1.

4. Survey Evidence

"*I think the best example of collaboration I have done is…where all the authors are from different countries and we met at the Bellagio Conference Center of the Rockefeller Foundation.*"

"*I think that it is absolutely indispensable to meet people in person to have effective collaborations.*"

"*Skype was not available…at the time we completed this work. We now use Skype or ITV connection to meet and discuss data with collaborators on a weekly basis.*"

"The international collaboration worked so well because of my frequent trips to Brazil during the project."[5]

For scientists to collaborate, they must meet and decide to work together, communicate during the collaboration, and ultimately combine their knowledge and skills to create sufficient new knowledge to generate a publishable paper.

Meeting and Communicating

We asked corresponding authors to answer the following question about their coauthors: "*How did you FIRST come in contact with each of these coauthors?*" For papers with up to six authors, we asked about each coauthor. For papers with more than six we asked about the first author and the last author if they were not the corresponding author and asked about randomly selected authors from the list of coauthors, so that we had information on a maximum of six collaborators.

Figure 4 displays the proportion of persons of each collaboration type who the corresponding author first met in one of five ways: advisor-student/post-doc; colleagues in the same department/institution; contacted without an introduction; conference, seminar or other meeting; visiting the department/institution. The figure shows that regardless of the form of collaboration, most first meetings occurred when the corresponding author and the other person worked in the same institution. For papers written in the same location, the predominant contact came through advisor-student or post-doc relationships. But over one third of the meetings came about as colleagues. For papers with authors from other US locations or foreign locations, the corresponding author met them through working in the same place, primarily as a colleague, but with nearly ten to sixteen percent meeting the person as a visitor. Conferences also accounted for a substantial proportion of the first meetings between corresponding authors on papers written with persons in other locations or in foreign locations.[6] Overall, Figure 4 shows broad similarity in the mode of meeting between non-collocated US authors and in the mode of

[5] The four quotes are based on comments from the open-ended section of our survey.
[6] The time series data in Appendix Figure A1 and A2 show that conferences have become a less important way to meet future coauthors, while students/postdocs have become more important, possibly due to their increased importance in the scientific production process.

meeting between US and foreign-located authors compared to the mode of meeting for coauthors in US collocated collaborations.

We asked corresponding authors the frequency with which they communicated with one or more of their coauthors from “every week” to “never”. Because collaborations that include persons in the same locale and persons in other locales as the corresponding author allow the corresponding author to meet face-to-face easily with some coauthors but only infrequently with coauthors in other locations, the question does not pin down differences associated with distance as well as we intended. To overcome this problem, we show in Figure 5 modes of communication between coauthors on two-authored papers, which differentiate properly communication between collocated, non-collocated US, and foreign coauthors.

The results show that the corresponding author relies extensively on face-to-face meetings when all authors are in the same location. But Figure 5 also shows that while face-to-face meetings are much lower for authors across distances, such meetings are still frequent. Among the 2-author papers, just over 50% of corresponding authors on international teams report meeting face-to-face at least a few times per year, while 64% of those on US non-collocated papers reported face-to-face meetings at least a few times a year.

By contrast, the figure shows no noticeable differences in using e-mail by distance. Corresponding authors in all forms of collaborations use e-mail frequently to communicate with their collaborators, approximately 40 weeks during the year. The differences in use of telephone vs. Internet (e.g. Skype) between US-based teams and international teams are readily explained by cost differences.

What Coauthors Bring to Collaboration

To understand what factors helped produce the collaborations, we asked the corresponding author to specify the unique contribution of each team member. Our question was “*Did any of the team members working on this article (including yourself) have access to one of the following resources that the other team members did NOT have which made it important for you to all work together on this topic?*” The possible choices were: access to data, material or components; data, material or components protected by intellectual property; a critical instrument, facility or infrastructure; funding; or unique knowledge, expertise or capabilities.

Figure 6 shows that the major factor cited for all types of collaborations was "unique knowledge, expertise, capabilities". That access to specialized human capital seems to drive collaborations, whether US or international, implies that a theory of collaboration should focus on the complementarity of skills and knowledge of collaborators just as the theory of trade focuses on comparative advantage, rather than on other factors. But there are still noticeable differences in the importance of other factors across forms of collaboration. Non-collocated and international teams were more likely to have a coauthor contributing data, material or components than US collocated teams– a pattern that has been increasing over time, according to Appendix Figure A3.

While most corresponding authors were able to answer the questions about the contribution and role of their coauthors, those on huge collaborations told a different story. As one respondent remarked, "Many of the questions are hard to translate to the field of experimental particle physics, where an international collaboration of hundreds of scientists work on the same project with funding from many countries. I can only guess, where the funding from each of the ~300 coauthors comes from, many of whom I have not even met. The published research is primarily the work of a single person (myself), but would not have been possible without having access to custom software and data provided by the collaboration."

Finally, taking advantage of the unique identification of authors in two-authored papers, we compare the specific contributions of foreign-located coauthors and domestic coauthors on those papers. US and foreign coauthors were equally likely to contribute "unique knowledge, expertise, or capabilities" and "data, material or components protected by intellectual property". Foreign coauthors are slightly more likely to contribute access to "data, material or components" or "a critical instrument, facility or infrastructure" while the US coauthor was slightly more likely to contribute funding.

Advantages and Challenges

We use a two-part strategy to assess the effects of the different forms of collaboration on the production and output of scientific activity in the papers in our survey. We asked the corresponding authors their views of the advantages and challenges on their collaboration. Then we used a regression analysis to relate the number of citations to the papers to attributes of the collaboration.

Table 4 summarizes the responses of corresponding authors on the advantages and challenges of the collaborations organized by type of collaboration. It records the average score on a five-point scale of agreement or disagreement with a set of statements that reflect the attributes of the collaboration. The corresponding authors agreed that their collaboration had substantial advantages in harnessing human capital to produce a scientific outcome. "Complementing our knowledge, expertise and capabilities" and "learning from each other" are the only two items with average scores greater than 4 in the table. The next highest score was that collaborations made the research experience more pleasant. Moreover, there is little variation in the responses between US non-collocated and international teams, with corresponding authors giving slightly higher scores to the knowledge advantages than the collocated teams.

Similarly, the groups ranked highly "Gaining access to data, material or components", though here the highest assessment came from the corresponding authors of US non-collocated teams. Another area of difference was in the higher score given by the corresponding authors of international teams and to lesser extent of US-non-collocated teams of the advantage of "Our research reached a wider audience" compared to US collocated only teams. Viewing "wider audience" in terms of the geographic distribution of citations, this suggests that the wider geographic distribution of authors, the wider is the distribution of citations, possibly even among papers with the same numbers of citations – a pattern that can be investigated further in the WoS data.

As for the challenges, US non-collocated and international teams tended to agree more that there was "Insufficient time for communication", "Problems coordinating with team members' schedules", and "Insufficient time to use a critical instrument, facility or infrastructure", but international teams did not report any greater problems in this regard than US non-collocated teams. Consistent with some of our earlier WoS results, geography would appear to be more than national boundaries in the way teams operated.

We also asked whether the corresponding authors viewed teams as having the optimal size. The responses, given in Appendix Table A2, show that most corresponding authors viewed their team as having the right size. Presumably the principal investigator(s) would have modified the team if they did not think that was the case. But there are some differences by collaboration type. US collocated teams were more likely to say that they needed additional collaborator (7.58% vs. 3.48% and 3.38% for US non-collocated and international); whereas international teams were more likely to say that fewer team members were needed (6.67% vs. 3.37% for US collocated). Reflecting the role of government policies, twenty four percent of the international teams received funding aimed at supporting cross-country collaboration, with 6.65% receiving US government funding, 4.64% receiving EU funding, and the remainder from other government sources.

As a second way to assess how the teams producing the papers in our survey performed, we combined our survey data with information on the surveyed paper from the WoS. We added to the Table 3 regressions of the number of citations on the characteristics of papers variables from the survey. Because publication of the paper preceded the survey, some of the corresponding author views of the collaboration will presumably have been affected by the success of the paper, which would give a distorted view of the link from collaboration to outcome. To deal with this problem, we limit analysis to the survey responses that seem least prone to be affected by the outcome – relatively objective questions about the way corresponding authors met coauthors, what coauthors contributed and funding support.

Table 5 gives the results of this analysis. Columns 1 and 2 replicate the regression estimates in Table 3 for the dichotomous international collaboration variable. While we don't present the regressions results for the larger sample of WoS papers in our 3 fields, the results in Table 5 show that the basic pattern found in the larger WoS sample is mirrored in the smaller survey paper sample, with however some differences. The positive coefficient on international collaborations in column 1 in Table 5 is larger than the coefficient in the comparable regression using the larger WoS sample papers in our 3 fields. The coefficients on the number of coauthors and number of references variables are positive and significant in column 2 of Table 5 but with the impact of coauthors larger than that of references, contrary to the result in the larger WoS sample.

The coefficients on the survey variables in columns 3, 4 and 5 show that papers in which at least one coauthor met at a conference had higher citations; that papers for which a coauthor contributed funding had lower citations, and that papers that got funding specifically for cross-country collaborations had lower citations.[7] A possible implication of these results is that

[7] We also estimated the model including dummies for whether the corresponding author didn't view the team size as optimal, and an average of the scores assessing the advantages and disadvantages to the collaboration, but found no effect of these measures on citations.

collaborations based on ideas or relations developed at conferences produce more cited and potentially better science than collaborations based on funding.

5. Conclusion: The Economics of Collaborations

This examination of scientific collaborations in particle physics, biotechnology, and nanotechnology has found that US collaborations increased across space, with the largest increase occurring across country lines; whereas the share of papers written by single scientists or by groups of scientists in single locations declined significantly. Our survey of corresponding authors shows strong similarities in the way collaborators first meet, communicate and work together, particularly in collaborations across cities and borders, save for very large physics projects. Consistent with studies for science broadly, citation rates in our three fields are higher the greater the number of collaborators. But our evidence is mixed on whether international collaborations are more productive in terms of citations than domestic collaborations. In biotech, international collaborations obtain more citations than domestic collaborations, but in nanotech they obtain fewer citations, and in particle physics there are no significant differences between international and domestic collaborations. In all three fields, papers with the same number of coauthors had lower citations if they were international collaborations, which suggests that the main advantage of international collaborations for papers with US authors is that they allow researchers to increase the number of collaborators more easily than if the supply of potential co-authors was limited to US-based scientists.

Why, then, has scientific research becoming increasingly collaborative? Viewing science as an aggregate process for producing new knowledge, the most plausible answer is that the knowledge base has become increasingly complex and specialized (Jones, 2010), and thus requires that increased numbers of researchers combine their expertise to make advances. Consistent with this, our survey of corresponding authors shows that access to specialized human capital is the main driver of collaborations. The growing number of references within papers suggests that each forward step in science builds on a large base of previous knowledge. And the positive link between numbers of references and citations suggests that the greater the knowledge that goes into a paper, the greater the scientific contribution of the paper. In short, the productivity advantage from collaborations appears to depend on the combination of ideas/knowledge from persons with different expertise, possibly along the lines laid out by

Weitzman (1998) in which the growth of useful knowledge comes largely from combining the growing supply of past ideas and knowledge in new ways.

Collaborations have costs as well as benefits. On one side are problems of coordinating the ideas of persons with different expertise or viewpoints or who are in different locations, and the expenses and difficulty of getting collaborators together or linking them with data and key pieces of machinery. Our survey finding that researchers meet most collaborators through personal connections made at their university or, to a lesser extent, at conferences, suggests that there is some role of chance in creating collaborations. The findings on use of the Internet and travel suggest further that those technologies have reduced the costs of collaborations, which ought to increase the amount of collaborative research work, particularly over distance.

To an individual researcher with career concerns, the biggest issue in collaboration is getting credit for a joint production. In a one-author paper, the one takes credit or blame. In a two-author paper, many fields adhere to the convention that the senior person’s name comes last and the junior person comes first, which potentially gives considerable credit to each. Freeman and Huang (2013) find that the impact factor of the placement of the paper and citations received depend more on the characteristics of the last named author, suggesting that the senior person has greater importance in gaining attention to the research and that junior persons can benefit from working with a more successful senior scientist. In papers with more than two authors, the decision on who is the first author and the placement of the non-first or non-last authors can create disputes in labs. On papers having huge numbers of names where tasks are highly specialized, the credit that a given author gains is presumably related to their specialty much like the credit that different members of a crew gain from a big movie production. Only people who know how the research (movie production) proceeded and what the particular person's function was would understand how to evaluate their receiving credit in the author (credit) list.

From the perspective of economic rationality, the decision of scientists to collaborate depends on both the productivity of the collaboration and the distribution of credit, much as decisions to engage in business partnerships depend on gains in profits and distribution of profits. To get some notion of the interplay of the factors underlying a collaboration, we consider the situation in which a scientist compares the value of collaborating on a paper with one or more other scientists compared to writing a paper by themselves. On the productivity/ citation side assume that a paper with n collaborators gains proportionately more citations than a solo-authored paper according to a linear productivity parameter $p > 1$ that links citations to numbers of authors. If a single authored paper gets C citations, a paper with two authors will get pC citations, a paper with three authors gains 2pC citations, and so on.

But whereas each author gets full credit for an individual paper they get only partial credit for joint work. Assume that a citation crediting function $\gamma(n)$ allocates credit for joint work with $\gamma(1) = 1 > \gamma(2) > \gamma(3)$ and so on. Someone seeking to maximize the number of citations credited to them would collaborate only if $p\,\gamma \geq 1$ – that is, if the gain in productivity from the collaboration exceeds the loss of credit associated with γ. If the crediting function was based on simple fractionalization of credit – with, say, each author in a two authored paper would be credited with 1/2 of the paper and thus 1/2 of the cites, p would have to exceed 2 for the two-authored paper to be worthwhile. Similarly, p would have to exceed n for any n-sized collaboration to be attractive. But estimates of the extent to which citations increase with number of authors falls far short of proportionality. Depending on field and specification, an increased number of authors raises citations by at most 1-2 citations per additional author (see the estimates in Table 3, which are in line with those in other studies). There would be little incentive to write a joint paper that gained 12 citations for which each author obtained credit for just 6 citations if a solo-authored paper could gain 10 citations.

How then might we balance the citation credit accounts to be consistent with the increased collaboration in science?

One possibility would be that scientists who collaborate with others are able to write so many more papers through division of labor than they would have by themselves so as to offset the lower credit set by the fractionalization crediting parameter. For most scientists, this seems unlikely. The average number of collaborators in science articles has roughly doubled in the past 4-5 decades while the number of papers written per researcher has not shown any such doubling. Most of the increase in papers over time has been associated with increased numbers of researchers rather than increased papers per researcher.

The other possibility, which we view as more plausible, is that the crediting function diverges greatly from fractionalization. The first author of a highly successful paper gains lots of credit. The last author also gains lots of credit. The authors in the middle of the author list presumably can less credit. But all gain credit for having been part of a successful team activity. We expect that the discount for working with others is, however, far less for the main authors and more nuanced for all authors than fractionalization. While we cannot test this interpretation with our data, it is testable with information on the future careers of persons who work on papers with different numbers of collaborators. We would expect first and last authors to benefit most from a successful collaboration in their future careers and for intermediate authors to benefit proportionate to their actual activity on the research, subject to the imperfections in markets and market information.

Finally, to the extent that the interplay between the productivity of working with other scientists and the distribution of credit affect collaboration decisions as hypothesized above, we would expect to find at most modest differences between the nature and effects of collaborations across national borders as within the US, as our survey and WoS data analysis seem to show.

References

Adams, J. (2013). “Collaborations: The Fourth Age of Research.” *Nature*, 497(7451), 557-560.

Adams, J. D., Black, G. C., Clemmons, J. R., & Stephan, P. E. (2005). “Scientific Teams And Institutional Collaborations: Evidence From US Universities, 1981–1999.” *Research Policy*, 34(3), 259-285.

Barrantes Bárbara S. Lancho, Vicente P. Guerrero Bote, Zaida Chinchilla Rodríguez, Félix de Moya Anegón (2012). “Citation Flows in the Zones of Influence of Scientific Collaborations” *Journal of the American Society for Information Science and Technology*, Volume 63, Issue 3, pages 481–489, March 2012.

deB. Beaver, Donald (2004). “Does Collaborative Research Have Greater Epistemic Authority?” *Scientometrics*, 60(3), 399-408.

Freeman, Richard B. 2010 “Globalization of Scientific And Engineering Talent: International Mobility of Students, Workers, and Ideas and The World Economy.” *Economics Of Innovation And New Technology*, Volume 19, issue 5, 201 pp. 393-406.

Freeman, R.B. and Huang, W. (2013). “Collaborating With People Like Me: Ethnic Co-authorship within the US.” Harvard mimeo.

Guerrero Bote, Vicente P. Carlos Olmeda-Gómez Félix de Moya-Anegón (2013). “Quantifying the Benefits of International Scientific Collaboration,” *Journal of the American Society for Information Science and Technology*, Volume 64, Issue 2, pages 392–404, February 2013.

Hsu, J.W., & Huang, D.W. (2011). “Correlation Between Impact and Collaboration.” *Scientometrics*, 86(2), 317–324.

Jones, B. F., Wuchty, S., & Uzzi, B. (2008). “Multi-University Research Teams: Shifting Impact, Geography, And Stratification In Science.” *Science*, 322(5905), 1259-1262.

Jones, Benjamin (2010). “As Science Evolves, How Can Science Policy?” NBER Working Paper No. 16002 Issued in May 2010.

Katz, J.S. And D.Hicks (1997). “How Much Is a Collaboration Worth? A Calibrated Bibliometric Model,” *Scientometrics*, 40:3, 541-554.

Lawani, S. M. (1986). “Some Bibliometric Correlates Of Quality In Scientific Research,” *Scientometrics*, 9:1-1 13-25.

Lee K, Brownstein JS, Mills RG, Kohane IS (2010). "Does Collocation Inform the Impact of Collaboration?" PLoS ONE 5(12): e14279. doi:10.1371/journal.pone.0014279

National Science Board (2012). Science and Engineering Indicators, 2012.

Rigby, John (2009). "Comparing the Scientific Quality Achieved By Funding Instruments For Single Grant Holders And For Collaborative Networks Within A Research System: Some Observations," *Scientometrics* 78:1, 145-164.

Sauermann, H., & Roach, M. (2013). "Increasing Web Survey Response Rates In Innovation Research: An Experimental Study Of Static And Dynamic Contact Design Features." *Research Policy,* 42(1), 273-286.

Scellato, G., Franzoni, C., & Stephan, P. (2012). "Mobile Scientists and International Networks," NBER Working Paper 18613.

Thomson Reuters (2012). Web of Science ®, prepared by THOMSON REUTERS ®, Inc. (Thomson®), Philadelphia, Pennsylvania, USA: © Copyright THOMSON REUTERS ® 2012. All rights reserved.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). "The Increasing Dominance Of Teams In Production Of Knowledge." *Science,* 316(5827), 1036-1039.

TABLES

Table 1. Distribution of papers by characteristics, Web of Science Papers and Survey Respondents

	(1) Papers, 1990-2010	(2) Papers in 2004, 2007, 2010	(3) Survey Sample, Papers in 2004, 2007, 2010	(3)-(2) Difference
Collaboration Type				
US Collaboration Only	66.29	63.65	62.25	-1.4
US Collocated	44.81	41.56	46.84	5.28
US Non-Collocated	21.47	22.09	15.41	-6.68
Int'l Collaboration	33.71	36.35	37.75	1.4
Int'l/US Collocated	24.04	26.04	26.94	0.9
Int'l/US Non-Coll.	9.68	10.31	10.81	0.5
Int'l Collaboration Survey			34.01	
Year				
2004	6.08	25.38	18.42	-6.96
2007	8.05	33.61	29.46	-4.15
2010	9.83	41.01	52.11	11.1
Field				
Nano	23.82	32.85	30.5	-2.35
Particle Physics	25.19	21.75	19.55	-2.2
Biotechnology	50.99	45.40	49.94	4.54
N	125,808	30,141	3,452	

Notes: (1) includes all papers in the Web of Science with more than 1 author, at least one US coauthor, and with journal subject categories of Nanoscience & Nanotechnology, Biotechnology & Applied Microbiology, and Physics, Particles & Fields, published from 1990-2010. (2) includes those papers in 2004, 2007, and 2010. (3) includes the respondents to our survey, which was a sample based on unique corresponding authors appearing in (2) that had more than 1 author.

Table 2: Estimated Relation Between Number of Coauthors and Number of References on Papers by Nature of Collaboration, By Field

	Particle Physics			Nano			Biotech		
	Coauthors	Coauthors	References	Coauthors	Coauthors	References	Coauthors	Coauthors	References
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
US Collaboration Only									
US Collocated									
US Non-Collocated		2.654**	3.453**		1.450**	-0.879**		1.688**	0.727**
		(0.150)	(0.377)		(0.033)	(0.232)		(0.029)	(0.179)
Int'l Collaboration	43.776**			1.331**			2.168**		
	(0.924)			(0.032)			(0.040)		
Int'l/US Collocated		12.017**	4.737**		1.458**	-0.963**		1.973**	0.275
		(0.641)	(0.313)		(0.033)	(0.272)		(0.032)	(0.189)
Int'l/US Non-Collocated		99.983**	4.590**		3.075**	0.168		5.015**	3.131**
		(2.091)	(0.400)		(0.073)	(0.400)		(0.126)	(0.359)
No. Coauthors			0.001			-0.060			0.435**
			(0.001)			(0.042)			(0.031)
Year Trend	-0.214*	-0.183*	0.796**	0.039**	0.038**	1.491**	0.078**	0.064**	0.535**
	(0.094)	(0.090)	(0.024)	(0.003)	(0.003)	(0.024)	(0.002)	(0.002)	(0.013)
Constant	433.018*	368.918*	-1.6e+03**	-73.670**	-71.578**	-3.0e+03**	-151.290**	-125.213**	-1.0e+03**
	(188.702)	(179.207)	(47.520)	(6.828)	(6.581)	(48.749)	(4.713)	(4.484)	(26.100)
R2	0.055	0.170	0.046	0.068	0.144	0.116	0.091	0.159	0.044
Nb. of Obs.	31,690	31,690	31,690	30,761	30,761	30,761	64,153	64,153	64,153

Notes: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, OLS estimation. Includes all papers in the Web of Science with more than 1 author, at least one US coauthor, and with journal subject categories of Nanoscience & Nanotechnology, Biotechnology & Applied Microbiology, and Physics, Particles & Fields, published from 1990-2010.

Table 3a: The Estimated Relation Between Number of Citations to a Paper and the Type of Collaboration That Produced the Paper, Particle Physics

	(1)	(2)	(3)	(4)
US Collaboration Only				
US Collocated				
US Non-Collocated				1.664*
				(0.691)
Int'l Collaboration	0.718	0.096	-1.212**	
	(0.469)	(0.452)	(0.464)	
Int'l/US Collocated				-1.418**
				(0.532)
Int'l/US Non-Collocated				1.402
				(0.856)
No. Coauthors		0.014**	0.014**	0.010**
		(0.003)	(0.002)	(0.002)
No. References			0.398**	0.396**
			(0.017)	(0.017)
Constant	24.030**	24.031**	15.404**	14.817**
	(1.953)	(1.945)	(1.894)	(1.901)
Year FE	Yes	Yes	Yes	Yes
R2	0.030	0.031	0.072	0.073
Nb. of Obs.	31,690	31,690	31,690	31,690

Notes: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, OLS estimation. Sample is all papers in the Web of Science with more than 1 author, at least one US coauthor, and with a journal subject category of Physics, Particles & Fields, published from 1990-2010.

Table 3b: The Estimated Relation Between Number of Citations to a Paper and the Type of Collaboration That Produced the Paper, Nano

	(1)	(2)	(3)	(4)
US Collaboration Only				
US Collocated				
US Non-Collocated				-3.971**
				(0.423)
Int'l Collaboration	-2.300**	-3.732**	-3.637**	
	(0.358)	(0.388)	(0.387)	
Int'l/US Collocated				-4.849**
				(0.470)
Int'l/US Non-Collocated				-6.305**
				(0.621)
No. Coauthors		1.074**	1.110**	1.294**
		(0.083)	(0.080)	(0.085)
No. References			0.295**	0.293**
			(0.068)	(0.068)
Constant	26.252**	21.712**	14.660**	14.747**
	(4.749)	(4.683)	(4.805)	(4.819)
Year FE	Yes	Yes	Yes	Yes
R2	0.039	0.045	0.068	0.070
Nb. of Obs.	30,761	30,761	30,761	30,761

Notes: + p < 0.10, * p < 0.05, ** p < 0.01, OLS estimation. Sample is all papers in the Web of Science with more than 1 author, at least one US coauthor, and with a journal subject category of Nanoscience & Nanotechnology, published from 1990-2010.

Table 3c: The Estimated Relation Between Number of Citations to a Paper and the Type of Collaboration That Produced the Paper, Biotech

	(1)	(2)	(3)	(4)
US Collaboration Only				
US Collocated				
US Non-Collocated				1.109*
				(0.531)
Int'l Collaboration	1.800**	-1.466*	-1.583*	
	(0.597)	(0.680)	(0.677)	
Int'l/US Collocated				-2.138**
				(0.647)
Int'l/US Non-Collocated				2.394
				(1.891)
No. Coauthors		1.506**	1.412**	1.333**
		(0.103)	(0.101)	(0.110)
No. References			0.193**	0.191**
			(0.015)	(0.015)
Constant	34.629**	29.522**	24.805**	24.917**
	(2.047)	(2.054)	(2.075)	(2.067)
Year FE	Yes	Yes	Yes	Yes
R2	0.025	0.032	0.036	0.036
Nb. of Obs.	64,153	64,153	64,153	64,153

Notes: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, OLS estimation. Sample is all papers in the Web of Science with more than 1 author, at least one US coauthor, and with a journal subject category of Biotechnology & Applied Microbiology, published from 1990-2010.

Table 4. Advantages and Challenges to Working with the Team

	US Collocated	US Non-Collocated	Int'l
Advantages			
Learning from each other	4.26	4.33	4.36
Complementing our knowledge, expertise and capabilities	4.39	4.58	4.57
Gaining access to data, materials or components	3.21	3.56	3.32
Gaining access to data, materials or components protected by IP	2.14	2.30	2.29
Our research reached a wider audience	3.24	3.37	3.48
The research experience was more pleasant	3.96	3.92	4.02
Challenges			
Insufficient time for communication	1.82	2.13	2.11
Less flexibility in how the research was carried out	1.73	1.99	1.93
Unable to unequivocally portray my contribution	1.55	1.59	1.65
Problems coordinating with team members' schedules	1.96	2.18	2.11
Insufficient time to use a critical instrument, facility or infrastructure	1.45	1.67	1.67
Observations	1,693	585	1,174

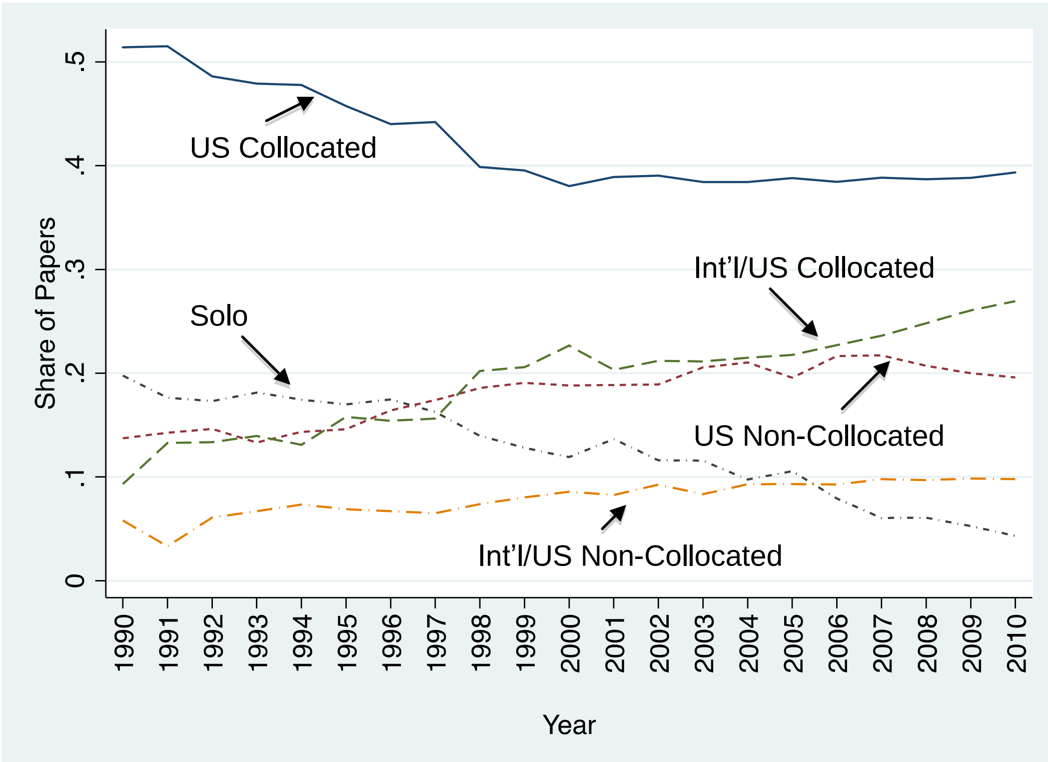
Notes: Respondents were asked to indicate their level of agreement with these statements regarding the main advantages/disadvantages of *"carrying out the research for this article with your team members"*, where 5 = Agree, 1 = Disagree.

Table 5: The Estimated Relation Between Number of Citations to a Paper and the Type and Characteristics of Collaboration, Survey Sample

	(1)	(2)	(3)	(4)	(5)
US Collaboration Only					
US Collocated					
US NonCollocated			-0.355	0.434	0.444
			(0.779)	(0.779)	(0.773)
Int'l Collaboration	0.878+	0.192	-0.579	0.370	0.495
	(0.529)	(0.538)	(0.649)	(0.600)	(0.586)
No. Coauthors		0.161*	0.157*	0.160*	0.161*
		(0.066)	(0.066)	(0.066)	(0.066)
No. References		0.099**	0.098**	0.099**	0.098**
		(0.015)	(0.015)	(0.015)	(0.015)
How They Met					
Advisor-Stu./Postdoc			-0.734		
			(0.656)		
Colleagues			0.592		
			(0.547)		
Visiting			0.703		
			(0.877)		
Conference			2.939**		
			(0.993)		
No introduction			0.575		
			(0.890)		
Coauthor Contributions					
Knowledge, etc.				0.498	
				(0.682)	
Funding				-1.327*	
				(0.553)	
Data, etc.				-0.305	
				(0.520)	
IP Data, etc.				0.124	
				(0.630)	
Instrument, etc.				0.166	
				(0.567)	
Cross-country funding					-1.207*
					(0.610)
Constant	17.433**	14.655**	14.654**	14.925**	14.676**
	(2.497)	(2.513)	(2.593)	(2.607)	(2.548)
R2	0.076	0.114	0.119	0.116	0.115
Nb. of Obs.	3,452	3,452	3,452	3,452	3,452

Notes: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, OLS estimation. All regressions include year, field, and year x field fixed effects. Sample is the survey sample described in Section 2. "How They Met" and "Coauthor Contribution" variables are dummies indicating whether any coauthor on the team met that way/contributed the resource.

Figure 1. Share of Papers by Collaboration Type



Notes: Includes all papers in the Web of Science with at least 1 US author, and with journal subject categories of Nanoscience & Nanotechnology, Biotechnology & Applied Microbiology, and Physics, Particles & Fields, published from 1990-2010.

Figure 2a: Share of Papers by Collaboration Type, Particle Physics

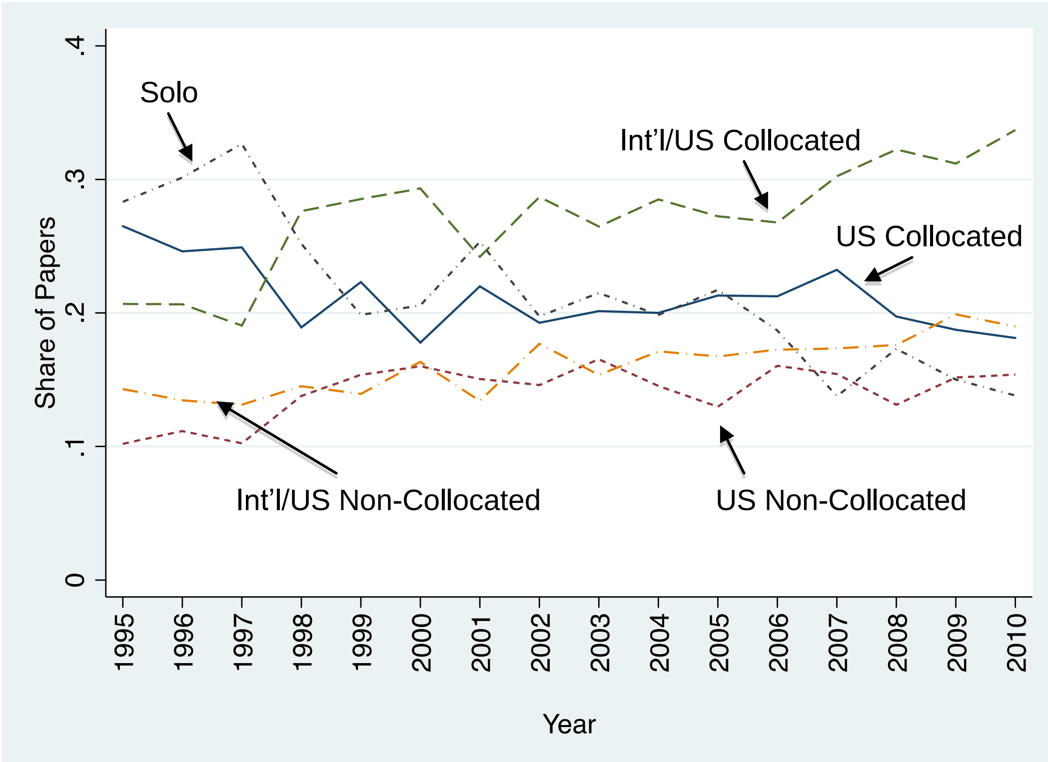


Figure 2b: Share of Papers by Collaboration Type, Nano

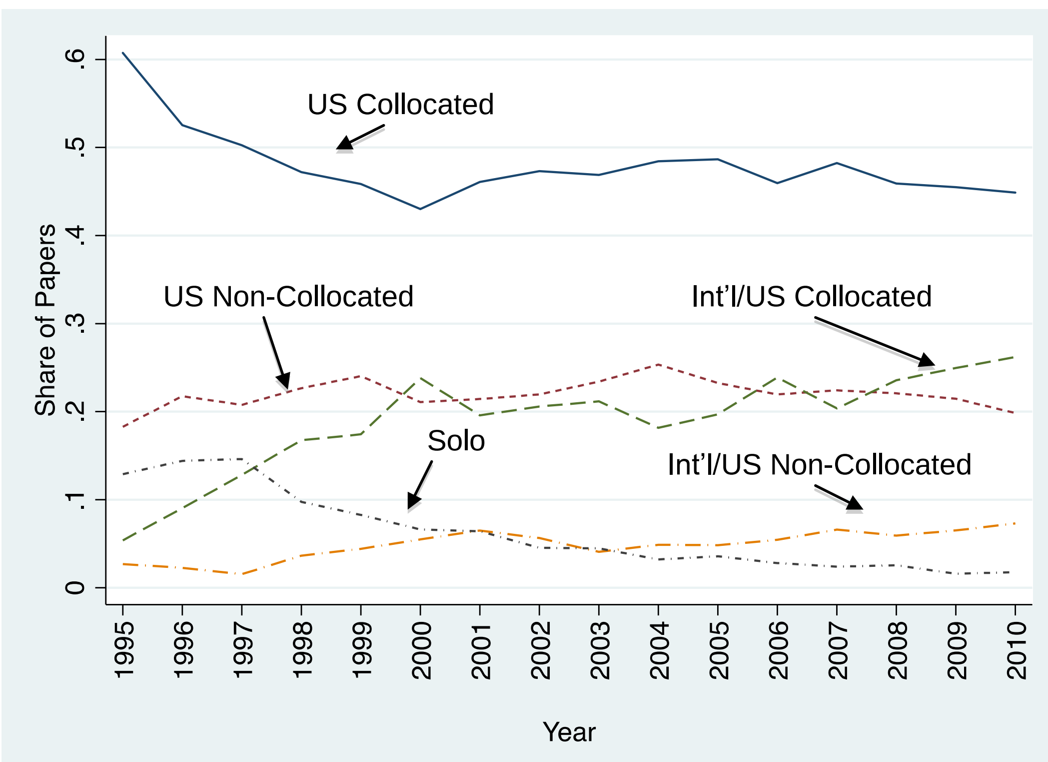


Figure 2c: Share of Papers by Collaboration Type, Biotech

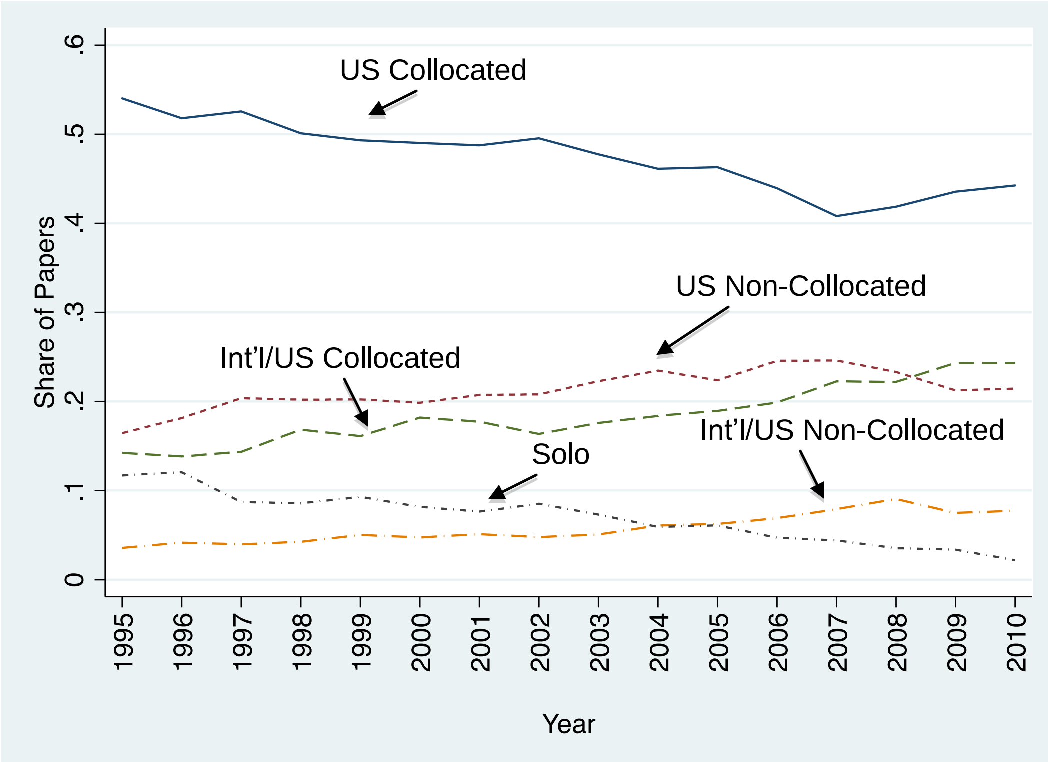
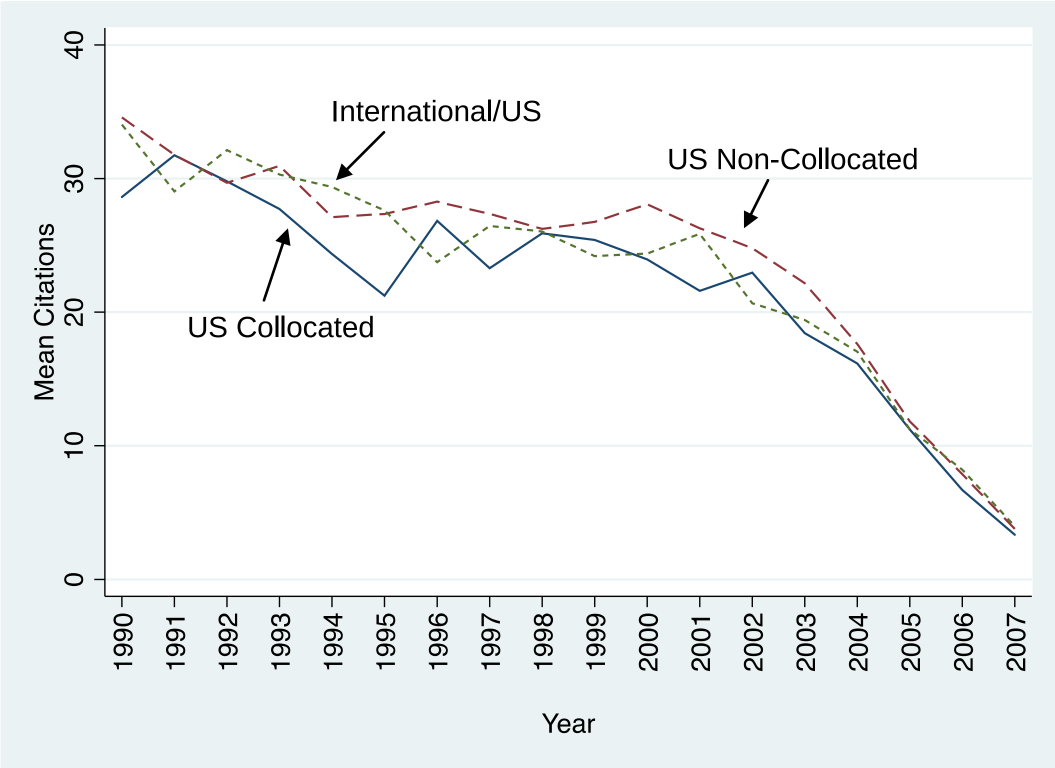
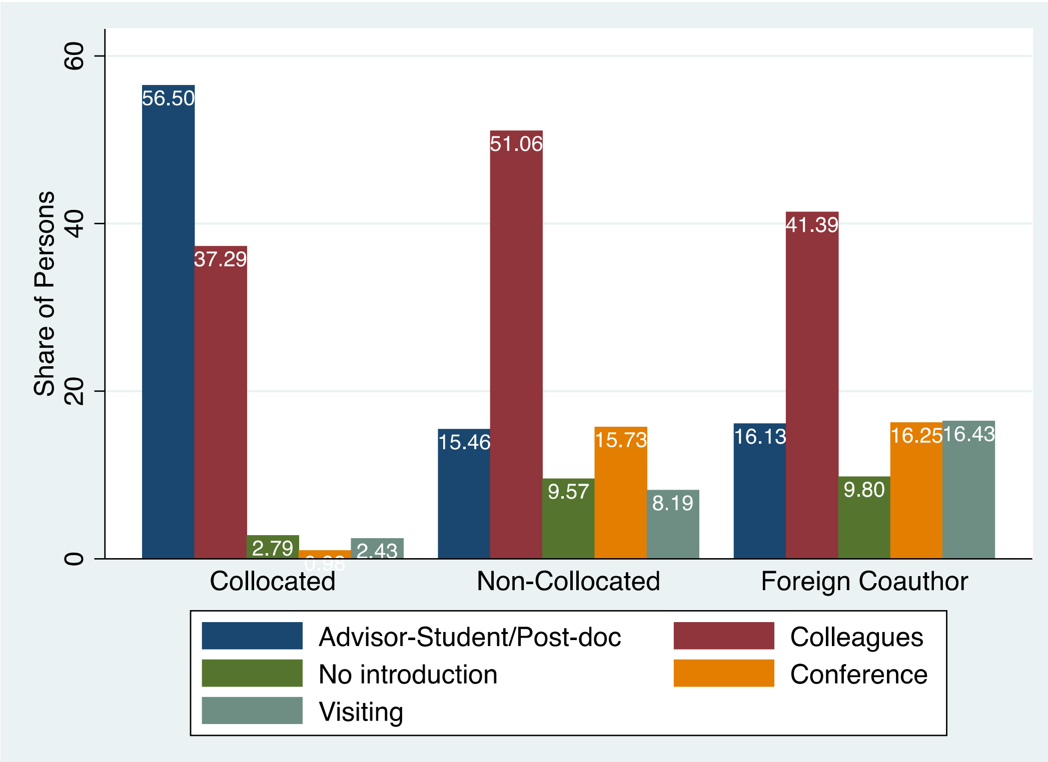


Figure 3. Citations By the Nature of Collaboration, All Fields by Year of Publication



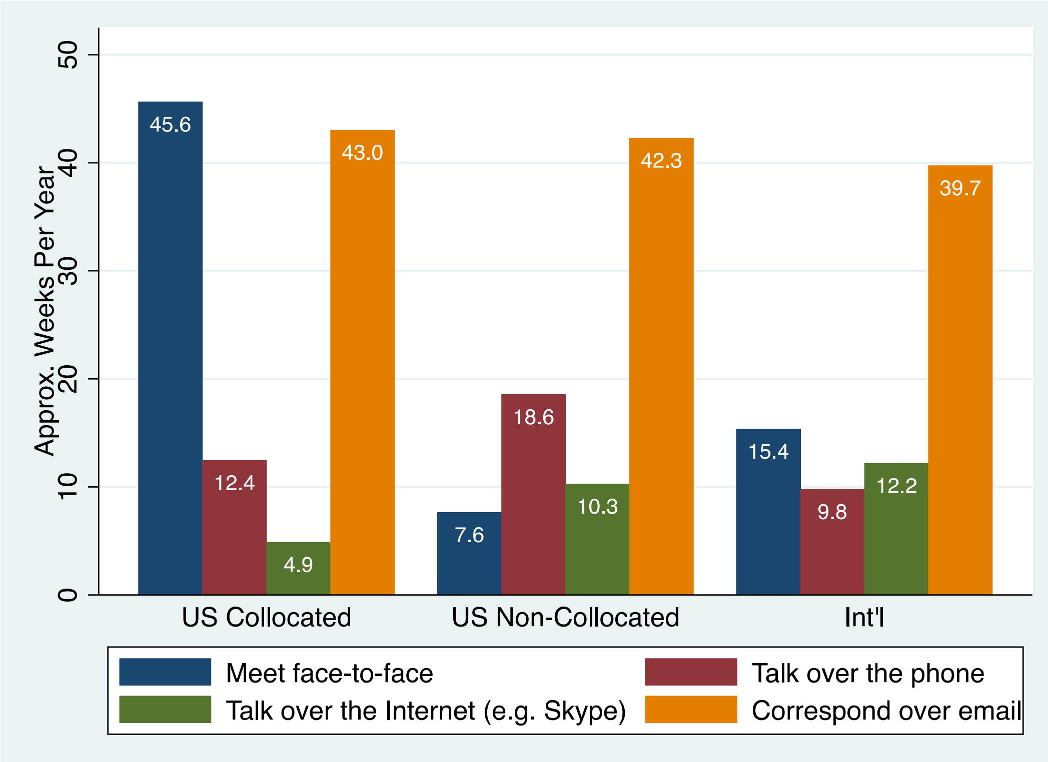
Notes: Figure shows forward citations of all papers in the Web of Science with at least 1 US author, and with journal subject categories of Nanoscience & Nanotechnology, Biotechnology & Applied Microbiology, and Physics, Particles & Fields, published from 1990-2010. Year indicates the year of publication of the cited paper.

Figure 4: Share of Persons Who Were First Met in a Given Way by the Nature of Collaboration



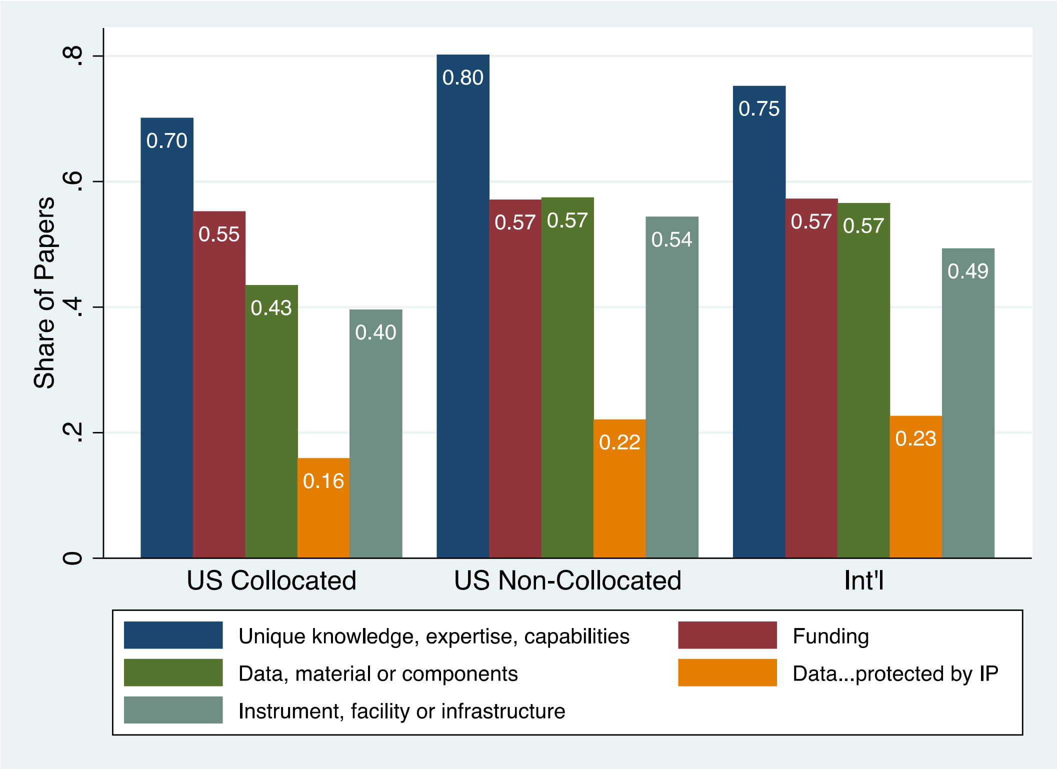
Notes: Share of all coauthors on papers for a given collaboration type. Question was phrased as "*How did you FIRST come in contact with each of these coauthors?*"

Figure 5. Overcoming Distance: Frequency of Communication Modes for 2-Author papers by the Nature of Collaboration (Approx. Weeks per Year)



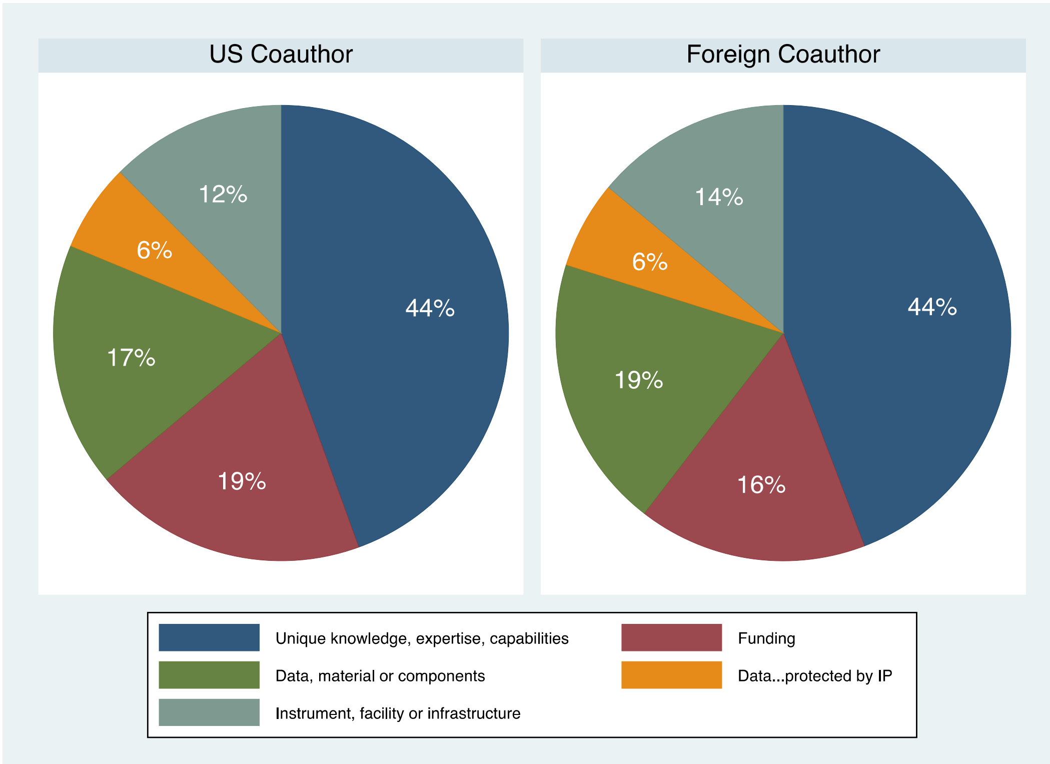
Notes: Question was phrased as "*When carrying out the research and writing for this article, how frequently did you use the following forms of communication with one or more of your coauthors?*" The possible choices were transformed into approximate number of weeks per year that each communication type was used: 6 = Every week (52), 5 = Almost every week (45), 4 = Once or twice a month (15), 3 = A few times per year (5), 2 = Less often than that (2), 1 = Never (0).

Figure 6: Contribution of Coauthors by the Nature of Collaboration



Notes: Share of papers for which the corresponding author reported at least one coauthor contributing the given resource. Question was phrased as "*Did any of the team members working on this article (including yourself) have access to one of the following resources that the other team members did NOT have, which made it important for you to all work together on this topic?*"

Figure 7. Contribution of US and Foreign Coauthors for 2-Author Papers



Notes: Share of US and foreign coauthors on 2-author papers only, as reported by the corresponding author.

APPENDIX

Table A1. Team Size Summary Statistics

	Particle Physics	Nano	Biotech
Mean	21.86	4.56	4.74
Standard Deviation	82.49	2.49	3.61
Maximum	1062	32	202
Percentiles			
10th	1	2	2
50th	3	4	4
75th	4	6	6
95th	100	9	11
99th	523	13	16
N	40,474	31,934	68,731

Notes: Measures of number of authors on papers in the Web of Science published 1990-2010, with a US author (including solo author papers), and with journal subject categories of Nanoscience & Nanotechnology, Biotechnology & Applied Microbiology, Physics, Particles & Fields.

Table A2. Optimal Team Size by Nature of Collaboration

	US Collocated	US Non-Collocated	All Int'l	Int'l with Cross-Country Funding
Yes	89.06	91.11	89.95	92.50
No, Additional	7.58	3.48	3.38	2.86
No, Fewer	3.37	5.40	6.67	4.64
N	1,663	574	1,154	280

Notes: Question was phrased as *"Do you think that the size of your team was optimal?"* The cross-country funding question was phrased as *"In carrying out the research for this article, did any of the coauthors receive funding that was specifically aimed at supporting cross-country scientific collaboration?"*

Figure A1. Share of Coauthors Who Were First Met at a Conference

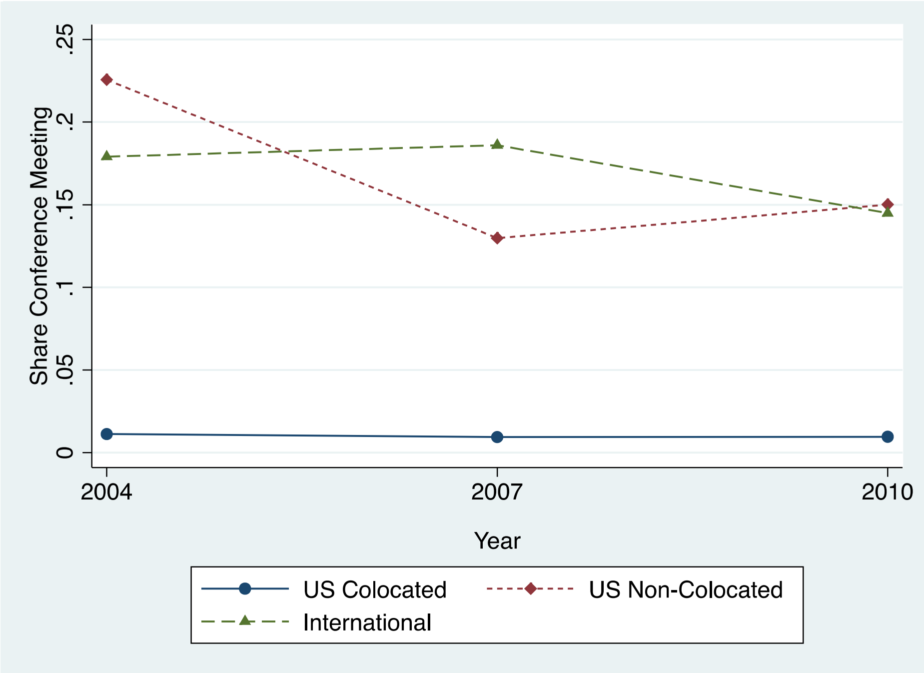


Figure A2. Share of Coauthors Who Were First Met as Advisor-Student/Postdoc

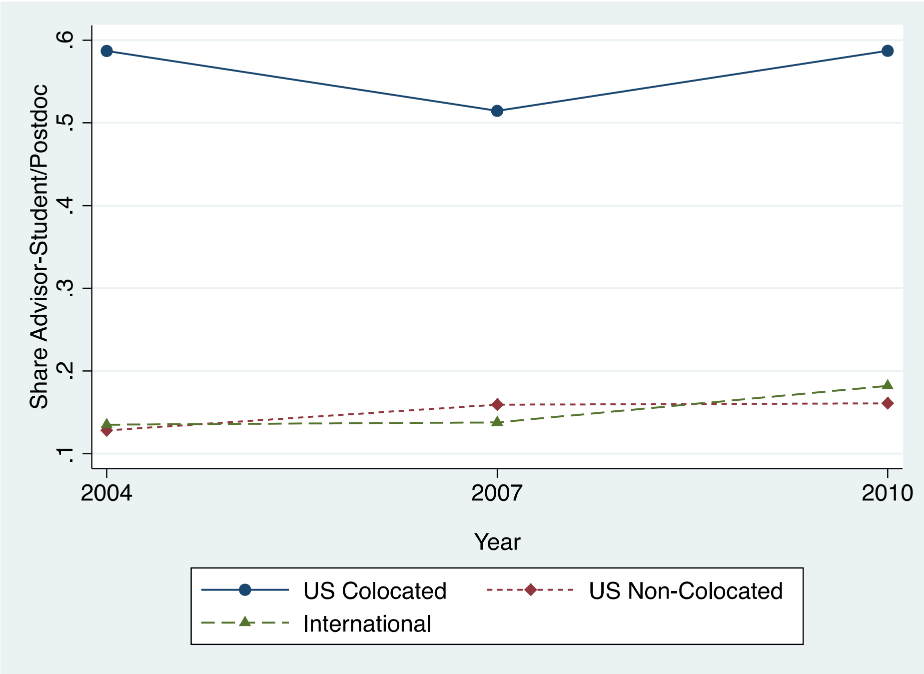


Figure A3. Share of Papers With a Coauthor Contributing Data, Material or Components