How Important is Location for Research in Science?

Shulamit Kahn*

Megan J. MacGarvie**

March 10, 2008

This paper asks whether scientists located outside the U.S. are at a disadvantage when it comes to research productivity, collaboration, and knowledge diffusion. The principal difficulties of comparing scientists inside the U.S. with those outside the U.S. arise from unobserved heterogeneity among scientists and the endogeneity of location choices. This paper makes use of a new and unique dataset of foreign-born U.S.-educated scientists that allows us to exploit exogenous variation in post-Ph.D. location induced by visa status. We thus are able to compare students who were required by law to leave the U.S. upon the completion of their studies with similar students who were allowed to remain in the U.S. We assess whether students who left the U.S. have more or fewer publications, patents, citations, and collaborators when compared with a control student with the same advisor. We also ask whether these students have more or fewer international collaborations.

---

* Boston University, ** Boston University and NBER. Email: skahn@bu.edu, mmacgarv@bu.edu. We thank the Institute for International Education and Jerry Murphy for data on Fulbrights. We would like to thank Gina Brandeis, Gerardo Gomez-Ruano, Dana Schulsinger, and Vanessa Wong for excellent research assistance. This project is funded by National Science Foundation Grant SBE-0738371.

The United States has the largest concentration of cutting-edge research scientists in the world, attracts more foreign graduate students than any other country, and is home to a disproportionate share of top scientists (Zucker and Darby 2007, Bound, Turner and Walsh 2006). If, as many papers suggest, knowledge diffusion and collaboration are enhanced by geographic proximity, then these facts alone will mean that the productivity of U.S.-based scientists will be elevated relative to those in other countries. Adding to this advantage is the ability of well-funded American universities and research institutes to devote considerable financial resources to increasingly expensive research laboratories and equipment.

There are several countervailing forces that might erase the advantages enjoyed by U.S. researchers. Other countries are educating increasing numbers of Ph.D. scientists and are attracting more star scientists. Other governments are making the development of stronger research capabilities a national priority, while the U.S. government has made some controversial policy choices that may have deterred some scientific explorations. At the same time, advances in communications technology and reductions in the cost of international travel have reduced geographic barriers to knowledge diffusion and to long-distance collaboration in science.

This paper asks whether scientists located outside the U.S. have in recent years been at a disadvantage when it comes to research productivity, collaboration, and knowledge diffusion. A first look at the data summarized in Figures 1 through 4 suggests the answer to this question is a resounding yes. The U.S.-located Ph.D. scientists in our sample produce more knowledge each year, as measured by journal publications, and this knowledge is diffused more broadly, as measured by forward citations to these articles. Scientists located abroad engage in fewer collaborative relationships with scientists at U.S. institutions, and their research is less likely to draw on the most recent scientific advances.

However, comparisons of scientists inside the U.S. with those outside are plagued by unobserved heterogeneity among scientists and the endogeneity of their location choices. Those scientists located in the U.S. and those outside are likely to vary widely in their research ability and proclivity both because of the job opportunities offered to them and because of the choices they make.

This paper makes use of a new and unique scientist-level dataset that allows us to exploit exogenous variation in post-Ph.D. location induced by visa status. We thus are able to compare foreign-born Ph.D. recipients who were required by law to leave the U.S. upon the completion of their studies with similar Ph.D. recipients who were allowed to remain in the U.S. We examine their research output in terms of the number of publications and the number of first-authored publications and their impact on science as measured by the number of forward citations. We also look more deeply at collaboration patterns, since collaboration is so crucial to scientific productivity. We find that the negative relationship between non-U.S. location and research output is eliminated after instrumenting for location. Our results suggest that the observed relationship between research output and location is principally due to selection bias, in which the most able researchers of foreign origin choose to locate in the U.S. However, instrumenting does not substantially change the results of our collaboration regressions, suggesting that location does have a significant causal impact on collaboration patterns.

Why Location May be Important

Both place and proximity matter for research productivity in science. In those geographic areas (typically metropolitan areas or states with one or more major universities) with greater stocks of knowledge (as measured by past articles, patent applications of scientists working there, presence of a star scientist etc.), we observe more new publications, patents and innovations by private companies and academics. To establish this correlation as geographic knowledge spillover (i.e. positive externalities) rather than geographic concentration of knowledge producers, authors have used a variety of strategies. Spillover is suggested by the increased likelihood to collaborate across sectors or institutions within the same geographic area (Zucker et al. 2007, Jaffe, Trajtenberg and Henderson 1993) and from the increased likelihood to cite articles and patents by others within the same geographical area (Jaffe, Trajtenberg and Henderson 1993, Agrawal, Cockburn and McHale 2006.) Other studies infer spillover from the tendency of new firms to locate near universities active in that field (Audretsch, Lehmann and Warning 2005, Zucker and Darby 2006, Zucker, Darby and Brewer 1998) or from the impact of exogenous changes in R&D funding (particularly in universities) on geographically-close companies (Jaffe 1989, Zucker, Darby, Furner, Liu and Ma 2007, Audretsch and Feldman 1996). However, Orlando (2004) and Thompson and Fox-Kean (2005) have contested the strength of some of this evidence.

Within academia, the quality of the university and department also has been shown to increase new knowledge creation. Thus, we observe that researchers at more highly ranked institutions publish more than those at lower-ranked institutions *ceteris paribus* and that the impact of location on an individual academic scientist's productivity is particularly important at the beginning of a Ph.D. scientist's career (Oyer 2006, Stephan and Levin 1992). The challenge in these studies is to establish that the research success of newcomers is due to the impact of the environment rather than simply evidence of clusters of productive researchers in excellent universities. To solve this, both Oyer (2006) and Stephan and Levin (1992) instrument for location of first job using demand and supply factors affecting the academic market in that field at the time of the initial placements or during the period of training.

There are many reasons that higher university quality might increase publication rates of newcomers. On the one hand, there are excellent potential collaborators and the direct exposure to the ideas and knowledge of cutting edge scientists. In addition, as Oyer (2006) notes, increased access to journal editors and reviewers, more physical resources, fewer teaching obligations, the high value put on successful research and the competitiveness of these environments all combine to increase the research productivity of newcomers to these institutions.

The U.S. has many of the best universities in the world. A Chinese ranking of the world's top Universities places the U.S. as having 15 and 17 of the top 20 universities in the world in natural sciences/math and engineering/computer science. (Shanghai Jiao Tong University 2008).[1] The U.S. also has the largest share of star scientists: Zucker and Darby (2007) identify the U.S. as having 50.2% of the stars in genetic-sequencing from 1973 to 1989.

As a consequence, if both the prestige of the university and the geographic proximity of many good scientists improve a scientist's research productivity, then foreign-born recipients of U.S. doctorates who return to home countries with lesser scientific communities will be less productive than those who remain in the U.S.. This diminished productivity

[1]The ranking is based on Nobel laureates and Fields medals prize winners, citations and publications. We thank Brown, Turner and Walsh (2006) for identifying this source.

might result because leaving the U.S. may involve diminished collaboration with Ph.D. advisors and other co-authors who remain in the U.S.. In addition, Ph.D. recipients who obtain university jobs in their home countries may face all the disadvantages of less prestigious universities already enumerated -- fewer highly published and well connected colleagues, fewer labs and resources, and lower publication norms. They also face higher costs of participating in U.S. conferences, seminars and meetings where they would have access to the network of scientists at American universities and where their research could find a wider audience. Finally, the availability of jobs where basic scientific research can be pursued may be lower in their country than in the U.S.

Moreover, initial career advantage tends to lead to later advantage in academia. This cumulative advantage, also called the Matthew effect, means that research scientists who have been productive in the past are more likely to be productive in the future. (See Stephan 1996 for a review of this literature.) Students who leave the U.S. post-Ph.D. for visa reasons are therefore likely to have their research career permanently affected.

Several trends may be working to moderate these factors. Kim, Morse and Zingales (2006) have found that co-authorship across long distances (albeit within the U.S.) has increased over the past decades. Consistent with this, in recent years collaborative ties have been shown to continue when a researcher changes geographic region (Agrawal, Kapur and McHale 2007) and “being in the same region or firm is found to have little additional effect on the probability of that knowledge flow (via patent citations) among inventors who already have close network ties” due to past collaborations (Singh 2005.) International collaboration has been subject to the same forces. Adams et al. (2005) find evidence of increased collaboration of S&E researchers in the U.S. with researchers in foreign universities during the nineties. Some of this may be due to an increasing propensity of U.S.-trained highly skilled immigrants to return to their countries, dubbed by Saxenian (2002) as a “brain circulation” replacing “brain drain.” “Brain circulation” also includes increasing professional and business links between highly skilled immigrants in the U.S. and their home countries, consistent with increasing international collaboration in academia. Indeed, Kerr (forthcoming) argues that international migration has enhanced knowledge diffusion, with non-U.S. inventors citing U.S. inventors of the same ethnicity 50% more often.

Above, we gave evidence of U.S.’s dominant position among the top world universities. However, the importance of being in a top university in terms of research productivity seems to be diminishing, at least in economics (Kim, Morse and Zingales 2006). Similarly, agglomeration effects in S&E research seem to be diminishing internationally. Within 14 OECD countries (including the U.S.), spillover effects of R&D spending by industry in one country has increasingly positive impacts on industry productivity in 13 other OECD countries suggesting that spillovers in science are becoming less localized and more internationalized (Keller 2002). Zucker and Darby (2006) find that there is no correlation between the level and the 1981-2004 growth rate of S&E stars across the 25 top S&E countries (including the U.S.), as increasingly the non-U.S. born stars living in the U.S. return to their home countries.

A final trend contributing to more equal research productivity of S&E scientists around the world is the growth of supply and demand for scientists and the increasing numbers of centers of scientific excellence outside the U.S.. On the supply side, the U.S. share of S&E Ph.D’s being awarded is dramatically decreasing, with Freeman (2006) documenting that in the past two decades, the major Asian Ph.D.-producing countries went from graduating less than half the number of Ph.D.s awarded by the U.S. to graduating more, and somewhat less dramatically, EU countries also moved from graduating less to graduating more S&E Ph.D.'s than the U.S.. While universities outside the U.S. have not made inroads into the top 20, between 2003 and 2007 they did gain slightly in their share of the top 100 universities although not in their share of the top 500 (Shanghai Jiao Tong University 2003 and 2007).

On the demand side, both Freeman (2006) and Kim, Morse and Zingales (2006) document the increasing numbers of highly skilled S&E jobs in the private labor market in other countries, as U.S. and multinational companies increase their non-U.S. employment of research scientists and off-shore additional high-level S&E jobs to foreign-owned companies. China has particularly accelerated its technological capabilities during the past decade.

Empirical approach: The Foreign Fulbright Program as an Instrument

In this paper, we study whether foreigners who received PhD's in the U.S. but then left the U.S., approximately 23% of graduating Ph.D.'s on temporary visas in 2005 as of the 2005 wave of the SED, were at a disadvantage with respect to research productivity, collaboration with the international scientific community, and knowledge acquisition and diffusion. Comparisons of U.S. and foreign scientists' research output will inevitably be plagued by endogeneity problems, as scientists' locations are likely to be influenced by unobserved characteristics correlated with these output variables. For example, the most productive foreign-born U.S.-educated scientists may be most likely to stay in the United States because they can choose from a wider range of options. To identify the effect of location on productivity and other aspects of scientific careers, we compare outcome variables for students on J-1 visas, specifically those who enter the U.S. through the Foreign Fulbright Fellowship program, who are required by law to leave the United States after finishing their doctorates for at least two years, relative to students who faced no such restrictions.[2]

For Fulbright status to be a useful instrument, we must establish that (1) far more Fulbright scholars leave the U.S. than other foreigners studying in the U.S. and (2) Fulbrights are similar to our control group with respect to potential research productivity.

Do Fulbright Fellows actually leave the U.S. as the conditions stipulated? It appears that the requirement to leave the country is quite stringent. It is possible to apply for a waiver of the foreign residency requirement if a student falls into one of several very restrictive categories.[3] They are sufficiently restrictive that almost all Foreign Fulbright recipients must fulfill the foreign residency requirement. A Fulbright recipient may delay their departure for a period, however, for educational purposes (i.e. a post-doc) and can apply for up to three years of "occupational or practical training" (OPT) immediately following the completion of doctoral studies.[4] Thus, in principle, a Foreign Fulbright recipient could remain in the U.S. for a substantial period of time following the completion of doctoral studies (up to 5-6 years) before having to leave the country. Moreover, after they spend two years abroad, they can

[2] Fulbright Fellows are the largest group of J-1 students. Note that after they fulfill their two year foreign residency requirement, they can apply for a work visa and return to the U.S..

[3] The first route is for the student to ask his country of origin to file a "no-objection" statement. While this approach may work for students whose J-1 status arose from scholarship funding from a foreign government, it is almost never considered grounds for waiving the foreign residence for Fulbrights whose funding comes from the U.S. government. (Conversation with BU ISSO January 2008) Waivers may also be obtained if an "Interested Government Agency (IGA)" files a request on behalf of the student, stating that the departure of the student will be detrimental to its interest and that of the public. Our conversations with experts suggest that these waivers are obtained only in rare and special circumstances. Medical doctors may also obtain a waiver if they agree to practice in a region of the U.S. with a shortage of health care professionals. A third reason for a waiver of the foreign-residency requirement is the threat of persecution, in which "an exchange visitor believes that he or she will be persecuted based on his/her race, religion, or political opinion if he/she were to return to his/her home country." Finally, applications for waivers may be filed on the basis of "Exceptional hardship to a United States citizen (or legal permanent resident) spouse or child of an exchange visitor." The State department warns "Please note that mere separation from family is not considered to be sufficient to establish exceptional hardship." http://travel.state.gov/visa/temp/info/info_1288.html (accessed February 17, 2008).

[4] OPT status allows students to work in their field of study for the purposes of obtaining on-the-job training.

apply for a work visa and return to the U.S.. In fact, a significant number of Ph.D. recipients did return to their home countries for two years, and then came back to the U.S. to take up a position at an American university or firm, or fulfilled the two year requirement in other ways.[5] However, for 81.5% percent of the Fulbright students in our sample, we were able to find evidence that they did spend time abroad after receiving their Ph.D., compared to 41.8 percent of our control group on non-Fulbrights. As Table 1 indicates, we observe our sample of 179 Fulbright scholars for a total of 1,626 person-years, and 74% of these years are spent outside the U.S. abroad. In contrast, the 179 controls – also foreign-born who had completed college in their home countries – spent less than 40% of their 1,714 observed person-years in the U.S..

While the Fulbright instrument is strongly correlated with the endogenous variable location, we still must face the second challenge of finding a control group that is similar to the Fulbright group in terms of potential research productivity. One potential problem is that students must apply to be in the Fulbright program. Students who feel strongly about the residency requirements of the visa may avoid applying for the fellowship. If this is the case, students who are most likely to become productive scientists may not apply, and our results will be biased towards finding that Fulbrights are less productive than randomly chosen non-Fulbright foreign students. A different problem is that the fellowship is merit-based. To the extent that Fulbright recipients are more accomplished than students in the non-Fulbright foreign student population, our findings may be biased in favor of finding that Fulbrights are

[5] Another way to fulfill the requirement is to accumulate a number of days in the home country that add up to two yrs. For example, a Ph.D. graduate could spend 6 summers there, or several semester-long visits, while under OPT status. Fulbrights who meet the requirement in this way are difficult to identify, unless they post an extremely detailed C.V. on their websites. In some cases, we were unable to find any evidence that a Fulbright fulfilled the foreign residency requirement. In these cases, we have no way of knowing whether the requirement was met through several visits adding up to two years or whether the requirement was waived.

more productive than other students. Under the first scenario (in which Fulbright status is negatively correlated with future productivity), estimates of the effect of location on research output obtained using Fulbright status as an instrument will be biased downwards. If Fulbrights are on average *more* productive than control students, the I.V. estimates will be biased upwards.

We deal with this issue in two ways. First, we create a control group of non-Fulbright students that is as similar as possible to the Fulbrights. Each Fulbright is matched to a randomly-selected student of foreign origin graduating in the same field at the same time from the same university (and, whenever possible, with the same advisor). Second, we control for each student's research productivity in the six years prior to Ph.D. completion. Students with more articles and more first-authored articles published while in graduate school are likely to have higher post-Ph.D. research output than students with fewer or no articles published during that time.

Data

The central piece of data in our project is information on the names, countries of origin, and fields of study of Foreign Fulbright Fellows who entered Ph.D. programs in science and engineering disciplines at U.S. universities in from the late 1980s through 1996. These data were obtained from volumes of *Foreign Fulbright Fellows: Directory of Students* published annually by the Institute for International Education from 1993 to 1996. The volumes published in 1994-96 list students starting programs in those years. The 1993 volume lists all the Foreign Fulbrights *enrolled* in U.S. graduate programs in that year and thus include some students who had started their Ph.D. programs during the 1980s. We started by collecting the names of all Foreign Fulbright Fellowship recipients in those years. From this group, we have so far identified 179 Fulbrights whose post-Ph.D. locations could be found via web searches.[6]

For each Fulbright in this group, we identified a "control" student – a student of foreign origin who did *not* have a Fulbright Fellowship. Our goal was to collect a sample of foreign students who did not have J-1 visas, and thus were not required to leave the U.S. after finishing their studies, but were otherwise identical to our Fulbright students. In an effort to make the Fulbright and control groups as similar as possible, we chose a control student for each Fulbright whose current location could also be found on the web and who graduated from the *same* program in the *same* year and, whenever such a student existed, with the same advisor,.[7] We obtained information on advisors, year of Ph.D., and field of study from the *Proquest Dissertations and Theses* database for both Fulbrights and controls.

To identify country of origin of possible controls, we looked at the Ph.D. dissertations themselves, viewed on Proquest. (see the Data Appendix for a detailed description of these data). When the student's undergraduate institution was listed in the dissertations, country of origin was based on that. This comprises a majority of our control sample. For the rest of our control sample, the country of origin was identified from the acknowledgements section of the dissertation, or from information on a student's country of origin or undergraduate degree drawn from a CV or bio found on the web.

Since students who receive substantial funding from their home country's government may also qualify for J-1 status and be subject to the foreign residency requirement, we

[6] We continue to collect data on Fulbright and control students. We expect to be able to increase the number of students in our database as the collection effort continues.
[7] In cases where there was no control student with the same advisor in the same year, we identified a student with the same advisor graduating within 3 years before or after the Fulbright. If no students met the latter criteria, we chose a student graduating in the same year in the same major field, but with a different advisor.

checked the “acknowledgements” section of potential control students’ dissertations and their CV’s for evidence of foreign governmental funding. When we found evidence of funding from a foreign government, we did not use the student as a control.

When several potential control students were identified for a single Fulbright, we chose students who came from the same or similar countries represented in the Fulbright sample. Table 2 lists the countries of origin of our Fulbright and Control samples. It is clear that the distribution of students across countries in the treatment and control groups, while similar, is not identical. There are several reasons for this. First, it is clear that the distribution of Fulbrights is affected by political factors. Thus, there are 9 Fulbright Scholars from Colombia but none from Venezuela, 5 from Thailand but none from Indonesia, China or India. We avoided choosing controls from China and India. However, when a suitable control could not be found from another country, we allowed students of Chinese and Indian origin in the sample. Finally, because many students from certain countries receive government funding, we were less likely to select controls from these countries. The differences in the countries of origin of the Fulbrights and control variables highlights the importance of including geographical control variables in our statistical analysis.

For each student in our sample, we collected a detailed history of all the student’s post-Ph.D. locations. This information was obtained in many cases from C.V.’s posted on the web. We also used information on authors’ affiliations listed in publications posted on the web. In other cases, we pieced together the student’s career history from multiple pieces of information found on the web (e.g., conference programs, course catalogues, faculty websites). If we were able to find evidence on a student’s location at different points in time but not for every year, we extrapolated the location information by at most two years.

The detailed histories were used to construct a dummy variable, *FORLOC*, which is equal to 1 if student *i* is located outside the U.S. in year *t*, and 0 otherwise. If we were unable to find information on a student's location in a given year, this variable is coded as missing.

We then collected data on the Fulbright and Control Ph.D.s' publication histories from *ISI's Web of Science*. Authors were identified using information on post-Ph.D. locations, authors' middle names, and fields of research. For each publication by an author, we obtained all information available on the publication record itself, including publication year, title, co-author names, author locations, complete backward citations, counts of forward citations, publication source, abstract, specific field (for example, Marine & Freshwater Biology), and keywords.

The data set includes 179 Fulbright Ph.D.s and 179 control Ph.Ds. We include data for each year that each Ph.D. is observed from their Ph.D. graduation year to 2007. Because there is a lag between when an article is submitted and when it is actually published, we generally lag the location variable *FORLOC* by two years, which leaves us with 2,654 observations. Throughout, the key right hand side variable *FORLOC* is a dummy for whether the researcher's location two years ago was in the U.S. or not.[8]

Measuring Research Output

In what follows, we analyze several aspects of the research output of the scientists in our sample. We focus on the following variables:

[8] We also experimented with different lag structures. Because FORLOC is highly serially correlated for each person, when more lags were included, their coefficients were typically insignificant. A two-year lag corresponded with what we expected to be the average lag between when the majority of the research was done and when it appeared in print (for the established journals in *Web of Science*) and also on average fit the equations the best. For Tables 6 and 7, we show specifications with the one year lag as well.

Publication counts: the number of articles on which the scientist is listed as a contributing author, by publication year. This is a measure of research output, but may be a noisy measure of research output for articles with multiple authors

First-authored publication counts: the number of articles on which the scientist is listed as the first contributing author, by publication year. This variable is a more direct indicator of the author's research output in fields in which there may be multiple authors and in which the first author is the major contributor to the research.

Citation-weighted total and first-authored publication counts: The total number of citations received by a publication, by publication year. Forward citations are an indicator of an article's impact, and we compute this for both the total articles published and the first-authored articles.

Median citation lag: The median difference between the articles' backward citations and its publication date. The longer the lag, the less likely that the article has been based on the most current science.[9]

Number of unique co-authors: This variable adds up the number of co-authors from each article published that year. The greater the number of coauthors, the more collaborators that the researcher has.

Percentage of publications with at least one foreign (non-U.S.) collaborative relationship, with at least one U.S. collaborative relationship, with at least one home country collaborative relationship, and with at least one third-country collaborative relationship: We expect foreign researchers to collaborate with foreign co-authors. However, a key question is

[9] Adams, Clemmons, and Stephan (2006) use the *modal* citation lag as a measure of how quickly scientific knowledge diffuses. While the modal lag is a more attractive measure, it is not as useful in our context because the typical author has only one or two articles per year. With low article counts, the number of unique years cited is low, and the modal lag is a noisier estimate of the vintage of the cited knowledge than the median.

whether they also continue to maintain ties with American researchers, either pre-existing ties with their advisors or their fellow students, or new ties. When any of these four collaboration variables are the dependent variable, the specification is limited to those person-years when one or more publications are observed.

Because the ISI database does not report an institution for each author, and instead reports each unique institution listed on the paper, we are only able to identify *collaborative relationships* between our scientists and at least one author at another institution. We thus compute the percentage of publications that have, for example, at least one co-author at a U.S. institution that is not the scientist's own institution. Similarly, we compute the percentage of publications that have at least one co-author at a *non-U.S.* institution that is not the scientist's own institution, and so on.

Percentage of publications co-authored with the scientist's thesis advisor: Students leaving the U.S. may be less likely to maintain collaborative relationships with thesis advisors due to the difficulties of long-distance collaboration. Alternatively, those outside the U.S. may be more dependent on thesis advisors as a link to the U.S. research network, and thus may co-author a larger share of papers with past advisors.

Table 5 displays the output and collaboration variables categorized both by present residence – U.S. or not – and by Fulbright status. The publication and citation data by present location confirm our expectations. Ph.D. scientists in the U.S. do publish more articles and are more highly cited, and differences are substantial. The publication and citation data by Fulbright status tell a very different story. Here, although the control scientists are much more likely to be living in the U.S., their publication and citation records are not very different from the Fulbrights' records and, when there are differences, the Fulbrights have the higher levels of publications and citations. This observation is suggestive of what we later find to be our major result, although we need much more analysis to establish this.

As expected, the average collaboration values indicate that scientists in the U.S. are networked to many more co-authors, and while those outside the U.S. have fewer co-authors overall, they have more collaborators from their home country (although not from third countries). In contrast to publication and citation variables, the control v. Fulbright comparison shares these general characteristics.

As predicted, the scientists located outside the U.S. do have longer median citation lags. The Fulbright v. control comparison is in the same direction, but differences are small.

Exogenous Control variables

The sample was constructed with the aim of choosing controls that are observationally identical to the Fulbright students. Nevertheless, in the regressions we include control variables to account for any differences that may exist between treatment and control groups.

Ln(real GDP per capita of the home country): The GDP per capita of the home country may affect these researchers in different ways. Richer countries may provide better research facilities and academic jobs for returning Ph.D.'s. Researchers in richer countries may be more able to travel abroad to access foreign networks. GDP may also be correlated with whether the student is located outside the U.S. post-Ph.D., as richer countries typically have more attractive job opportunities for scientists.

Region of origin dummies: This might pick up differences in regional resources and in cultural expectations. Europeans are more likely to move across European countries and access a wider variety of resources and possible employers. Note that "Europe" is somewhat of a misnomer, in that it includes Canada, Australia and New Zealand. The excluded region is Asia.

Field dummies: Fields differ widely in the number of co-authors per article, the number of articles published a year, and even in conventions regarding citing precedents. The excluded field is "Agricultural Sciences/Natural Resources."

Employment sector dummies: Jobs were categorized as being in government, industry, or academia (excluded category). To some extent, this might pick up one of the reasons that scientists in foreign locations are less productive, the scarcity of good academic jobs. In additional specifications (not reported), these dummies were excluded and made no qualitative differences to our conclusions.

Table 3 shows the similarity in fields between the Fulbrights and controls. Since the control was chosen from the same department, the distribution across fields of study should be exactly identical. There are small differences, however, since many dissertations list more than one field and often the fields specified are quite narrowly defined. While in our data we include only the first field listed on the Proquest dissertation record, different students of the same advisor and department thesis may list different narrowly defined fields and, even if the fields listed are identical, might choose to list them in different order.

Calendar year and years from Ph.D. dummy variables: These are included in all specifications (with the exception of the GMM I.V. specifications) but are not separately reported. [10]

[10] Due to the difficulty of getting the model to converge when many dummies are included, the year and the number of years post-Ph.D. are included in their original form rather than as dummies in the estimates obtained from the GMM I.V. The results of the un-instrumented Poisson regressions are very similar whether we include these variables as dummies or in their original form.

Table 4 lists Ph.D. year and we once again see a similar but not identical distribution between Fulbrights and controls. The differences are due to the fact that when there were no foreign students graduating in the same year as a Fulbright, we tried to find the closest available foreign student within three years of the Fulbright's Ph.D. receipt. Note also that since we identify Fulbrights in Ph.D. programs starting in 1993, there are no Fulbrights in our sample who graduated before then.

Number of articles and of first-authored articles published during graduate school: The number of pre-graduation publications measures individual-specific variation in past research productivity and hence inherent research potential. The inclusion of this variable in the regression is similar in spirit to the *pre-sample mean estimator* proposed by Blundell, Griffith and Windmeijer (2002) as an alternative to the fixed-effects Poisson model when regressors are predetermined and series are highly persistent.

Un-instrumented Estimation and Results

We first estimate the relationship between location and our research indicators in an un-instrumented model. Because we have panel data and our dependent variables are counts (number of publications, number of citations, etc.), we estimate Poisson models with robust standard errors clustered by scientist. We chose Poisson for its robustness, but Negative Binomial models yielded similar results.[11]

[11] Wooldridge (2002) explains that if the underlying distribution is truly Negative Binomial, the Negative Binomial estimator is more efficient than the Poisson, but if the distributional assumption is wrong, the Poisson is still consistent as long as the conditional mean is correctly specified. He writes, "On balance, because of its robustness, the Poisson QMLE has the edge over the NegBin1 for estimating the parameters of the conditional mean." (p. 657) In practice, we found that there was essentially no difference between results obtained using Negative Binomial model and those obtained from the Poisson model. The former are available upon request.

Tables 6-10 contain the results of Poisson regressions with robust standard errors clustered by student. Fixed effects for the year, number of years since Ph.D., the student's region of origin, the field of specialization, and the type of job are included. Table 6 focuses on publication counts. In columns 1 and 2, the dependent variable is the total number of articles published in year t on which the scientist was an author. There appears to be a negative but insignificant relationship between lagged foreign location and total publications for both last year's location and the location two years previous. Columns 3-6 contain the results of the regression when the dependent variable is the number of *first-authored* publications in year t. Here, the foreign location dummy lagged by one year is significant and is associated with a 30% reduction in the number of first-authored publications. The foreign location dummy lagged two years is negatively but only marginally significantly related to the number of first-authored publications. The final column controls for total publication count in year t so that the coefficient on the foreign location dummy can be thought of as measuring its impact on the percentage of publications that are first-authored. Results are qualitatively similar to column (4).

The number of articles published while in graduate school is, as expected, positively and significantly associated with the number of articles published post-Ph.D. Students from Latin America and the Middle East or Africa have fewer publications on average than students from Asia or Europe/Canada/Australia/New Zealand. As expected, scientists employed in government or industry have fewer publications than those employed in academic jobs.

Table 7 contains similar regressions in which the dependent variable is citation-weighted publications. There results here are stronger and more significant than for publications. There is a significantly negative relationship between foreign location (whether lagged one or two years) and the number of citations received by articles published by scientist *i* in year *t*. The effect, an approximately 40-50% reduction in the number of citation-weighted publications, is similar for citation-weighted total publications and for citations received by first-authored publications. This effect persists after controlling for the publication count in year t, implying that the number of citations per article, as well as the total number of citations, is negatively associated with foreign location. The control variables enter as in Table 7.

In Table 8, the foreign location dummy is broken down by region, with separate dummies for location in Europe/Canada/Australia/New Zealand and location in all other foreign countries. The region of origin dummies are omitted from this regression, since only a small number of students located outside the U.S. are found outside their broadly-defined regions of origin. The results show that most of the estimated relationship between foreign location and research output measured by publications or citations appears to be driven by students located outside Europe/Canada/Australia/New Zealand. Columns 5-8 present results from regressions similar to those in Tables 6 and 7, but in which students from Europe/Canada/Australia/New Zealand are dropped from the sample, and the results imply similar though stronger negative effects of foreign location on research output.

Columns 1-3 of Table 9 presents estimate poisson models of the relationship between the median backward citation lag and post-Ph.D. location using only the observations with at least one publication.[12] We find that scientists located outside the U.S. on average cite older literature than scientists inside the U.S. Being outside the U.S. increases the median lag by

[12] By doing this, we no longer have a one-to-one match between controls and Fulbrights so there may be more unobserved heterogeneity in this smaller sample.

approximately 20%, though the effect is diminished after controlling for the number of publications and forward citations. The number of forward citations is negatively associated with the citation lag, suggesting that articles that themselves receive more citations (and are perhaps of higher quality) tend to cite more recent articles. One interpretation of the longer backward citation lag in foreign locations is that recent scientific breakthroughs take longer to reach scientists outside the U.S. because distance impedes knowledge flows. Another interpretation, not inconsistent with the first, is that those outside the U.S. tend to specialize in less dynamic, slower-moving sub-fields of research.

Coefficients on other control variables confirm expectations. For instance, scientists in government also cite older literature, and those in Computer Science and Physics appear to have particularly fast-moving citation cycles.

An analysis of collaboration patterns by location is presented in Table 10. After controlling for covariates, the relationship between foreign location and the number of collaborative relationships (column 1) is negative and significant. Columns 2-6 allow us to track more specifics about collaboration patterns. Location outside the U.S. exerts a particularly strong negative influence on collaborations with scientists at U.S. institutions. Instead, scientists located abroad tend to collaborate more with non-U.S. scientists (at institutions excluding their own institution). There is no significant difference between scientists located in the U.S. and abroad when it comes to collaborations with past theses advisors.

Instrumented Estimation and Results

Since whether or not U.S.-educated Ph.D. recipients stay in the U.S., return home or go to a third country is obviously related to their research capabilities, our most important results are the instrumented results of Tables 11-13. In Table 11, we use instrumental variable regressions using Limited Information Maximum Likelihood (which is not a count data methodology) to estimate the unbiased influence of being in a foreign location. All dependent variables are logged. Our (single) instrument is the Fulbright dummy representing whether or not the Ph.D. recipients was required to leave the U.S.

In Appendix Table A1, we report results of the regression of the foreign location dummy on the Fulbright instrument. Table A1 presents several regressions of FORLOC on the instrument and control variables. A simple regression of FORLOC on the Fulbright dummy has an adjusted R-squared of 0.13, and the F-statistic associated with the Fulbright dummy is 50.86. Adding all the control variables reduces this F-statistic somewhat, however at 36.91 (column 3) it is still well above the “rule of thumb” critical value for weak instruments of 10 (Staiger and Stock(1997), Stock and Yogo (2005)). Adding the interaction of Fulbright status with the log of the home country's real GDP per capita does not improve the predictive power of the first stage, so we do not include this instrument in the regressions reported in Tables 11-13. The preferred specification, found in column 3 of the appendix table, is the first stage of the regressions in columns 1, 2, 4, and 5 of Table 11. The other columns in that table have an additional control variable, but are otherwise the same and have essentially the same F-statistic.

Instrumented, being in a foreign location has no effect on either publications or citations, neither total nor first-authored: the coefficients on *FORLOC* are all close to zero and statistically insignificant and vary in sign. We conclude from this that the negative effect of being in a foreign location on research output observed in earlier regressions is pure selection bias. Stated another way, if we were to randomly assign students to foreign locations upon the completion of their doctorates, we would not expect to see any difference in research output or its dispersion between foreign-located and domestically-located students.

In contrast, some of the other control variables that had been insignificant became significant. For instance, the impact of (total) articles published during grad school now has a larger impact on publications. Differences between regions of the world became smaller, however.

We also experimented with adding a second instrument interacting the Fulbright dummy with the log of real GDP per capita in the student's country of origin. This interaction term was meant to capture the fact that Fulbrights from richer countries may be more likely to remain in their home countries once the two-year foreign residency requirement is complete, whereas Fulbrights from countries with less well-developed science infrastructure may be more likely to return to the U.S. after completion of the requirement. This term was insignificant in the first stage so we do not include these specifications in the table. The results on foreign location do not differ dramatically depending on which sets of instrument we use.

Results in Table 12 are also IV estimation, but use a count-data instrumental variables model developed by Mullahy (1997), a GMM model for count data with endogenous variables and multiplicative error terms. We have also experimented with a similar GMM/IV model for count data with an additive error term (see Windmeijer 2006) and obtained similar results.[13]

The results using the Mullahy model in Table 12 are very similar to those in Table 11. Again, the estimated effect of foreign location on publications, citations, first-authored publications and citations to first-authored publications is essentially zero.

Table 13 re-estimates the collaboration equations instrumenting foreign location with the Fulbright dummy. The methodology used is IV with LIML since columns 2-6 are not count data. Comparing Tables 10 and 13, being in a foreign country no longer has a significant detrimental effect on the number of co-authors. In fact, the new coefficient, while insignificant, has changed signs. The coefficients indicating some collaboration with the U.S. are now larger in magnitude but somewhat less accurately measured. As before, location outside the U.S. reduces the likelihood of collaboration with scientists at U.S. institutions while increasing the likelihood of collaborating with scientists in non-U.S.institutions both in their home country and in third countries.

We have also rerun an IV version of the median citation model using the Mullahy approach, again instrumenting foreign location with the Fulbright dummy. Column 4 of Table 9 contains the results. Compared to the non-instrumented version in column 3, the foreign location dummy falls slightly and the standard error increases so that this coefficient is no longer significant at the 5% level. However, the fact that this coefficient changes so little after instrumenting suggests that the impact of location on the tendency to cite recent literature is not purely driven by selection bias.

[13] Angrist (2001) has shown that this model gives a consistent estimate of the local average treatment effect (LATE) in a model with a binary instrument and endogenous treatment variable. Which an endogenous *dummy* variable, the LATE is just the effect of the treatment on the treated.

Overall, our major finding is that instrumenting eliminates the negative effect of foreign location on research output. This result is somewhat surprising. While we originally expected our I.V. results to display a weaker correlation between location and research output than the uninstrumented results, we did not expect the difference in output between U.S. and foreign scientists to be eliminated completely.

We began this research expecting that those who would be most hurt by being outside the U.S. would be those who return to less developed countries, if for no other reason than that funding for research in the U.S. is considerably more generous than in most less developed countries. Our final table, Table 14, contains the results of the GMM I.V. regressions for publications and citations after excluding scientists from Europe, Canada, Australia and New Zealand. LIML estimates provided qualitatively similar results. This regression is not directly comparable with others in the paper due to the fact that in dropping students from Europe (etc.), we no longer have a one-to-one match between controls and Fulbrights, so that there may be more unobserved heterogeneity in this restricted sample.

In Table 14, we concentrate on the coefficients in columns 2-4 of this table, since the impact of FORLOC on total publications (column 1) was not significant even in the uninstrumented estimation using the entire sample of Table 6. The coefficient estimates on FORLOC in Table 14 (instrumented, less wealthy countries) falls between the corresponding ones in Table 8 and Table 12. They were large in magnitude and significantly negative in the uninstrumented estimation of only less wealthy countries (Table 8), are reduced in magnitude and measured more imprecisely in this instrumented estimation of only less wealthy countries, but they remain considerably more negative than the point estimates from the instrumented regressions of all countries of Table 12. This suggests that although there was clearly selection influencing the uninstrumented estimates, there *may* be some causal impact of location on research output for students from less wealthy countries who pursue scientific careers in their home countries that we have not been able to precisely estimate. As this research progresses, our focus will be on estimating this effect more precisely.

A final caveat is that our instrumented results can give us only the impact of the treatment on the treated, i.e. the effects of being in a foreign location estimates on Fulbrights or people like them. It is possible that, despite our attempts to control for unobserved quality, Fulbrights may be different than other foreign students in the U.S. in terms of inherent research ability, in which case they may not be informative for the population of scientists as a whole. However, we feel that any differences in unobserved quality between Fulbrights and controls in our dataset, after controlling for characteristics, are likely to be small. As a result, while our estimates may <u>understate</u> the true causal impact of location on research output, we expect that the understatement is not severe.

Conclusion

In this paper, we have examined whether newly-minted Ph.D.s of foreign origin who obtain their degrees in the U.S. maximize their post-Ph.D. contributions to science if they remain in the U.S.. A naïve comparison of post-Ph.D. publication records for a sample of such students would suggest that those who remain in the U.S. are at an advantage, based on higher rates of publication, citation, collaboration with U.S.-based scientists, and access to the most recent research. However, an analysis which uses exogenous variation in post-Ph.D. location to identify the causal effect of location on research output suggests that the observed differences in output between U.S. and non-U.S. scientists in our sample are principally due to endogeneity bias, particularly for more wealthy countries. This suggests that randomly assigning Ph.D. recipients to U.S. or non-U.S. institutions would have little impact on their post-Ph.D. productivity. This finding is quite surprising in light of the wide variation in support for research, quality of colleagues, and incentives for productivity at institutions around the world. However, it may reflect the dual factors of increasingly easy international collaboration and communication via the internet, and of the increasing numbers of research centers around the world in both the academic and private sectors.

While our estimates of the causal effect of location on research output are precisely estimated and close to zero for a sample that includes Ph.D. graduates taking jobs in rich countries, our estimates are less precise when it comes to the impact of location in poorer countries. Our focus for ongoing research will be on estimating this effect more precisely.

In addition to research output, we examined the impact of location on collaboration patterns and access to recent science. We found, in both uninstrumented and instrumented models, that scientists outside the U.S. have fewer collaborative relationships with scientists at U.S. institutions, and more relationships with scientists abroad and in their home countries. Distance does not appear to affect collaboration with past thesis advisors. Scientists outside the U.S., however, appear to cite a somewhat older literature, suggesting a barrier to access to the most recently published research.

References

Adams, J., R. Clemmons, and P. Stephan (2006). "How Rapidly Does Science Leak Out?" *NBER Working Paper No. 11997.*

Adams, James, Grant Black, Roger Clemmons, Paula Stephan (2005). "Scientific Teams and Institutional Collaborations: Evidence from U.S. Universities, 1981-1999." *Research Policy,* 34(3), 259-85.

Agrawal, Ajay, Iain Cockburn, and John McHale (2006). "Gone but not forgotten: Knowledge Flows, Labor Mobility, and Enduring Social Relationships." *Journal of Economic Geography* 6(5), 571-91.

Agrawal, Ajay, Devesh Kapur, and John McHale (2007). "Birds of a Feather - Better Together? Exploring the Optimal Spatial Distribution of Ethnic Inventors." *NBER Working Paper 12823.*

Angrist, Joshua (2001). "Estimations of Limited Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice." *Journal of Business and Economic Statistics*, 19(1), 2-16.

Audretsch, David and Maryann Feldman (1996). "R&D Spillovers and the Geography of Innovation and Production." *American Economic Review.* 86(3), 630-40.

Audretsch, David, Erik Lehmann and Susanne Warning (2005). "University spillovers and new firm location." *Research Policy* 34, 1113–1122.

Blackburn, Robert T., Charles E. Behymer; David E. Hall (1978). "Research Note: Correlates of Faculty Publications" *Sociology of Education,* 132-141.

Blundell, Richard, Rachel Griffith and Frank Windmeijer (2002). "Individual effects and dynamics in count data models," *Journal of Econometrics*, 108(1), 113-131.

Bound, John, Sarah Turner and Patrick Walsh (2006). *Internationalization of U.S. Doctorate Education*. NBER: mimeo.

Freeman, Richard (2006). "Does Globalization of the Scientific/Engineering Workforce Threaten U.S. Economic Leadership?" in Adam Jaffe, Josh Lerner and Scott Stern editors. *Innovation Policy and the Economy.* Cambridge MA: NBER/MIT Press.

Jaffe, Adam, Manuel Trajtenberg, and Rebecca Henderson (1993). "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *Quarterly Journal of Economics*, 108:3, 577–598.

Jaffe, Adam (1989). “Real Effects of Academic Research.” *American Economic Review* 79(5), 957-70.

Kerr, William R. (forthcoming)"Ethnic Scientific Communities and International Technology Diffusion." *Review of Economics and Statistics.*

Kim, E. Han, Adair Morse and Luigi Zingales (2006). Are Elite Universities Losing Their Competitive Edge? *NBER Working Paper 12245.*

Keller, Wolfgang (2002). “Geographic Localization of International Technology Diffusion.” *The American Economic Review* 92(1), pp. 120-142.

Mullahy J. (1997) “Instrumental variable estimation of Poisson regression models: application to models of cigarette smoking behavior”, *Review of Economics and Statistics* 79, 586-593.

Orlando, Michael J. (2004) “Measuring spillovers from industrial R&D: on the importance of geographic and technological proximity.” *Rand Journal of Economics*, 35(4): 777-786.

Oyer, Paul (2006). “Initial Labor Market Conditions and Long-Term Outcomes for Economists.” *Journal of Economic Perspectives*, 20(3), 143.

Saxenian, Annalee (2002). “Brain Drain or Brain Circulation: How High-Skill Immigration Makes Everyone Better Off.” *The Brookings Review*, Winter 20(1): 28-31.

Shanghai Jiao Tong University. *Academic Rankings of World Universities* 2003, 2007, and 2008. *www.arwu.org*.

Singh, Jasjit (2005). “Collaborative Networks as Determinants of Knowledge Diffusion Patterns” *Management Science*, 51(5): 756-770.

Staiger, D., and J. H. Stock (1997) “Instrumental Variables Regression with Weak Instruments.” *Econometrica*, 65, 557-586.

Stephan, Paula and Sharon Levin (1992). *Striking the Mother Lode in Science: The Importance of Age, Place, and Time.* New York and Oxford: Oxford University Press.

Stephan, Paula (1996) “The Economics of Science.” *Journal of Economic Literature*, 34(3), 1199-1235.

Stock, James H. and Motohiro Yogo (2005) “Testing for Weak Instruments in IV Regression,” in *Identification and Inference for Econometric Models: A Festschrift in Honor of Thomas Rothenberg*. Donald W.K. Andrews and James H. Stock, eds. Cambridge University Press, 80–108.

Toole, Andrew and Dirk Czarnitzki (2007). “Biomedical Academic Entrepreneurship through the SBIR Program.” *Journal of Economic Behavior and Organization.* 63(4), 716-38.

Thompson, Peter; Fox-Kean, Melanie (2005). “Patent Citations and the Geography of Knowledge Spillovers: A Reassessment.” *American Economic Review* 95 (4): 461-464.

Windmeijer, Frank (2006) "GMM for panel count data models," CeMMAP working papers CWP21/06, Centre for Microdata Methods and Practice, Institute for Fiscal Studies, revised.

Wooldridge, J. (1997) “Distribution-Free Estimation of Some Nonlinear Panel Data Models.” *Journal of Econometrics*, 90, 77-91.

Zucker, Lynne and Michael Darby (2007). Star Scientists, Innovation and Regional And National Immigration. *NBER Working Paper 13547.*

Zucker, Lynne and Michael Darby (2006). Movement of Star Scientists and Engineers and High-Tech Firm Entry. *NBER Working Papers 12172*

Zucker, Lynne, Michael Darby and Marilynn Brewer (1998). “Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises.” *American Economic Review*, 88(1): 290-306.

Zucker, Lynne, Michael Darby, Jonathan Furner, Robert Liu and Hongyan Ma (2007). “Minerva Unbound: Knowledge Stocks, Knowledge Flows and New Knowledge Production.” *Research Policy* 36, 850-63.

Figure 1

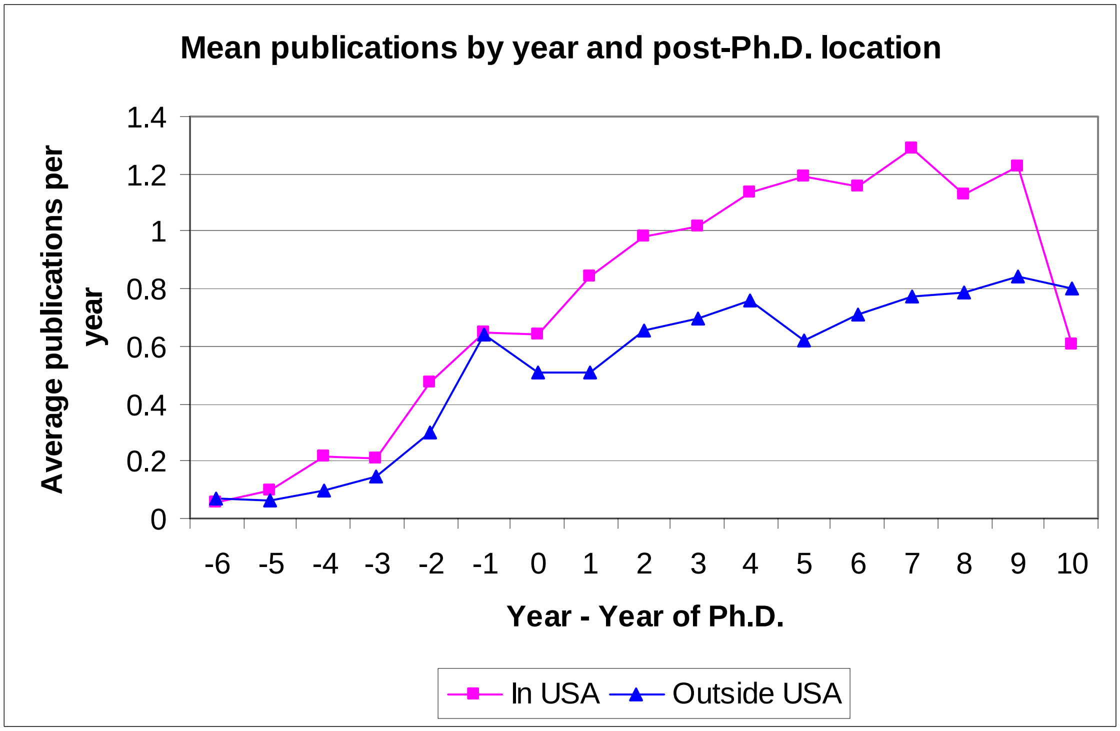
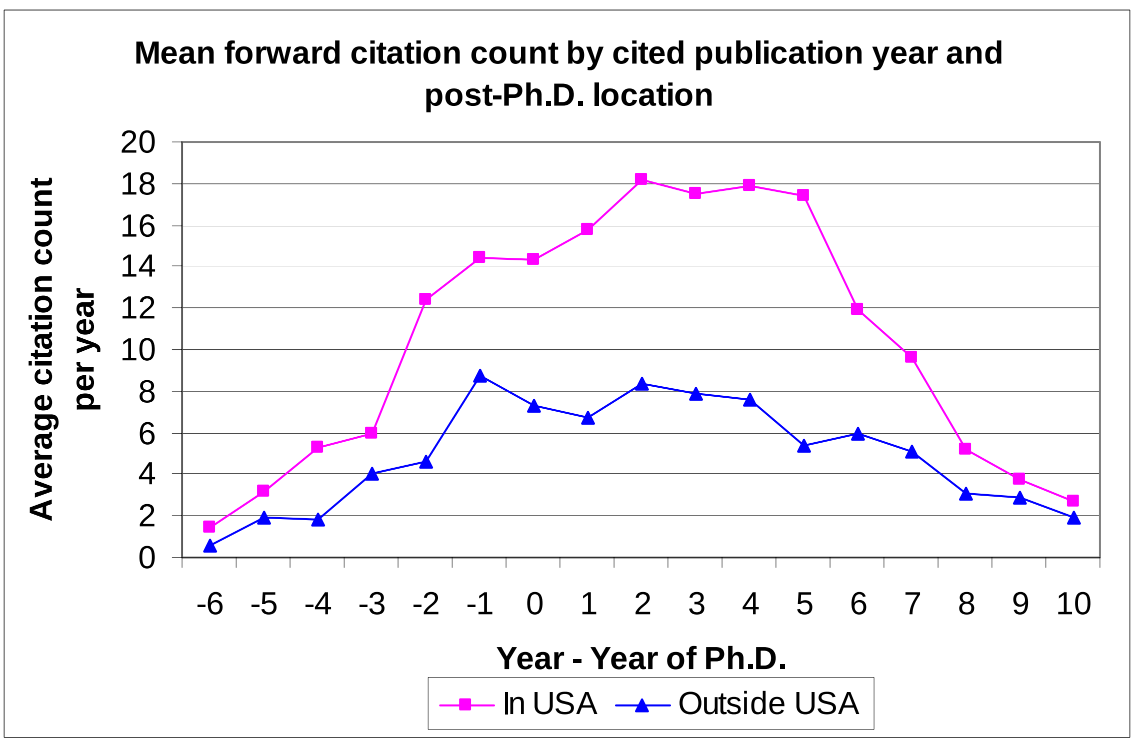


Figure 2



Note: In these graphs, scientists are classified as "Outside USA" if they were ever located outside the US during our sample period.

Figure 3

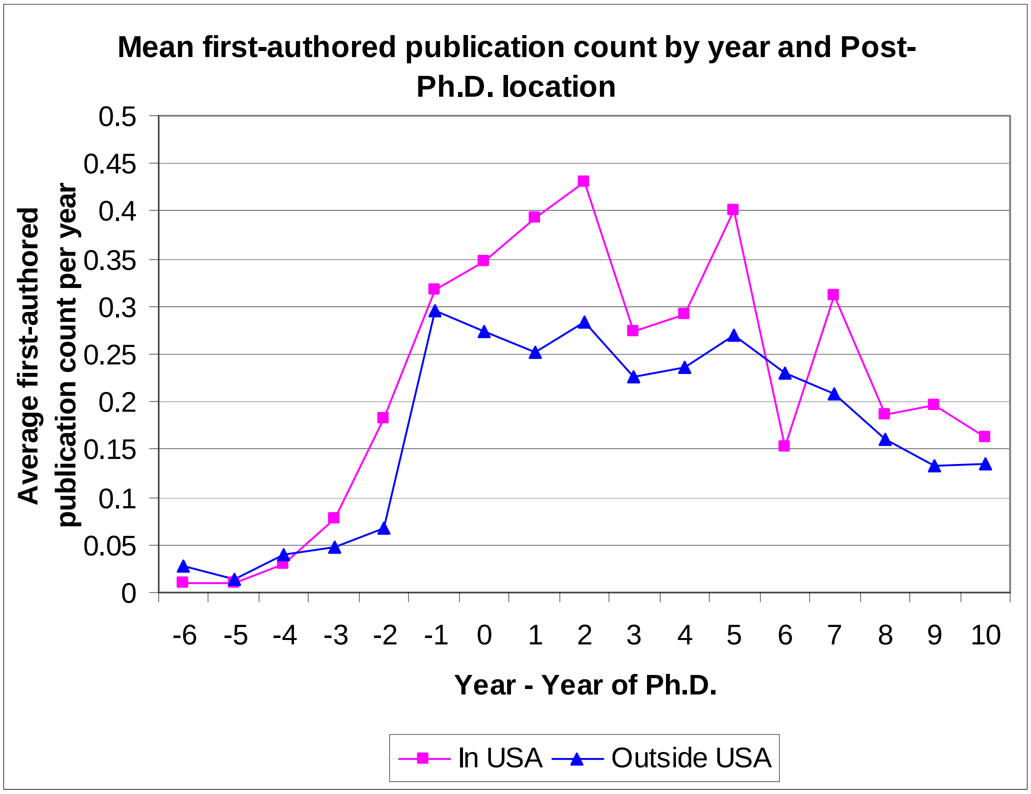
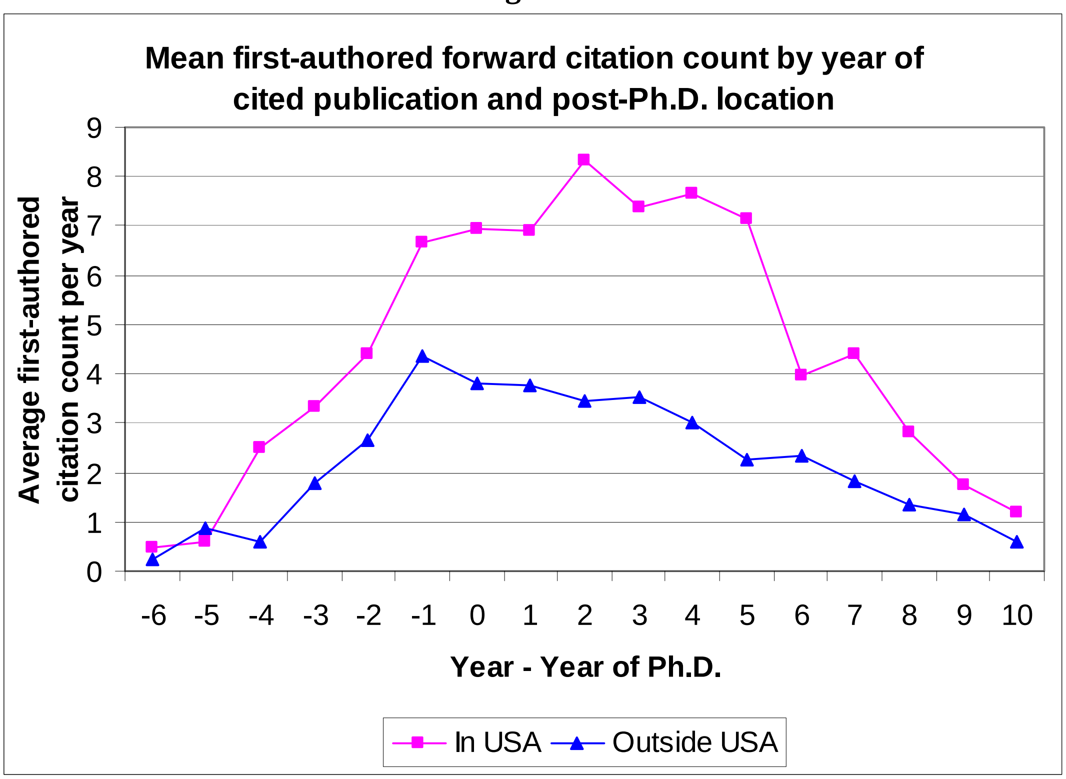


Figure 4



Note: In these graphs, scientists are classified as "Outside USA" if they were ever located outside the US during our sample period.

Table 1: Share of post-PhD years spent outside U.S.

	Mean	Std.Dev.	Count
Overall			
Control	0.392	0.488	1714
Fulbright	0.743	0.437	1626
Students originating in Asia			
Control	0.347	0.477	613
Fulbright	0.687	0.467	67
Students originating in Europe			
Control	0.432	0.496	630
Fulbright	0.753	0.432	699
Students originating in Latin America			
Control	0.538	0.500	251
Fulbright	0.778	0.416	559
Students originating in Middle East/Africa			
Control	0.240	0.428	221
Fulbright	0.668	0.472	301

Table 2: Number of Fulbright and Control Students, by country and region of origin

Region/Country of origin	Control	Fulbright	Total
Asia	**65**	**6**	**71**
Bangladesh	1	0	1
China	18	0	18
India	17	0	17
Indonesia	3	0	3
Japan	4	0	4
Korea	7	0	7
Pakistan	1	0	1
Philippines	3	1	4
Singapore	1	0	1
Sri Lanka	1	0	1
Taiwan	5	0	5
Thailand	4	5	9
Europe/Canada/Aust/NZ	**64**	**72**	**136**
Australia	0	5	5
Austria	1	2	2
Bulgaria	1	1	1
Canada	8	8	8
Croatia	2	3	3
Czech Republic	3	4	4
Denmark	2	3	3
Finland	2	5	5
France	1	1	1
Germany	7	7	7
Greece	4	12	12
Hungary	1	2	2
Iceland	1	8	8
Ireland	2	3	3
Italy	2	5	5
Lithuania	0	1	1
Macedonia	1	1	1
Netherlands	3	8	8
Norway	1	5	6
Poland	0	1	1
Portugal	1	13	13
Romania	3	4	4
Russia	5	5	5
Spain	2	9	9
Sweden	1	4	4
Switzerland	1	2	2
UK	2	7	7
Ukraine	4	4	4
Yugoslavia	3	3	3

Region/Country of origin	Control	Fulbright	Total
Latin America	**27**	**67**	**94**
Argentina	3	4	7
Bolivia	0	1	1
Brazil	6	0	6
Chile	1	0	1
Colombia	4	9	13
Costa Rica	0	1	1
Guatemala	1	0	1
Haiti	0	1	1
Mexico	9	50	59
Peru	2	1	3
Venezuela	1	0	1
Middle East/Africa	**23**	**34**	**57**
Armenia	1	0	1
Botswana	0	1	1
Cote D'Ivoire	1	3	4
Egypt	2	0	2
Ethiopia	2	4	6
Ghana	1	2	3
Iraq	1	0	1
Israel	2	2	4
Jordan	1	0	1
Kenya	0	3	3
Lesotho	0	1	1
Malawi	0	1	1
Morocco	0	2	2
Nigeria	1	0	1
Solomon Islands	0	1	1
South Africa	0	7	7
Tanzania	1	1	2
Togo	0	2	2
Turkey	6	1	7
Uganda	1	3	4
Zaire	1	0	1
Zimbabwe	2	0	2

Table 3: Number of Fulbrights and Controls, by Field of Study

Field of Study	Controls	Fulbrights	Total
Agricultural Sciences/Natural Resources	20	24	44
Biological/Biomedical/Health Sciences	44	43	87
Engineering	43	41	84
Computer & Information Sciences	13	11	24
Mathematics	12	15	27
Chemistry	6	4	10
Geological & Earth Sciences	9	9	18
Physics	19	19	38
Ocean/Marine/Enviro Sci	13	13	26
Total	179	179	358

Table 4: Number of Fulbrights and Controls, by Year of Ph.D.

Year of Ph.D.	Controls	Fulbrights	Total
1991	1	0	1
1992	2	0	2
1993	5	4	9
1994	10	15	25
1995	7	17	24
1996	22	18	40
1997	37	23	60
1998	27	26	53
1999	23	26	49
2000	17	14	31
2001	10	18	28
2002	7	9	16
2003	6	6	12
2004	3	2	5
2005	2	1	3
Total	179	179	358

Table 5: Publication variables, by post-Ph.D. location and Fulbright status

Variable	**Obs**	**Mean**	**Std.Dev.**	**Min**	**Max**
Scientists located in USA					
Publications	1459	1.085675	2.734012	0	34
Fwd citations	1459	14.1172	40.9138	0	227
First-authored publications	1459	0.361892	0.805894	0	8
Fwd cites to first-authored publications	1459	6.134339	16.70111	0	63
Number of co-authors	1459	4.836189	11.05386	0	52
Pubs with at least one U.S. collaborative relationship	1459	0.7	2.26494	0	32
Pubs with at least one non-U.S. collaborative relationship	1459	0.375343	1.794117	0	32
Pubs with at least one collaborative relationship in home country	1459	0.110959	0.789767	0	13
Pubs with at least one collaborative relationship in third country	1459	0.339041	1.760346	0	32
Pubs with advisor	1459	0.267808	1.311699	0	25
Median citation lag	503	6.99006	3.506667	1	21
Scientists located outside USA					
Publications	1881	0.641148	2.271795	0	48
Fwd citations	1881	4.73631	20.77922	0	227
First-authored publications	1881	0.196704	0.636321	0	9
Fwd cites to first-authored publications	1881	2.049442	9.372617	0	63
Number of co-authors	1881	2.313663	6.695781	0	52
Pubs with at least one U.S. collaborative relationship	1881	0.302499	1.591346	0	44
Pubs with at least one non-U.S. collaborative relationship	1881	0.3126	1.951567	0	45
Pubs with at least one collaborative relationship in home country	1881	0.200425	1.785795	0	45
Pubs with at least one collaborative relationship in third country	1881	0.197767	1.416549	0	39
Pubs with advisor	1881	0.1042	0.42123	0	4
Median citation lag	457	8.944201	4.998207	1	34.5
Control Scientists					
Publications	1776	0.833896	2.535606	0	48
Fwd citations	1776	8.673986	30.20195	0	227
First-authored publications	1776	0.258446	0.649322	0	5
Fwd cites to first-authored publications	1776	3.947635	13.58487	0	63
Number of co-authors	1776	3.583896	9.274332	0	52
Pubs with at least one U.S. collaborative relationship	1776	0.517727	2.178376	0	44
Pubs with at least one non-U.S. collaborative relationship	1776	0.339899	2.072816	0	45
Pubs with at least one collaborative relationship in home country	1776	0.158132	1.633435	0	45
Pubs with at least one collaborative relationship in third country	1776	0.288689	1.934296	0	39
Pubs with advisor	1776	0.159257	0.618744	0	9
Median citation lag	515	7.881553	4.441858	1	34.5
Fulbright Scientists					
Publications	1765	0.768839	2.324532	0	34
Fwd citations	1765	8.301983	31.47871	0	227
First-authored publications	1765	0.264023	0.76375	0	9
Fwd cites to first-authored publications	1765	3.455524	12.40368	0	63
Number of co-authors	1765	2.935977	8.151686	0	52
Pubs with at least one U.S. collaborative relationship	1765	0.393768	1.504591	0	25
Pubs with at least one non-U.S. collaborative relationship	1765	0.31898	1.566264	0	29
Pubs with at least one collaborative relationship in home country	1765	0.159773	1.131627	0	25
Pubs with at least one collaborative relationship in third country	1765	0.206232	0.97559	0	16
Pubs with advisor	1765	0.175071	1.115979	0	25
Median citation lag	471	8.008493	4.303998	1	33

note: the counts of forward citations and co-authors are truncated at the 99th percentile of the distributions due to a small number of extreme outliers.

Table 6: Publication Counts by year

Estimation method: Poisson
Fixed effects for year of publication and years since Ph.D. included;
robust standard errors clustered by scientist in parentheses.

	(1)	(2)	(3)	(4)	(5)
	articles published in year t		first-authored articles published in year t		
FORLOC lagged	-0.214 (0.179)		-0.355 (0.175)**		
FORLOC lagged 2 years		-0.203 (0.182)		-0.311 (0.185)*	-0.195 (0.112)*
ln pub count					1.689 (0.097)***
ln real GDP per cap. of home country	-0.082 (0.116)	-0.084 (0.120)	-0.149 (0.126)	-0.156 (0.136)	-0.077 (0.088)
Number of articles published in grad school	0.120 (0.023)***	0.119 (0.025)***	0.007 (0.031)	0.013 (0.035)	-0.143 (0.033)***
First-authored articles published in grad school	0.102 (0.057)*	0.100 (0.061)	0.297 (0.063)***	0.280 (0.070)***	0.241 (0.044)***
Region of origin: Europe	-0.072 (0.348)	-0.062 (0.366)	-0.100 (0.330)	-0.144 (0.366)	0.105 (0.204)
Region of origin: Latin Am.	-0.583 (0.301)*	-0.608 (0.320)*	-0.719 (0.342)**	-0.724 (0.375)*	-0.054 (0.230)
Region of origin: Middle East/Africa	-0.774 (0.316)**	-0.800 (0.339)**	-0.465 (0.313)	-0.456 (0.355)	0.395 (0.197)**
Employed in Government	-0.891 (0.260)***	-1.007 (0.277)***	-0.365 (0.336)	-0.505 (0.372)	0.189 (0.306)
Employed in Industry	-0.777 (0.264)***	-0.879 (0.274)***	-0.666 (0.268)**	-0.751 (0.308)**	-0.196 (0.190)
Bio/Biomed/Health Sciences	-0.637 (0.346)*	-0.660 (0.371)*	-0.435 (0.364)	-0.382 (0.423)	0.351 (0.221)
Engineering	-0.481 (0.306)	-0.491 (0.329)	-0.312 (0.312)	-0.261 (0.357)	0.435 (0.184)**
Computer & Information Sciences	-0.617 (0.594)	-0.579 (0.619)	-0.859 (0.433)**	-0.755 (0.474)	-0.155 (0.445)
Mathematics	-0.355 (0.352)	-0.334 (0.368)	0.473 (0.350)	0.669 (0.398)*	1.109 (0.178)***
Chemistry	-0.035 (0.491)	-0.032 (0.532)	-0.398 (0.575)	-0.348 (0.650)	0.103 (0.250)
Geological & Earth Sciences	-0.779 (0.418)*	-0.819 (0.473)*	-0.438 (0.453)	-0.701 (0.574)	0.227 (0.481)
Physics	-0.523 (0.436)	-0.545 (0.450)	-0.429 (0.394)	-0.353 (0.448)	0.301 (0.270)
Ocean/Marine/Enviro Sci	-0.453 (0.388)	-0.462 (0.414)	0.442 (0.416)	0.497 (0.478)	1.060 (0.238)***
Constant	0.565 (0.989)	0.445 (1.039)	0.486 (1.049)	0.477 (1.131)	-1.766 (0.801)**
Observations	2998	2654	2998	2654	2654

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 7: Forward citation counts as of 2008

Estimation method: Poisson
Fixed effects for year of publication and years since Ph.D. included;
robust standard errors clustered by scientist in parentheses.

	(1)	(2)	(3)	(4)	(5)	(6)
	fwd cites to articles published in year t			fwd cites to first-authored articles published in year t		
FORLOC lagged	-0.639 (0.268)**			-0.680 (0.266)**		
FORLOC lagged 2 years		-0.733 (0.294)**	-0.563 (0.192)***		-0.810 (0.295)***	-0.699 (0.233)***
ln pub count			1.781 (0.107)***			
ln first-authored pub count						2.038 (0.169)***
ln real GDP per cap. of home country	-0.040 (0.199)	-0.050 (0.205)	-0.005 (0.156)	-0.178 (0.201)	-0.184 (0.212)	-0.193 (0.153)
number of articles published in grad school	0.123 (0.029)***	0.119 (0.032)***	-0.050 (0.017)***	0.087 (0.030)***	0.080 (0.032)**	0.137 (0.023)***
first-authored articles published in grad school	0.069 (0.053)	0.076 (0.059)	0.094 (0.033)***	0.131 (0.061)**	0.137 (0.066)**	-0.131 (0.042)***
Region of origin: Europe	0.299 (0.449)	0.277 (0.478)	0.303 (0.272)	0.283 (0.470)	0.219 (0.519)	0.517 (0.334)
Region of origin: Latin Amer	-0.404 (0.327)	-0.403 (0.361)	0.141 (0.195)	-0.527 (0.378)	-0.489 (0.408)	-0.085 (0.276)
Region of origin: ME/Africa	-0.870 (0.345)**	-0.910 (0.386)**	-0.223 (0.258)	-0.742 (0.389)*	-0.736 (0.451)	-0.433 (0.286)
Employed in Government	-0.872 (0.361)**	-0.809 (0.367)**	-0.063 (0.277)	-0.746 (0.398)*	-0.703 (0.411)*	-0.239 (0.373)
Employed in Industry	-0.308 (0.331)	-0.550 (0.397)	-0.237 (0.260)	-0.367 (0.316)	-0.594 (0.388)	-0.281 (0.314)
Bio/Biomed/Health Sciences	-0.112 (0.347)	-0.216 (0.391)	0.087 (0.232)	-0.162 (0.371)	-0.170 (0.451)	0.021 (0.265)
Engineering	-0.634 (0.363)*	-0.584 (0.395)	-0.244 (0.275)	-0.723 (0.393)*	-0.640 (0.444)	-0.404 (0.323)
Computer & Information Sci	-0.651 (0.834)	-0.517 (0.849)	-0.148 (0.387)	-0.952 (0.820)	-0.729 (0.841)	-0.192 (0.784)
Mathematics	-0.185 (0.807)	-0.208 (0.811)	0.060 (0.718)	-0.330 (0.711)	-0.202 (0.738)	-0.568 (0.724)
Chemistry	0.432 (0.441)	0.264 (0.508)	0.033 (0.241)	0.298 (0.483)	0.285 (0.566)	0.132 (0.338)
Geological & Earth Sciences	-0.968 (0.496)*	-1.213 (0.543)**	-0.451 (0.430)	-1.230 (0.721)*	-1.984 (0.738)***	-1.483 (0.618)**
Physics	0.037 (0.473)	0.041 (0.519)	0.425 (0.311)	-0.085 (0.464)	0.038 (0.531)	0.157 (0.405)
Ocean/Marine/Enviro Sci	-0.098 (0.485)	-0.112 (0.530)	0.098 (0.274)	0.139 (0.504)	0.192 (0.571)	0.105 (0.335)
Constant	2.754 (1.641)*	2.696 (1.717)	1.116 (1.281)	3.524 (1.631)**	3.812 (1.757)**	2.877 (1.270)**
Observations	2998	2654	2654	2998	2654	2654

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 8: Foreign location dummy broken down by current region

Estimation method: Poisson
Fixed effects for year of publication and years since Ph.D. included;
robust standard errors clustered by scientist in parentheses.

	(1) articles published in year t	(2) fwd cites to articles published in year t	(3) first-authored articles published in year t	(4) fwd cites to first-authored articles published in year t	(5) articles published in year t	(6) fwd cites to articles published in year t	(7) first-authored articles published in year t	(8) fwd cites to first-authored articles published in year t
FORLOC lagged 2 yrs: Asia, ME, Africa, Latin America	-0.312 (0.219)	-0.921 (0.293)***	-0.564 (0.217)***	-1.069 (0.300)***				
FORLOC lagged 2 yrs: Europe, Canada, Aus & NZ	-0.236 (0.256)	-0.539 (0.440)	-0.253 (0.232)	-0.508 (0.414)	**Dropping students from Europe/Can/Aus/NZ**			
FORLOC lagged 2 years					-0.155 (0.246)	-0.910 (0.354)**	-0.540 (0.225)**	-1.219 (0.392)***
ln real GDP per cap. of home country	0.016 (0.119)	0.148 (0.158)	-0.158 (0.093)*	-0.051 (0.128)	-0.083 (0.133)	-0.077 (0.171)	-0.191 (0.169)	-0.121 (0.205)
Number of articles published in grad school	0.138 (0.025)***	0.132 (0.033)***	0.020 (0.034)	0.089 (0.032)***	0.115 (0.026)***	0.108 (0.032)***	0.000 (0.039)	0.061 (0.035)*
First-authored articles published in grad school	0.086 (0.056)	0.078 (0.058)	0.283 (0.065)***	0.150 (0.062)**	0.163 (0.088)*	0.140 (0.109)	0.389 (0.095)***	0.274 (0.114)**
Employed in Govt.	-0.909 (0.258)***	-0.661 (0.342)*	-0.423 (0.377)	-0.540 (0.392)	-1.143 (0.558)**	-1.060 (0.830)	-1.327 (0.760)*	-2.815 (0.882)***
Employed in Industry	-0.855 (0.282)***	-0.502 (0.382)	-0.766 (0.320)**	-0.549 (0.377)	-0.820 (0.463)*	-0.989 (0.527)*	-0.641 (0.487)	-0.966 (0.512)*
Bio/Biomed/Health Sci	-0.640 (0.393)	-0.167 (0.374)	-0.385 (0.433)	-0.118 (0.416)	-0.964 (0.472)**	-0.612 (0.505)	-0.535 (0.589)	-0.548 (0.673)
Engineering	-0.414 (0.349)	-0.543 (0.426)	-0.270 (0.371)	-0.622 (0.443)	-0.732 (0.374)*	-0.744 (0.510)	-0.567 (0.408)	-0.937 (0.614)
Computer & Info. Sci	-0.479 (0.630)	-0.415 (0.857)	-0.762 (0.487)	-0.628 (0.843)	-0.181 (0.712)	-0.009 (0.883)	-0.708 (0.598)	-0.296 (0.936)
Mathematics	-0.189 (0.389)	-0.085 (0.832)	0.705 (0.420)*	-0.073 (0.747)	-0.078 (0.553)	-2.707 (0.612)***	0.760 (0.617)	-1.847 (0.675)***
Chemistry	0.161 (0.501)	0.368 (0.495)	-0.222 (0.581)	0.409 (0.517)	-0.194 (0.705)	0.127 (0.570)	-0.754 (0.969)	-0.010 (0.756)

Geological & Earth Sciences	-0.865	-1.199	-0.747	-1.982	-1.701	-2.139	-1.605	-19.376
	(0.484)*	(0.546)**	(0.584)	(0.738)***	(0.667)**	(1.029)**	(0.700)**	(0.684)***
Physics	-0.450	0.150	-0.308	0.150	-1.057	-0.723	-0.570	-0.885
	(0.454)	(0.505)	(0.444)	(0.503)	(0.692)	(0.698)	(0.752)	(0.852)
Ocean/Marine/Envir Sci	-0.438	-0.091	0.450	0.219	-0.495	-0.192	0.505	0.073
	(0.410)	(0.521)	(0.430)	(0.517)	(0.446)	(0.574)	(0.559)	(0.679)
Region of origin: Latin Am.					-0.612	-0.442	-0.586	-0.433
					(0.281)**	(0.332)	(0.339)*	(0.379)
Region of origin: Middle East/Africa					-0.852	-1.074	-0.516	-0.820
					(0.308)***	(0.373)***	(0.400)	(0.473)*
Constant	-0.708	0.829	0.286	2.473	0.655	3.242	1.231	3.598
	(1.082)	(1.440)	(0.858)	(1.201)**	(1.004)	(1.361)**	(1.316)	(1.574)**
Observations	2654	2654	2654	2654	1589	1589	1589	1589

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 9: Median citation lag

Fixed effects for year of publication and years since Ph.D. included in Poisson estimates.
Year and years since Ph.D. included in GMM estimates.
Robust standard errors clustered by scientist in parentheses .

	(1)	(2)	(3)	(4)
Estimation method:		Poisson ML		IV GMM
FORLOC lagged 2 years	0.202 (0.057)***	0.196 (0.057)***	0.140 (0.055)**	0.120 (0.072) *
ln pub count		-0.128 (0.042)***	0.020 (0.048)	0.000 (0.022)
ln fwd cite count			-0.100 (0.020)***	-0.044*** (0.011) ***
ln real GDP per cap. of home country	-0.007 (0.048)	-0.009 (0.047)	-0.001 (0.045)	0.000 (0.015)
Number of articles published in grad school	-0.014 (0.008)	-0.008 (0.008)	-0.006 (0.008)	-0.004 (0.003)
First-authored articles published in grad school	0.001 (0.017)	-0.000 (0.017)	-0.002 (0.018)	0.003 (0.006)
Employed in Government	0.255 (0.085)***	0.220 (0.085)***	0.220 (0.078)***	0.105 (0.039) ***
Employed in Industry	-0.005 (0.097)	-0.016 (0.097)	-0.005 (0.085)	-0.024 (0.030)
Biological/Biomedical/Health Sciences	-0.114 (0.071)	-0.126 (0.073)*	-0.107 (0.069)	-0.013 (0.025)
Engineering	-0.100 (0.092)	-0.121 (0.093)	-0.161 (0.087)*	-0.039 (0.030)
Computer & Information Sciences	-0.322 (0.139)**	-0.341 (0.133)**	-0.422 (0.130)***	-0.155 (0.049) ***
Mathematics	0.140 (0.114)	0.118 (0.114)	0.044 (0.099)	0.079 (0.039) **
Chemistry	-0.107 (0.100)	-0.096 (0.096)	-0.097 (0.092)	-0.008 (0.036)
Geological & Earth Sciences	0.033 (0.162)	0.020 (0.168)	-0.010 (0.168)	0.032 (0.063)
Physics	-0.206 (0.111)*	-0.224 (0.110)**	-0.212 (0.103)**	-0.062 (0.038)
Ocean/Marine/Enviro Sci	-0.143 (0.103)	-0.161 (0.106)	-0.159 (0.113)	
Region of origin: Europe	-0.019 (0.121)	-0.027 (0.121)	-0.017 (0.116)	-0.037 (0.035)
Region of origin: Latin America	-0.008 (0.078)	-0.041 (0.077)	-0.049 (0.077)	-0.035 (0.032)
Region of origin: Middle East/Africa	0.118 (0.092)	0.092 (0.094)	0.061 (0.094)	0.031 (0.034)
Constant	2.374 (0.432)***	2.517 (0.427)***	2.596 (0.416)***	18.487 (10.456)*
Observations	776	776	776	776

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 10: Collaboration Variables

Estimation method: OLS

Fixed effects for year of publication and years since Ph.D. included

Robust standard errors clustered by scientist in parentheses.

	(1) Ln(# coauthors)	(2) Share of pubs w collaboration with U.S.	(3) Share of pubs w collaboration with abroad	(4) Share of pubs w collaboration with home	(5) Share of pubs w collaboration with third co	(6) Share of pubs w collaboration with advisor
FORLOC lagged 2 years	-0.138	-0.207	0.183	0.159	0.068	-0.047
	(0.069)**	(0.042)***	(0.045)***	(0.042)***	(0.043)	(0.041)
ln real GDP per cap. of home country	-0.029	-0.043	0.051	0.031	0.030	-0.060
	(0.048)	(0.033)	(0.030)*	(0.030)	(0.025)	(0.031)*
Number of articles published in grad school	0.084	0.015	0.017	-0.001	0.019	-0.003
	(0.010)***	(0.006)***	(0.010)*	(0.008)	(0.009)**	(0.007)
First-authored articles published in grad school	-0.087	-0.022	-0.020	0.002	-0.016	0.001
	(0.024)***	(0.011)*	(0.016)	(0.016)	(0.018)	(0.013)
Region of origin: Europe	-0.023	0.120	-0.034	-0.057	-0.004	0.131
	(0.137)	(0.083)	(0.073)	(0.067)	(0.073)	(0.079)*
Region of origin: Latin America	-0.370	-0.029	-0.140	-0.107	-0.105	0.038
	(0.128)***	(0.067)	(0.073)*	(0.066)	(0.063)*	(0.061)
Region of origin: Middle East/Africa	-0.244	0.028	-0.065	-0.116	0.041	0.082
	(0.123)**	(0.070)	(0.064)	(0.056)**	(0.068)	(0.083)
Bio/Biomed/Health Sci	-0.023	0.029	-0.083	-0.123	0.106	-0.002
	(0.114)	(0.068)	(0.068)	(0.060)**	(0.063)*	(0.079)
Engineering	-0.462	-0.052	-0.200	-0.114	-0.044	-0.007
	(0.108)***	(0.080)	(0.060)***	(0.058)**	(0.058)	(0.088)
Computer & Information Sciences	-0.307	-0.075	-0.117	0.007	0.059	-0.070
	(0.284)	(0.109)	(0.126)	(0.127)	(0.123)	(0.098)
Mathematics	-0.764	-0.156	-0.166	-0.097	-0.037	-0.144
	(0.143)***	(0.090)*	(0.112)	(0.103)	(0.079)	(0.084)*
Chemistry	0.177	-0.117	-0.007	0.103	-0.016	-0.154
	(0.292)	(0.110)	(0.151)	(0.142)	(0.098)	(0.095)
Geological & Earth Sciences	-0.388	0.025	-0.122	-0.201	0.123	0.153
	(0.156)**	(0.133)	(0.107)	(0.075)***	(0.144)	(0.127)
Physics	-0.155	-0.005	-0.018	-0.096	0.121	-0.177
	(0.156)	(0.086)	(0.106)	(0.079)	(0.104)	(0.078)**
Ocean/Marine/Enviro Sci	-0.431	0.251	-0.206	-0.184	0.006	-0.001
	(0.143)***	(0.075)***	(0.087)**	(0.090)**	(0.068)	(0.090)
Employed in Government	-0.290	0.069	-0.079	-0.178	0.093	0.080
	(0.121)**	(0.104)	(0.070)	(0.053)***	(0.102)	(0.091)
Employed in Industry	-0.097	0.047	-0.056	-0.148	0.045	0.013
	(0.115)	(0.067)	(0.063)	(0.042)***	(0.074)	(0.067)
Constant	2.089	1.139	-0.386	-0.180	-0.343	0.982
	(0.429)***	(0.303)***	(0.246)	(0.230)	(0.214)	(0.281)***
Observations	793	793	793	793	793	793
R-squared	0.43	0.20	0.21	0.16	0.14	0.17

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 11: I.V. results on publications and forward citations

Estimation method: LIML
Dependent variables measured in logs
Fixed effects for year of publication and years since Ph.D. included;
robust standard errors clustered by scientist in parentheses.
FORLOC lagged 2 years instrumented by Fulbright status.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	ln pub count	ln fwd cite count		Ln pub count	Ln fwd cite count	
First-stage F-stat	36.91	36.91	36.90	36.91	36.91	36.93
FORLOC lagged 2 years	0.019 (0.124)	-0.051 (0.260)	-0.082 (0.137)	0.009 (0.070)	0.073 (0.198)	0.058 (0.127)
Ln pub count			1.605 (0.075)***			
Ln first-authored pub count						1.811 (0.098)***
Ln real GDP per cap. of home country	-0.051 (0.027)*	-0.065 (0.060)	0.017 (0.039)	-0.032 (0.019)*	-0.077 (0.052)	-0.020 (0.032)
Number of articles published in grad school	0.097 (0.022)***	0.1702 (0.038)***	0.017 (0.016)	-0.002 (0.011)	0.077 (0.024)***	0.079 (0.021)***
First-authored articles published in grad school	0.085 (0.041)**	0.185 (0.084)**	0.049 (0.047)	0.119 (0.028)***	0.184 (0.063)***	-0.031 (0.041)
Region of origin: Europe	-0.022 (0.070)	0.003 (0.143)	0.037 (0.071)	-0.006 (0.039)	0.027 (0.119)	0.037 (0.077)
Region of origin: Latin Am	-0.130 (0.069)*	-0.177 (0.134)	0.032 (0.063)	-0.075 (0.038)**	-0.178 (0.105)*	-0.042 (0.066)
Region of origin: Middle East/Africa	-0.183 (0.071)***	-0.330 (0.137)**	-0.036 (0.077)	-0.054 (0.045)	-0.176 (0.112)	-0.078 (0.064)
Bio/Biomed/Health Sci	-0.039 (0.069)	-0.008 (0.147)	0.054 (0.081)	0.010 (0.038)	0.082 (0.101)	0.064 (0.065)
Engineering	-0.037 (0.062)	-0.176 (0.125)	-0.116 (0.069)*	0.011 (0.032)	-0.026 (0.083)	-0.045 (0.050)
Computer & Information Sci	-0.053 (0.085)	-0.294 (0.178)*	-0.208 (0.092)**	-0.010 (0.037)	-0.019 (0.123)	-0.001 (0.098)
Mathematics	0.081 (0.090)	-0.171 (0.232)	-0.301 (0.181)*	0.154 (0.058)***	0.070 (0.168)	-0.208 (0.156)
Chemistry	0.059 (0.176)	0.034 (0.299)	-0.061 (0.100)	-0.014 (0.090)	0.123 (0.260)	0.148 (0.141)
Geological & Earth Sciences	-0.087 (0.078)	-0.264 (0.155)*	-0.124 (0.071)*	-0.020 (0.034)	-0.150 (0.090)*	-0.114 (0.063)*
Physics	-0.014 (0.091)	0.056 (0.198)	0.077 (0.096)	0.002 (0.043)	0.105 (0.136)	0.102 (0.090)
Ocean/Marine/Enviro Sci	-0.014 (0.088)	-0.053 (0.203)	-0.031 (0.100)	0.100 (0.062)	0.131 (0.154)	-0.051 (0.083)
Employed in Government	-0.116 (0.044)***	-0.178 (0.099)*	0.008 (0.056)	-0.037 (0.030)	-0.123 (0.071)*	-0.057 (0.050)
Employed in Industry	-0.149 (0.045)***	-0.190 (0.100)*	0.048 (0.054)	-0.069 (0.025)***	-0.097 (0.071)	0.028 (0.049)
Constant	0.669 (0.227)***	1.219 (0.501)**	0.145 (0.343)	0.405 (0.165)**	1.090 (0.452)**	0.357 (0.293)
Observations	2654	2654	2654	2654	2654	2654

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 12: I.V. results on publications and forward citations

Estimation method: GMM

Dependent variables measured in levels;
Robust standard errors in parentheses.

	(1) Publications	(2) Fwd citations	(3) First-authored publications	(4) Fwd cites to first-authored publications
FORLOC lagged 2 years	0.152 (0.150)	0.030 (0.213)	-0.150 (0.219)	-0.057 (0.281)
Number of articles published In grad school	0.075*** (0.007)	0.077*** (0.009)	0.024** (0.010)	0.052*** (0.010)
First-authored articles published in grad school	0.093*** (0.016)	0.083*** (0.019)	0.180*** (0.022)	0.130*** (0.023)
Ln real GDP of home country	-0.081*** (0.031)	-0.047 (0.051)	-0.110** (0.044)	-0.113* (0.068)
Year	0.006 (0.009)	-0.060*** (0.014)	-0.038*** (0.012)	-0.077*** (0.017)
Years since Ph.D.	-0.006 (0.010)	-0.066*** (0.016)	-0.019 (0.014)	-0.067*** (0.021)
Region of origin: Europe	-0.031 (0.073)	0.020 (0.102)	0.032 (0.091)	0.041 (0.118)
Region of origin: Latin Am	-0.319*** (0.075)	-0.347*** (0.106)	-0.369*** (0.105)	-0.406*** (0.147)
Region of origin: Middle East/Africa	-0.422*** (0.076)	-0.570*** (0.103)	-0.179* (0.094)	-0.487*** (0.128)
Employed in Government	-0.494*** (0.094)	-0.482*** (0.144)	-0.182 (0.147)	-0.389** (0.198)
Employed in Industry	-0.420*** (0.063)	-0.354*** (0.100)	-0.332*** (0.083)	-0.318*** (0.117)
Bio/Biomed/Health Sci	-0.156** (0.070)	-0.125 (0.091)	-0.001 (0.094)	-0.050 (0.126)
Engineering	-0.230*** (0.066)	-0.466*** (0.094)	0.020 (0.089)	-0.385*** (0.116)
Computer & Information Sci	-0.109 (0.134)	-0.149 (0.219)	-0.237 (0.150)	-0.205 (0.234)
Mathematics	-0.043 (0.079)	-0.129 (0.175)	0.488*** (0.100)	-0.080 (0.168)
Chemistry	0.044 (0.122)	0.088 (0.161)	-0.091 (0.157)	0.141 (0.177)
Geological & Earth Sciences	-0.494*** (0.132)	-0.907*** (0.158)	-0.336 (0.213)	-1.335*** (0.291)
Physics	-0.143* (0.085)	-0.195 (0.127)	0.417*** (0.105)	0.014 (0.155)
Constant	-12.350 (18.655)	122.790*** (27.111)	77.259*** (23.328)	156.046*** (33.766)
Observations	2654	2654	2654	2654

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 13: I.V. results on collaboration

Estimation method: LIML

Fixed effects for year of publication and years since Ph.D. included;
robust standard errors clustered by scientist in parentheses.
Instrument for FORLOC: Fulbright dummy
First-stage F-statistic: 14.95

	(1) Ln(# coauthors)	(2) Share of pubs w collaboration with US	(3) Share of pubs w collaboration with abroad	(4) Share of pubs w collaboration with home	(5) Share of pubs w collaboration with third co	(6) Share of pubs w collaboration with advisor
FORLOC lagged 2 years	0.496 (0.305)	-0.330 (0.155)**	0.350 (0.169)**	0.262 (0.125)**	0.263 (0.177)	0.022 (0.167)
Ln real GDP per cap. of home country	-0.065 (0.061)	-0.036 (0.032)	0.041 (0.033)	0.025 (0.031)	0.018 (0.027)	-0.064 (0.030)**
Number of articles published in grad school	0.085 (0.011)***	0.015 (0.006)***	0.017 (0.010)*	-0.001 (0.007)	0.019 (0.010)**	-0.003 (0.007)
First-authored articles published in grad school	-0.078 (0.024)***	-0.024 (0.012)*	-0.017 (0.017)	0.004 (0.015)	-0.013 (0.018)	0.002 (0.014)
Region of origin: Europe	-0.056 (0.156)	0.127 (0.083)	-0.043 (0.077)	-0.062 (0.068)	-0.014 (0.078)	0.127 (0.080)
Region of origin: Latin Am	-0.537 (0.166)***	0.003 (0.073)	-0.184 (0.086)**	-0.134 (0.071)*	-0.156 (0.084)*	0.019 (0.073)
Region of origin: Middle East/Africa	-0.275 (0.122)**	0.034 (0.069)	-0.074 (0.065)	-0.121 (0.055)**	0.031 (0.069)	0.079 (0.082)
Bio/Biomed/Health Sci	0.084 (0.143)	0.008 (0.073)	-0.055 (0.067)	-0.106 (0.057)*	0.139 (0.076)*	0.010 (0.086)
Engineering	-0.493 (0.126)***	-0.046 (0.079)	-0.208 (0.059)***	-0.120 (0.054)**	-0.053 (0.070)	-0.011 (0.088)
Computer & Information Sci	-0.207 (0.305)	-0.095 (0.105)	-0.091 (0.131)	0.023 (0.127)	0.090 (0.137)	-0.059 (0.105)
Mathematics	-0.767 (0.166)***	-0.155 (0.083)*	-0.167 (0.112)	-0.097 (0.102)	-0.038 (0.089)	-0.144 (0.083)*
Chemistry	0.324 (0.322)	-0.145 (0.116)	0.031 (0.148)	0.127 (0.133)	0.029 (0.130)	-0.137 (0.101)
Geological & Earth Sciences	-0.620 (0.218)***	0.070 (0.132)	-0.184 (0.121)	-0.239 (0.085)***	0.052 (0.162)	0.127 (0.132)
Physics	-0.065 (0.179)	-0.023 (0.085)	0.006 (0.102)	-0.081 (0.076)	0.148 (0.108)	-0.167 (0.084)**
Ocean/Marine/Enviro Sci	-0.475 (0.197)**	0.260 (0.065)***	-0.218 (0.090)**	-0.191 (0.093)**	-0.008 (0.078)	-0.006 (0.088)
Employed in Government	-0.422 (0.185)**	0.095 (0.104)	-0.114 (0.079)	-0.200 (0.068)***	0.053 (0.102)	0.065 (0.095)
Employed in Industry	-0.109 (0.112)	0.049 (0.066)	-0.059 (0.064)	-0.150 (0.043)***	0.042 (0.075)	0.012 (0.067)
Constant	2.027 (0.537)***	1.152 (0.299)***	-0.403 (0.272)	-0.190 (0.244)	-0.362 (0.227)	0.975 (0.271)***
Observations	793	793	793	793	793	793

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 14: I.V. Results on publications and forward citations, excluding scientists from Europe

Estimation method: GMM

Dependent variables measured in levels; Robust standard errors in parentheses.

Dependent variable:	(1) Publications	(2) Fwd citations	(3) First-authored publications	(4) Fwd cites to first-authored publications
FORLOC lagged 2 years	0.239	-0.326	-0.303	-0.418
	(0.179)	(0.282)	(0.284)	(0.461)
Number of articles published in grad school	0.070***	0.067***	0.016	0.041***
	(0.008)	(0.010)	(0.012)	(0.012)
First-authored articles published in grad school	0.110***	0.114***	0.241***	0.183***
	(0.020)	(0.025)	(0.034)	(0.035)
ln real GDP per cap. of home Country	-0.085**	-0.049	-0.131***	-0.132*
	(0.036)	(0.059)	(0.051)	(0.078)
Region of origin: Latin America	-0.309***	-0.204*	-0.313***	-0.267
	(0.075)	(0.111)	(0.104)	(0.162)
Region of origin: Middle East/Africa	-0.391***	-0.512***	-0.115	-0.397***
	(0.075)	(0.100)	(0.096)	(0.130)
Employed in Government	-0.656***	-0.582***	-0.476*	-1.205***
	(0.159)	(0.225)	(0.281)	(0.450)
Employed in Industry	-0.386***	-0.489***	-0.294**	-0.447**
	(0.099)	(0.170)	(0.128)	(0.209)
Biological/Biomed/Health Sci	-0.169*	-0.123	-0.036	-0.020
	(0.099)	(0.136)	(0.143)	(0.221)
Engineering	-0.394***	-0.477***	-0.180	-0.428***
	(0.081)	(0.109)	(0.111)	(0.156)
Computer & Information Sci	0.157	0.150	-0.083	0.148
	(0.169)	(0.256)	(0.214)	(0.292)
Mathematics	0.033	-1.334***	0.643***	-0.845***
	(0.142)	(0.302)	(0.185)	(0.320)
Chemistry	0.133	0.154	-0.130	0.176
	(0.151)	(0.195)	(0.207)	(0.246)
Geological & Earth Sciences	-1.008***	-1.403***	-0.590**	-30.289
	(0.201)	(0.335)	(0.300)	0.000
Physics	-0.075	-0.055	0.482***	0.182
	(0.094)	(0.134)	(0.117)	(0.174)
Constant	19.483	109.561***	127.281***	148.882***
	(22.184)	(30.892)	(31.303)	(42.513)

* significant at 10%; ** significant at 5%; *** significant at 1%

Appendix A1: First-stage regressions of $FORLOC_{t-2}$ on control variables and instruments.

Preferred specification in column 3

Dummies for year and years since PhD included

	(1)	(2)	(3)	(4)
Adjusted R-squared	0.1227	0.1396	0.2314	0.2322
First-stage F-stat	50.86	51.90	36.91	18.59
Instruments:	Fulbright dummy	Fulbright dummy	Fulbright dummy	Fulbright & Fulbright X ln RGDP
Fulbright dummy	0.348***	0.348***	0.307***	0.630
	(0.049)	(0.048)	(0.051)	(0.478)
ln real GDP per cap. of home country			0.093**	0.107**
			(0.037)	(0.043)
number of articles published in grad school		-0.010	-0.004	-0.003
		(0.010)	(0.011)	(0.011)
first-authored articles published in grad schl		-0.028	-0.032	-0.032
		(0.024)	(0.027)	(0.027)
Fulbright dummy X ln real GDP				-0.035
			0.000	(0.051)
Region of origin: Europe			-0.061	-0.065
			(0.083)	(0.082)
Region of origin: Latin America			0.081	0.071
			(0.080)	(0.081)
Region of origin: Middle East/Africa			-0.054	-0.071
			(0.093)	(0.095)
Employed in Government			-0.153**	-0.148*
			(0.076)	(0.077)
Employed in Industry			-0.031	-0.025
			(0.075)	(0.074)
Biological/Biomedical/Health Sciences			-0.305***	-0.301***
			(0.098)	(0.098)
Engineering			-0.055	-0.044
			(0.109)	(0.109)
Computer & Information Sciences			-0.286*	-0.273
			(0.173)	(0.173)
Mathematics			0.000	-0.002
			(0.123)	(0.121)
Chemistry			-0.108	-0.105
			(0.091)	(0.090)
Geological & Earth Sciences			-0.065	-0.061
			(0.117)	(0.118)
Physics			0.170**	0.170**
			(0.078)	(0.079)
Ocean/Marine/Enviro Sci			-0.056	-0.054
			(0.063)	(0.063)
Constant			-0.100	-0.225
			(0.327)	(0.379)

* significant at 10%; ** significant at 5%; *** significant at 1%

Appendix B

DATA APPENDIX

Fulbright Data

The names of Fulbrights were obtained from volumes of *Foreign Fulbright Fellows: Directory of Students* published annually by the Institute for International Education (IIE) from 1993 to 1996.

Location Search Procedure

First, we entered data from the IIE volumes on the Fulbright Student's name, graduate institution, field of study, and country of origin. Then, we searched for these students in the *Proquest* database (described below) to find their date of graduation (for those who completed their studies) and advisor name. For those Fulbrights successfully completing their programs, we then performed searches on Google, Google Scholar, LinkedIn, and/or Web of Science to obtain as much information as possible on all the student's post-Ph.D. locations and affiliations. The search time was limited to 20 minutes. If a student was not found at all on the web within 20 minutes, the searcher moved on to the next name.

For the students found on the web, we then searched for controls. We focused on universities listing biographical information or prior degrees, because of the difficulty of otherwise finding information on students' countries of origin. We searched for controls obtaining Ph.D.s in the same year, with the same advisor, at the same institution as the Fulbright. Click on the name of the student's advisor. If this step failed (i.e. there are no foreign students with the same advisor graduating in same year), we looked for a student with the same advisor graduating within 3 years of the Fulbright. When choosing controls, we alternated students graduating before the Fulbright with those graduating after the Fulbright so that on average controls graduate at the same time as Fulbrights. If this step failed, we choose a control graduating in the same year in the same field of study (e.g. Biochemistry) at the same university.

For schools that did not list prior degrees, if we found a potential control student, we looked them up on the web. If we could find their current location and evidence that they came from a foreign country (i.e. foreign undergraduate degree or biography), we recorded their name, year of Ph.D., current location, and estimated country of origin.

Proquest Dissertations and Theses

The *Proquest Dissertations and Theses* database is a database of almost all dissertations filed at over 700 U.S. universities. We obtained information from this database on students' full names, advisors, fields of study, Ph.D. completion dates, and undergraduate institution and/or country of birth. Starting in the1990's, ProQuest began publishing online the full text of the first 24 pages of the dissertation.

Several universities require students to list biographical information in the front matter of the dissertation. Table A1 lists these universities, which were identified by checking dissertations filed at the universities that are major producers of scientists and engineers in the United States. At some universities, the information includes a full biographical sketch (e.g., Ohio State, NC State), but in most cases, the information is limited to a list of previous degrees. Figures A1 and A2 present examples of this information drawn from dissertations filed at the University of Illinois and the Ohio State University.

The biographical information contained in these dissertations can be used to identify the country of origin of the student. Under the assumption that most students attend undergraduate programs in their country of origin, we treat the country of undergraduate degree as the country of origin. Using this information as a proxy for the nationality of the student will of course introduce some error, since not all students receiving undergraduate degrees do so in their country of origin. However, evidence from the NSF's *Survey of Earned Doctorates* suggests that the country of undergraduate degree is a very good proxy for the country of origin. For students completing doctorates in 2003 and 2004, the *SED* lists the country of undergraduate degree. For 84.9% of students, the country of undergraduate degree is the same as the country of citizenship. However, there is considerable heterogeneity across countries in the extent to which students pursue undergraduate studies outside their countries of origin. Table A2 presents, for a selected list of countries, the share of students responding to the *SED*'s questions who remained in their home country for undergraduate study. Students from Germany and Japan have the lowest rates of staying at home among the major producers of U.S. graduate students (73% and 74%, respectively). However, the countries that send the most students (China, India, Taiwan, Korea, and Canada) have high stay-at-home rates for undergraduate study (98%, 93%, 89%, 76%, and 82%, respectively). Furthermore, counts of the number of doctoral recipients by country of origin, university and year computed from a ProQuest sample have a correlation of 0.948 with analogous counts obtained from the *SED.*

The data on country of origin is only available beginning in the late 1990's when universities began submitting digital copies of dissertations to be posted on the web by ProQuest. However, by 1996 or 1997 almost all dissertations are available in digital format.

Publication Data

We obtained publication histories from *ISI's Web of Science*. Authors were identified using information on post-Ph.D. locations, authors' middle names, and fields of research. For each publication by an author, we obtained all information available on the publication record itself, including publication year, title, co-author names, author locations, complete backward citations, counts of forward citations, publication source, abstract, specific field (for example, Marine & Freshwater Biology), and keywords.

It should be noted that our information on the number of forward citations received by an article includes self-citations. The median backward citation lag also includes self-citations. In future work, we intend to remove these citations. However, this requires downloading bibliographic data on each specific citing article, which is a very time-consuming process.

The *ISI Web of Science* database does not cover every scientific journal published worldwide. It lists articles from 6,650 scientific journals. Among Thomson's criteria for including a journal in the index are "The journal's basic publishing standards, its editorial content, the international diversity of its authorship, and the citation data associated with it."[14] Journals must typically publish on-time, implying a substantial backlog of articles forthcoming. They must publish bibliographic information in English, and must include full bibliographic information for cited references and must list address information for each author. Thomson also looks for international diversity among contributing authors, but regionally focused journals are evaluated on the basis of their specific contribution to knowledge. The number of citations received by the journal is a key factor in evaluation for inclusion in the index, with preference going to highly cited journals or journals whose contributing authors are cited highly elsewhere.

The ISI selection procedure is designed to select the most relevant scientific journals, independent of the location of their editorial offices. Since such a large share of cutting-edge science research takes place in the U.S., there will inevitably be a high share of journals in this index based in the U.S.. Journals that do not publish bibliographic information in English are less likely to be included, so articles written abroad and published in low-profile regional journals with limited readership beyond the region (as evidenced by a failure to publish bibliographic information in English) will be excluded from our data. As a result, our publication data should be viewed as information on scientists' participation in the international scientific community, rather than raw article counts. Still, the large number of journals included, and the special consideration given to regionally-focused journals means that most of the relevant journals in

[14] "The Thomson Scientific Journal Selection Process"
http://scientific.thomson.com/free/essays/selectionofmaterial/journalselection/ (accessed March 11, 2008)

which our scientists publish will be included. We examined the publication records of some of our scientists located outside the U.S., and found that even what might seem like relatively obscure journals (e.g. *Revista Chilena de Historia Natura, Revista Brasileira de Ciência do Solo, Acta Pharmacalogica Sinica*, etc.) were all included in the ISI index. We do not feel that a systematic bias is introduced by restricting our attention to journals included in the ISI index. However, as this project continues, we plan to verify this by comparing the CV's of our sample scientists against the ISI index.

We made sure to collect information on Fulbright and Control publications at the same time, ideally on the same day. We did this to avoid biasing the data to include more pubs and cites for one of the groups because they were collected later and had more time to appear in the database.

Table B1: Universities listing biographic info in thesis

AUBURN
BOSTON U
CALIFORNIA STATE U
CLARK
CORNELL U
FLORIDA INSTITUTE OF TECHNOLOGY
FORDHAM
GEORGE WASHINGTON U
GEORGETOWN U
KANSAS STATE
LOUISIANA STATE U
NC STATE
OH STATE
OK STATE
SYRACUSE
TEXAS A&M
U ARKANSAS
U CALIFORNIA
U CINCINATTI
U COLORADO
U CONNECTICUT
U FLORIDA
U ILLINOIS
U MAINE
U MASSACHUSETTS
U MASSACHUSETTS AT AMHERST
U MISSOURI
U NEVADA
U OREGON
U PITTSBURGH
U SOUTH ALABAMA
U SOUTH CAROLINA
U VIRGINIA

Table B2
Share of Ph.D. students at U.S. universities who received undergraduate degrees in their countries of citizenship

AUSTRALIA	85.00%
BRAZIL	96.02%
CANADA	82.51%
CHINA	98.35%
EGYPT	96.38%
FRANCE	82.05%
GERMANY	73.05%
GREECE	80.51%
INDIA	92.71%
IRAN	88.33%
ISRAEL	88.46%
JAPAN	73.51%
MEXICO	89.19%
NIGERIA	60.61%
PHILIPPINES	87.23%
SOUTH KOREA	76.33%
TAIWAN	89.19%
THAILAND	87.28%
TURKEY	95.57%
U.K.	63.64%
Weighted average across these countries	89.50%
Weighted average across all countries	84.79%

Figure B1

ALGORITHMS AND ARCHITECTURES FOR SOFT-DECODING REED-SOLOMON CODES

BY

ARSHAD AHMED

B.E., Regional Engineering College, Trichy, 1998
M.E., Indian Institute of Science, Bangalore, 2000

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2006

Urbana, Illinois

Figure B2

VITA

January 31, 1973 .... Born – Da-An, Jilin Province, China

September 1989 – July 1993 .... Bachelor of Science in Electrical Engineering, Nanjing University of Science and Technology, Nanjing, China

September 1993 – April 1996 .... Master of Science in Electrical Engineering, Nanjing University of Science and Technology, Nanjing, China

Septmeber 2002 – present .... Ph.D student, Analog VLSI Laboratory, Department of Electrical and Computer Engineering, the Ohio State University, Columbus, Ohio

Since June 2006 .... RFIC design engineer, Freescale Semiconductor Inc., Boca Raton, Florida

PUBLICATIONS

Research Publications

P. Zhang, and M. Ismail "A New RF Front-End and Frequency Synthesizer Architecture for 3.1–10.6 GHz MB-OFDM UWB Receivers", *Proc.* 48^{th} *Midwest Symposium on Circuit and System*, vol.2, pp.1119–1122, August 2005.

C. Garuda, X. Cui, P. Lin, S. Doo, P. Zhang, and M. Ismail "A 3–5 GHz Fully Differential CMOS LNA with Dual-gain Mode for Wireless UWB Applications", *Proc.* 48^{th} *Midwest Symposium on Circuit and System*, vol.1, pp.790–793, August 2005.

Y. Yu, L. Bu, S. Shen, B. Jalali-Farahani, G. Ghiaasi, P. Zhang, and M. Ismail "A 1.8V Fully Integrated Dual-band VCO for Zero-IF WiMAX/WLNA Receiver in 0.18μm CMOS", *Proc.* 48^{th} *Midwest Symposium on Circuit and System*, vol.1, pp.187–190, August 2005.

vii