How Important is U.S. Location for Research in Science?

Shulamit Kahn*

Megan J. MacGarvie**

December 22, 2008

Abstract

This paper asks whether scientists located outside the U.S. are at a disadvantage when it comes to research productivity. 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. We make 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 citations, and collaborators when compared with a control student with the same advisor. We examine their research output in terms of the number of publications, first-authored publications, publications in high-impact journals, and the publications' impact on science as measured by the number of forward citations. Instrumenting for location using visa status and allowing richer and poorer countries to have different impacts, we find that the negative relationship between non-U.S. location and research output is present and large for poorer countries but completely eliminated when the researcher is located in a richer country, with two one exceptions. Foreign location negatively impacts last authorship and collaboration with Americans even for those located in the richest countries. Further, allowing for heterogeneity in the treatment effect of foreign location on research output on these same countries, we find that the negative effect on publications of being abroad is largest for those with the lowest estimated propensity of being abroad, those who – given their observable characteristics -- would be expected to remain in the US.

---

* Boston University, ** Boston University and NBER. Email: skahn@bu.edu, mmacgarv@bu.edu. We thank the Institute for International Education and Jerry Murphy for the Fulbright directories. We would like to thank Olesya Baker, Gina Brandeis, Maria Burtseva, Gerardo Gomez-Ruano, Dana Schulsinger, and Vanessa Wong for excellent research assistance. We thank Richard Freeman, Rodrigo Canales, and participants at the Spring 2008 meetings of the NBER Productivity group and the MIT BPS Jr Faculty Mini-conference for helpful comments. 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 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 who received U.S. doctorates but 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 from our sample of 446 foreigners who received U.S. science Ph.D.'s during the 1990's and early 2000's summarized in Figures 1 and 2 suggests the answer to this question is a resounding yes. Compared to those located outside the U.S., the U.S.-located U.S.-educated foreign Ph.D. scientists in our sample produce more knowledge each year, as measured by their average journal publications, and this knowledge is diffused more broadly, as measured by forward (i.e. later) citations to these articles. Furthermore, U.S.-educated Ph.D. scientists located abroad conduct research that 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 endogeneity of their location choices. Those scientists located in the U.S. and those outside are likely to differ widely in their inherent research ability and proclivity. Better researchers may be more likely to receive U.S. job offer. and/or those most interested in research may be more likely to remain in the U.S.

This paper makes use of a new dataset that follows the post-Ph.D. careers of foreign scientists who came to the U.S. for their doctorate. It is unique in being the only data set of which we are aware that tracks the career progression of individual U.S.-trained Ph.D. scientists, whether they leave the U.S. or not.[1]

Our sample has been carefully crafted to exploit exogenous variation in post-Ph.D. location induced by visa status. It does this by comparing 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 and prestige of publications and the individuals' contribution to these publications as measured by first and last authorship. We measure these publications' impact on science by their number of forward citations, the scientists' connection to cutting-edge science by the median lag of publications' backward citations (i.e. articles cited in the publication), and their links to the American scientific community by co-authorship with Ph.D. advisors and others in the U.S. In all regressions, we control for scientists' pre-graduation research output, which we believe to be a good proxy for inherent research potential.

Instrumenting for location using visa status and allowing richer and poorer countries to have different impacts, we find that the negative relationship between non-U.S. location and research output is present and large for poorer countries but completely eliminated when the researcher is located in a richer country, with two exceptions. Even for those located in the

[1] One can obtain information on foreign-born scienstists who remain in the U.S. from the NSF's SESTAT database. Also, Michael G. Finn's research provides valuable information on the stay rates of Ph.D.s. of foreign origin.

richest countries, foreign location negatively impacts both last authorship and collaboration with Americans. Further, allowing for heterogeneity in the treatment effect of foreign location on research output on these same countries, we find that the negative effect on publications of being abroad is largest for those with the lowest estimated propensity of being abroad, those who – given their observable characteristics -- would be expected to remain in the US.

Why Location May be Important

Both place and proximity matter for research productivity in science. In those geographic areas in the U.S. (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 both 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 problem, 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 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 respectively. (Shanghai Jiao Tong University 2008).[2] 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

[2]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.

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 might result because leaving the U.S. may involve diminished collaboration with Ph.D. advisors and other contacts made during graduate school that is not replaced by local collaborators with similar knowledge and ability. In addition, Ph.D. recipients who obtain university jobs in their home countries may face all the other 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 had 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 beginning 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 both increase their non-U.S. employment of research scientists and off-shore some 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 wish to isolate the impact of being in the U.S. on research productivity, impact and collaboration and hence knowledge acquisition and diffusion. As noted in the introduction, however, comparisons of U.S. and foreign scientists' research output will inevitably be plagued by selection bias, as scientists' locations are likely to be influenced by unobserved characteristics correlated with productivity. 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.

The strategy we use to identify the separate effect of location on productivity, collaboration etc. is to identify pairs of foreign-born U.S.-Ph.D. recipients in science from the same department in the same university graduating during the same period university (and, whenever possible, with the same advisor) – one of whom has a J-1 visa and is required by law to leave the United States for at least two years after finishing his/her doctorate, and one of whom faced no such restrictions. Many U.S. doctoral recipients with J-1 visas entered the U.S. through the Foreign Fulbright Fellowship program; and we use Fulbright award information to identify these students.

The Fulbright Program for Foreign Students is sponsored by the U.S. Department of State and administered by bi-national Fulbright Commissions/Foundations or U.S. Embassies. Potential Foreign Fulbright Fellowship recipients apply to and are selected by committees in their home countries. Selection criteria for the Fellowships are determined by each country's committee. Once students are selected, the Institute for International Education (IIE) works with the student and national Fulbright commission to facilitate their academic placement at U.S. universities.

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) our Fulbright sample are similar to our control group with respect to potential research productivity and proclivity at graduation.

Do Fulbright Fellows actually leave the U.S. as the conditions stipulate? The requirement to leave the country after the completion of studies 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] These categories 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) on-the-job immediately following the

[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).

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 years, before having to leave the country. Moreover, after they spend two years abroad, they can apply for a work visa and return to the U.S. The two years outside the U.S. need not be 730 consecutive days, but could be a combination of summers and/or semester-long visits while in a post-doc or in OPT. In fact, a substantial number of Ph.D. recipients with Fulbrights 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. Only 12.1 percent of our Fulbright sample appeared to have remained in the U.S. continuously and thus not have fulfilled their foreign residency requirement, although even they very well could have been fulfilled the requirements in short segments. For the other 87.9 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 41.3 percent of our control group on non-Fulbrights. As Table 1 indicates, we observe our sample of 223 Fulbright scholars for a total of 2,053 person-years, and 77.9% of these years are spent outside the U.S. In contrast, the 223 controls – also foreign-born who had completed college in their home countries – spent only 33.3% of their 2,116 observed person-years outside the U.S. This stay rate of approximately 66.6% for control students is very similar to the average stay rate estimated in a much larger sample by Finn (2007), who found that of students receiving their doctorates in 1998 (close to the average year of Ph.D. in our sample), 67% were observed in the U.S. in 2003.

While the Fulbright instrument is strongly correlated with the endogenous variable location, we still must face the second challenge of establishing that the Fulbright group and the control group are similar to the Fulbright group in terms of *potential* research productivity and proclivity at graduation. Our matching of each Fulbright with a control by university,

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

department, time period and, when possible advisor, was done precisely in order to create two groups with identical research potential. Foreign students have told us that while the foreigners with the highest observable ability obtain funding from sources that do not have location restrictions, these dimensions of ability are reflected in the quality of the university that they attend.

Nevertheless, we must consider the possibility that the Fulbright and control, although so carefully matched, may have different research productivity/proclivity. The direction of bias is not obvious, however: arguments can be made in both directions.[5] It is tempting to wave our hands and dismiss these concerns on the grounds that the negative bias is likely to counteract the positive. However, given the possibility of bias, we include all exogenous control variables in both stages of our IV analysis. Importantly, we control for students' research output while in grad school, which we believe to be a good proxy for inherent ability. We have also compared our results of this matched sample with a different matching strategy for our sample based on estimated propensity scores of being abroad.

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

[5] For instance, foreigners with high preferences for research might want to remain in the US and not accept funding that restricts future funding, but preferences may be unobservable to university admissions. On the other hand, in some countries, primarily developing countries, the Fulbright committee chooses the university a Fulbright recipients attend. We spoke to several foreign graduate students in the US who believed that Fulbrights were often placed in lesser universities than they would otherwise have gone to.

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 identified 223 Fulbrights whose post-Ph.D. locations could be found via web searches.

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.[6] 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 A 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 checked the "acknowledgements" section of potential control students' dissertations and their CV's for

[6] 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.

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 8 Fulbright Scholars from Colombia but none from Chile, 5 from Thailand but none from Indonesia, China or India. We thus 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 location 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 final data set includes 223 Fulbright Ph.D.s and 223 control Ph.Ds. We include data for each year that each Ph.D. is observed from their Ph.D. graduation year to 2007. The key right hand side variable is the researcher's location. Because of the time between when research is performed and when it is published, we have lagged this variable, experimenting with a one and a two year lag. Results are qualitatively similar. We display results for one year lagged non-US location, *LAGFORLOC,* choosing it over two-year lags both because a one-year lag corresponds with scientists' impressions of the average publication lag (for the established journals in *Web of Science*) and because this gives us more observations than we would have with two year lags and therefore more significant results. This leaves us with 4,169 observations.[7]

The match between treated (i.e. Fulbright) and control students was made with the goal of choosing controls that are as similar as possible along the characteristics relevant to our study. The criteria we used for matching were based on our priors about the characteristics that are most relevant for future research output (institution, advisor/field, date of graduation, and where possible region of origin). However, it is possible that, due to the inherent difficulties of finding controls that are identical to the Fulbrights along every dimension except with regard to visa status, there may be differences between controls and Fulbrights that introduce bias. For example, our matched pairs are very often not from the same region of origin, because we felt it

[7] We also experimented with simultaneous lags of different lengths. Because FORLOC is highly serially correlated for each person, when more lags were included, their coefficients were typically insignificant.

was more important to match on advisor or field than on region of origin. In an ideal setting, we would have a large dataset with rich data on the characteristics of U.S.-trained scientists inside and outside of the U.S. before and after graduation (matched to publication data) with J-1 visas and without. With such a dataset, we could experiment with matching on different characteristics or use a methodology like propensity score matching to select (possibly multiple) controls. Such a dataset does not exist, which is why we have painstakingly hand-collected the data we have described here matched Fulbrights to controls on the characteristics we view as most important for establishing similarity between controls and Fulbrights. However, to investigate the possibility that our dataset is biased relative to one that matches on other characteristics, below we compare results using our matched sample to samples matched with nearest neighbors in field and region or matched on propensity score.

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:

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.

Last-authored publication counts: the number of articles on which the scientist is listed as the last contributing author, by publication year. Since typically, the Principal Investigator (PI)_on a research grant will be the final author listed, this variable is an indicator of the author's ability to secure research funding.

Publications in high-impact journals: We classify a journal as "high-impact" if it is in the top ten percent of journals ranked by ISI's impact factors (as of 2007) or if it is in the top ten percent of journals ranked by total citations received. A list of the journals meeting these criteria is available upon request.

Forward citation counts: The total number of citations received by articles published that year as of 2008. Publications have different impacts on their field. This variable in a sense weights publications by the number of their citations. Forward citations are an indicator of an article's impact, and we compute this for the total articles published, for first-authored publications, for last-authored publications, and for publications in high-impact journals.

Due to the extreme skewness of their distributions, publication counts are winsorized at the 99th percentile and and citation counts at the 95th percentile. Results obtained using raw publication and citation counts were qualitatively similar to the ones we report here.

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.[8] Analyses of citation lags are limited to those person-years when one or more publications are observed.

Share of publications co-authored with the scientist's thesis advisor: The percentage of publications that list the scientist's thesis advisor as one of its co-authors. 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. Analyses of this and the following co-authorship variable forward citations are limited to those person-years when one or more publications are

[8] 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.

observed. In addition, most but not all students' advisors are listed on *Proquest*, and for students whose advisors are not listed, this variable takes on missing values.

Share of publications with at least one U.S. collaborative relationship: The percentage of publications that have at least one co-author at a U.S. institution that is not the scientist's own institution. A key question is whether recipients of U.S. doctorates who leave the U.S. continue to maintain ties with American researchers, either pre-existing ties with their advisors or their fellow students, or new ties. Because the ISI database does not link authors with institutions, instead listing any institutions associated with one or more co-author, we can compute the share of publications with at least one non-U.S. coauthor but cannot compute the share of co-authors from the U.S.

Table 3 displays the publication, citation and U.S. collaboration variables categorized both by present residence – U.S. or not – and by Fulbright status. The 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. However, the publication and citation data by Fulbright status tell a somewhat different story. Although the control scientists are much more likely to be living in the U.S., Table 3 shows that differences between controls' and Fulbrights' publications, citations and U.S. collaborations are smaller, averaging about 50% of the US/non-US spread. This observation is suggestive of what we later find in our multivariate instrumented compared to un-instrumented estimation.

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.

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. We include controls for the total number of publications and the number of first-authored publications while in a doctoral program. We extend this time period through the year after the date of completion of the doctorate, because these articles are very likely to reflect dissertation research rather than new work performed following graduation.[9] 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.

Ranking of Ph.D. institution: We include the relative ranking of the U.S. Ph.D. institution (by field) as a control for the quality of Ph.D. training. We use data as of 1995 from the National Research Council's report *Research Doctorate Programs in the United States: Continuity and Change.* Note that a lower rank signifies higher quality.

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.

Table 4 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 thesis department 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.Ds: Both variables are included in all specifications (with the exception of the Poisson I.V. specifications).[10] Table 5 lists Ph.D. year and we once again

[9] Pre-doctorate high impact and last-authored publications never had a significant effect on later research, perhaps because there were so few of them.

[10] We have also estimated the un-instrumented Poisson regressions with separate dummies for each year and for each year since Ph.D. The results are very similar whether we include these variables as continuous variables or as

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.

Gender: We obtained data on the gender of the scientist using information from web searches (e.g. photographs, the use of personal pronouns in web bios), using a web-based algorithm for identifying the probable genders of given names when no other information was available.[11]

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.

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 practically identical results.[12]

---

dummies. Due to the difficulty of getting the model to converge when many dummies are included, year and post-Ph.D. years are included as quantitative variables rather than as dummies in the estimates obtained from the GMM I.V. and for consistency, in the un-instrumented regressions as well.

[11] The gender-guessing program is found at: http://www.gpeters.com/names/baby-names.php

[12] 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 our study, we find 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.

The top half of Tables 6 and 7 contain the coefficients on foreign location variables for un-instrumented Poisson regressions, with robust standard errors clustered by student. Control variables include year, years since completion of Ph.D., pre-graduation publication variables, log of the university/department rank, gender, field and job sector. (Appendix B contains the complete regressions for the first specification in these tables. Full estimates from other specifications are available on request from the authors.) Table 6 focuses on publication counts and Table 7 on forward citations.

In the first specification, the single location variable is lagged foreign location (*LAGFORLOC*). Foreign location has a negative and statistically significant relationship on both total publications and citations, suggesting that after controlling for covariates, scientists outside the U.S. publish approximately 30% fewer articles in a given year and these articles are cited 37% less (row 1, Tables 6 and 7).[13] Limiting the analysis to first-authored publications -- i.e. those that the scientist had the major role in the research -- impacts on both publications and citations are a bit smaller (24% and 32% respectively) but still significant.

In many scientific fields, last-authorship signifies that the person was the PI who obtained the funding. Given that our sample consists of scientists in the years after they receive their Ph.D., it is not surprising that there are few last authors in our sample[14] and, correspondingly, that results for this dependent variable, although of comparable sign and magnitude of the first-authored publications and citations, is significant only at the 10% level for both publications and citations. When first and last authored articles are combined, the results resemble that for first-authored alone. Finally, *LAGFORLOC* has the largest and most significant impact on publications and citations *in high-impact journals*, at approximately 54% for publications and 50% for citations.

[13]Foreign location variable lagged two years finds qualitatively similar results. Note that the percentage effects from Poisson coefficients are calculated $\exp(\beta) - 1$.

[14]23% of our sample were *ever* last authors.

Coefficients of control variables were as expected. Of particular interest, first-authored publications in graduate school – a proxy for pre-Ph.D. research ability and proclivity -- are associated with more and more-highly cited post-Ph.D. publications. However, total publications while in graduate school typically did not have a significant effect when the dependent variables are first- or last-authored publications or citations. In other words, only work done primarily by the student himself or herself signals later research likelihood.

While in principle we would want to use a set of dummies for year and years since graduation, in practice we found it difficult to estimate the instrumented model for certain dependent variables (notably last-authored pubs and high-impact pubs) including a full set of dummies. The same can be said for a more detailed set of field dummies. We found that the results did not differ substantially depending on whether we used the more restricted or the more detailed set of control variables, and as a result we used the restricted set in all our regressions. This is likely due to the fact that our sample of controls and Fulbrights is evenly balanced in terms of year and field characteristics. It is unlikely that all foreign students have equal impacts of doing science outside the U.S. One important factor likely to affect the magnitude of any negative impact of being abroad is the real GDP per capita of the country in which the scientist is located. One might expect that those who would be most hurt by being outside the U.S. would be those who are in less developed countries, if for no other reason than that opportunities and funding for scientific in these countries is considerably less generous than in the U.S. We thus hypothesize that the effect of being outside the U.S. is heterogeneous and depends on the per capita GDP of the country the person is in. In the next section, we instrument for this variable since the income of the scientist's country is likely to be endogenous. Here, we report the uninstrumented results for comparison.

In order to examine whether the impact of being outside the U.S. depends on the wealth of the country one is in, we provide two additional specifications in Tables 6 and 7. In the first, we divide foreign countries into two groups: those in the richer continents: Europe, Canada, Australia, New Zealand Israel, and Asia (since most of those in our data base from Asia are from Japan, and India); and those in Latin America, Africa and the Middle East (excluding Israel). In the final specification, the (lagged) log GDP per capita of the foreign country appears in the equation in addition to *LAGFORLOC*. From this last specification, we calculate the effect of being in a foreign location at various points of the income per capita distribution.

The results in Tables 6 and 7 for these specifications indicate the sharp difference between those in rich and poor countries. For those in poor continents, or in countries in the lower three quarters of the income distribution, being abroad has a large deleterious impact on both publications and citations. The effects for the poorer countries are much larger than those for all countries pooled. For instance, those at the 50th percentile of the income distribution have 41% fewer total publications and first-or-last authored publications than those in the US. They are even more disadvantaged with respect to citations, where they have 53% fewer publications and 46% fewer first-or-last authored publications. As before, the largest impact of being abroad at the lower 50th percentile country is on publications in high impact journal articles and citations to those articles (54% and 58% lower respectively.) The impacts on those in the poorer continents are practically identical to the impacts at the 50th percentile rank.

In contrast, the difference between being abroad in a rich country and being in the U.S. is much smaller and statistically insignificant even at the 10% significance level for total, first and last authored publications and forward citations. For publications, these impacts range from 5% (first authored publications) to 17% (total publications); for citations, they range from 9% fewer citations to last authored articles to 21% fewer total citations.

High-impact publications seem less influenced by the foreign country's income and remain significantly lower in rich countries than in the U.S. Those scientists located abroad have 54% fewer high impact publications than those who remained in the US, whether their location was at the top or the bottom of the income distribution. On the other hand, citations to high-impact articles do differ by income level, but even those in the richest 10% of countries have a highly significant 40% fewer citations to high impact articles.

Column 1 of Table 8 presents regression models of the relationship between the median backward citation lag and post-Ph.D. location using only the observations with at least one publication.[15] Scientists located outside the U.S. on average cite older literature than scientists inside the U.S. *LAGFORLOC* increases the median lag by 1.2 years, approximately a 14.5% longer median backward citation lag.[16] One interpretation of this 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.

Dividing scientists abroad by the income of the country they are in, once again it is in the poorest countries where the largest impact is seen, amounting to 20% increase in the median lag. The median citation lag at the higher GDP levels, however, is smaller in magnitude and insignificantly different in value.

Columns (2) and (3) of Table 8 investigates whether being abroad impedes collaboration with those in the U.S. Lagged foreign location has no impact at all on the share of publications co-authored with the scientist's thesis advisor. Its impact on co-authorship with anyone in the U.S. is negative and insignificant. Dividing by country GDP indicates that, unlike with other

[15] 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.

[16] Although the effect is diminished slightly 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.

measures we have investigated here, the negative impact of foreign location is largest and most significant in poor countries.

To guard against any bias which may be introduced by our matching procedure, we have also re-computed Tables 6, 7 and 8 matching on the propensity score and alternatively, matching on nearest-neighbor field and region (see Appendix C). The results for these alternate matching procedures do not differ from our original matching procedure in a systematic or substantial way, and we are thus reassured that our preferred approach is not substantially biased relative to alternative matching approaches.

Instrumented Estimation and Results

Since whether or not U.S.-educated Ph.D. recipients stay in the U.S. , go to a rich country or go to a poor county is obviously related to their research capabilities and therefore the un-instrumented results are biased estimates of the causal impact of location on research output, our most important results are the instrumented results in the bottom half of Tables 6 and 7. We use a count-data instrumental variables model developed by Mullahy (1997), a GMM model for count data with endogenous variables and a multiplicative error term. Again, standard errors are clustered by scientist.[17] Angrist (2001) has shown that the Mullahy model gives a consistent estimate of the local average treatment effect (LATE) in a model with a binary instrument, endogenous treatment variable and no covariates.

Our instruments are: (1) a Fulbright dummy representing whether or not the Ph.D. recipients was required to leave the U.S. (2) home country (lagged) GDP and (3) dummies for home continent. While Fulbrights are required to leave the U.S. for at least two years, Fulbright policy does not require that they return to their home country. Some countries do stipulate that the Fulbright-funded scientist must return home while others do not. The majority of those

[17] We used Stata's ivpois function to estimate these models, modified to allow for clustered standard errors.

Fulbrights abroad are indeed in their home country. Only an average of 13% of the Fulbrights in our sample are observed in countries other than their home country or the U.S. in any each year, and less than 17% are *ever* observed in a third country post-Ph.D. receipt. Consequently, home region and lagged income per capita of these scientists' home country are likely to be powerful instruments.[18] The F-statistic measuring the power of these instruments in predicting *LAGFORLOC* is 42.36. The F-statistic measuring the power of these instruments in predicting the (lagged) GDP of the present country is 44.61. Both of these are well above the "rule of thumb" critical value for weak instruments of 10 (Staiger and Stock(1997), Stock and Yogo (2005)). In all specifications that include more than one instrument, the instruments pass the tests of over-identifying restrictions.

We first consider the coefficient on (instrumented) *LAGFORLOC* with no income differentiation (Tables 6 and 7, bottom half first row). We expected the instrumented coefficients to be less negative than the un-instrumented ones due to foreign location being negatively correlated with ability and propensity to publish. Instead, the point estimates are all still negative and much larger in magnitude than were the coefficients in the un-instrumented results. For example, the point estimate of the impact of foreign location on total publications un-instrumented was -.359 (implying a reduction of 30%), but instrumented becomes -.784 (-54%). The largest jump is in last-authored publications, where the coefficient's magnitude changes from -.377 to -1.790. Significance levels on foreign location on average *are* lower in the instrumented results, but coefficients do remain significant. The same patterns are observed

[18] In Appendix B, we report the first stage results of the regression of the foreign location dummy on the instruments and F statistics for the other two specifications. The first column is the first stage of instrumented results in which *LAGFORLOC* is regressed on the Fulbright dummy and the other control variables.

for citations, with point estimates increasing as much as fourfold and significance levels slightly falling.[19]

Above, we argued that those who would be most hurt by being outside the U.S. would more likely be those who return to less developed countries. We thus hypothesize that the treatment effect is heterogeneous, and depends on the per capita GDP of the destination country, as a proxy for resources devoted to science at the national level.

In the un-instrumented results, the impact of being outside the US depended heavily on the income of the person's location. In the instrumented results at the bottom of Tables 6 and 7, we again first divide the impact of being outside the US by those in rich and poor regions, and in the final specification add a control for log GDP per capita of the current country. However, here, these variables are instrumented.

As with the un-instrumented results, the wealth of the country in which the scientist is located does indeed mitigate the effects of foreignness. Again, the instrumented results for these specifications with income are counter-intuitively larger than the un-instrumented ones.

Total publications by those in a foreign country at the 50th percentile of real per capita GDP are 57% lower (coefficient -.834) than total publications by those in the US, first authored publications are 62% lower (coefficient -.962) lower, and high-impact publications are 80% lower (coefficient -1.587). Citations to these publication measures – total, first authored and high impact – are all approximately 80% lower for those in a country at the 50th GDP percentile than for scientists in the U.S.

Scientists in rich countries fare much better. For those in countries at the 90th percentile of the income distribution, neither total, first-authored nor high impact publications, nor forward citations to these publications, are significantly lower than for those in the U.S. P-values for these six output categories average 64%. The point estimate for the most basic measure of

[19] We had some difficulty estimating the model with the full set of control variables when last-authored publications was the dependent variable, and in these regressions we use a modified set of controls with the pregrad publication counts in levels instead of in logs.

output -- total publications -- is only 8.0% lower in these rich foreign countries than in the U.S. The point estimates for total citations and for high-impact publications/ citations are not as small, but *are* smaller than they were for the 90th percentile in the uninstrumented results. (Compare, for instance, 18% fewer high-impact publications in the instrumented results to 54% fewer high-impact publications in the un-instrumented results.) For first-authored publications and citations, magnitudes are substantially greater in the instrumented results than they were in the uninstrumented results, but remain statistically insignificant.

Finally, the impact of foreign location on last authorship – both counts and citations – is quite different than the impact of other research measures. At the 50th percentile of the income distribution or for those in a poor region, the instrumented point estimate of being outside the U.S. is much greater than the un-instrumented one, more than 250% greater. This difference in last authorship between those in rich countries compared to the US, which had been wildly insignificant in the uninstrumented result, becomes significant at the 10% pevel (p=.0715) in the instrumented results. Citations to these articles showed similarly large increases in magnitude and similarly moved towards significance, although they did not achieve it at conventional levels.

In most scientific fields, last authorship goes to the principal investigator, the one who has obtained funding and supervises the research. The great majority of the observations in our sample come from scientists less than ten years away from doctorates. Being a PI at this career stage is relatively rare even among those Ph.D.'s who remain in the U.S., averaging 19% of the sample years and smaller among those located abroad (13%). These results suggest that while scientific scholarship proceeds in rich countries outside the U.S., funding and large labs are much more prevalent among scientists located in the U.S.

We have also rerun IV versions of the median citation lags and U.S.- publications co-authorship variables and included them in the bottom half of Table 8. Compared to the noninstrumented version in median citation lags, the point estimate of the coefficient on the foreign location dummy increases somewhat, but its standard error increases much more and as a result, foreign location becomes insignificant for the sample overall. However, dividing the impact of foreign location by country GDP, the instrumented results have the same pattern as the un-instrumented one: the citation lag is longest and significantly different from zero for the poorer countries only (here limited to those at or below the 50th percentile of countries). Again, as with publication measures, instrumenting increases rather than decreases point estimates.

In terms of collaboration with the US, there was no discernible impact of foreign location on the share of publications co-authored with the person's advisor, similar to the un-instrumented specifications. However, with IV, the impact of being abroad has a very different pattern than seen in the uninstrumented results. The negative impact of being abroad is very large, decreasing the share with US co-authors by more than 26 percentage points (compared to a mean of 49% of articles co-authored with at least one US author), and is quite similar across income levels.

To summarize the instrumented estimation, scientific research output, citations to this research and collaboration with US scientists suffers for scientists who leave the U.S. for poor countries, even after the selection bias is accounted for by instrumentation. However, agglomeration effects do not matter for researchers located in wealthy countries in terms of publications or citations, with two exceptions. Foreign location negatively impacts both last authorship and collaboration with Americans even for those located in the richest countries. First, access to funding do make it less likely that scientists outside the U.S. get funding for large labs, as indicated by the scarcity of last authorships in any country outside ths U.S. Second, collaboration seems to require propinquity. Even in the richest countries, collaboration with US scientists is rarer than for those actually living in the US.

The most perplexing aspect of our results is that, with the exclusion of our most comprehensive measures of research – total publications and citations to these publications -- coefficients in the instrumented results are larger in magnitude than in the un-instrumented ones, albeit less significantly. The estimation model we used was slightly different, with the un-instrumented poisson assuming an additive error and the instrumented GMM method assuming a multiplicative one. To check whether this accounts for the larger coefficients, we have estimated an alternative IV count model with additive errors (as developed by Windmeijer) and obtained somewhat smaller coefficients, but still ones larger than in the uninstrumented ones.

One possibility for the larger coefficients is that these results are biased by inherent differences between Fulbrights and controls. As discussed previously, there is still a possibility that Fulbright recipients are dissimilar in inherent research "quality" to their controls, despite being from the same field and institution and despite controlling for pre-graduation publications. We investigated this by comparing the two groups on the only directly observable measure of research quality in our data, publications before Ph.D. receipt (including the Ph.D. year itself). Here, research output is once again measured not just in terms of total articles published while in graduate school, but also first-authored articles and high-impact articles. Regressing any of these publication variables on Fulbright status, along with field and home region controls (again using Poisson estimation)[20], the impact of Fulbright is not distinguishable from zero, with t-statistics near or below 1. Point estimates *are* negative, however, which suggests that if our instrumented results remain biased, they will be biased away from zero so that we overstate the negative causal effect of foreign location. As a result, we are cautious in our interpretation of the magnitude of the negatives effects we observe, despite the fact that we control for observable quality via pre-graduation publications. However, this analysis lays to rest any concerns that the absence of any negative impact of non-U.S. location in wealthier countries (for anything besides

[20] Results available from authors.

last authorship) is due to a *positive* bias due to higher research quality among Fulbrights than among the controls.

Heterogeneous effects of Foreign Location

These results constitute evidence of heterogeneous effects of being abroad on research output, insofar as these effects vary by countries' per capita income. If the effects of being abroad differ along an observable dimension of variation across scientists, there is also likely to be *unobservable* heterogeneity in the impact of foreign location. As the literature on LATE (local average treatment effects) emphasizes, in I.V. the coefficient on *LAGFORLOC* (the treatment) can only pick up the impact of foreign location on those Fulbrights who would otherwise stay in the U.S. (the treated) as compared to the impact of foreign location on other Fulbrights and on the controls. If the unobservable attractiveness of U.S. location is not uniform across individuals, the instrumented results of Tables 6 and 7 may not necessarily reflect the true magnitude of being abroad on those most affected by foreign location. While our Fulbright instrument may give us an unbiased estimate of the average effect of foreign location based on those whose behavior changes as a result of the Fulbright's location restrictions (LATE), since the effect may differ across individuals, it would be more informative if we could distinguish impacts on heterogenous individuals.

This problem is described by the literature on heterogeneous treatment effects (Bjorklund and Moffitt 1987, Imbens and Angrist 1994, Heckman and Vytlacil 1999, etc.). Heckman, Urzua and Vytlacil (2006) and Moffitt (2008) develop methods in which the heterogeneous individual-specific benefit from the treatment is modeled as a function of the probability of participating in the treatment (the propensity score).

For our problem, we use the approach described in Moffitt (2008) to investigate heterogeneous treatment effects of foreign location on research output. Specifically, let y_i

represent the research outcome for each scientist in our sample, X_i the exogenous explanatory variables, Z_i the Fulbright dummy instrument, D the dummy *LAGFORLOC*, and P equal the predicted propensity score. A two stage method to identify heterogeneous impacts of foreign location on research output is to estimate:[21]

$$D = P(Z_i, X_i) + \mu_i$$

$$y_i = \alpha + X_i \gamma + X_i P_i \delta + \beta_1 P_i + \beta_2 P_i^2 + \beta_3 P_i^3 + \varepsilon_i$$

The estimated marginal treatment effect (MTE) measuring the heterogeneous effect of being in a foreign location is the derivative of this equation:

$$dy_i/dP = X_i \delta + \beta_1 + 2 \beta_2 P_i + 3 \beta_3 P_i^2$$

The results of this estimation are summarized by Figures 3 and 4, where the effect of being in a foreign country (the MTE) is graphed as a function of the propensity score for publications and citations respectively. The propensity score estimates are dominated by two factors besides Fulbright status: the per capita income of the home country (which increases the propensity to be abroad) and pre-graduation research output (which decreases it).

As noted by Moffitt (2008), the measured effect of the treatment is only an accurate representation in the range of propensity scores where there is a reasonable fraction of the observations and, simultaneously, where the instruments have some power, which in our case means in the range where we observe a reasonable number of both Fulbrights and controls. At low levels of propensity score, there are both very few Fulbrights and very few foreign-born Ph.D. scientists observed outside of the U.S., while at the high levels of propensity score, there are few controls and few scientists observed in the U.S.[22] Figures 3 and 4 are limited to the ranges of propensity scores within the interquartile ranges of both the Fulbrights and controls in

[21] Moffitt (2008) uses several different methods to estimate nonlinear functions of P_i. We use the simplest method using polynomials because it uses up the least degrees of freedom, a great benefit in our small sample.

[22] Only 10% of the observations with LAGFORLOC=1 have a propensity score of less than 0.35 and similarly, only 10% of those with LAGFORLOC=0 have a propensity score above 0.75. Only 2% of Fulbrights have a propensity score below 0.4 and only 16% of controls have a propensity score above 0.75.

our dataset where the estimated effects of being abroad are the most informative, which translates to propensity scores between 40% and 70%.[23]

In Figure 3, the impact of being abroad on publications rises monotonically, remaining below zero in the ranges of most observations. On the left of the distribution lie those who *ceteris paribus* would have been more likely to remain in the US, for instance because they are the most likely to do research (as indicated by their pre-Ph.D. publications) and come from relatively richer countries. In this range, the Fulbrights' locational behavior is most affected by their Fulbright status and their research output most vulnerable to obstacles they face in conducting research.

The negative impact of being abroad is approximately zero at the upper ranges of this graph. These Ph.D. recipients would have been likely to leave the U.S. anyway. They are unlikely to publish even if they remain in the U.S. and therefore have low opportunity costs of being abroad.

The impact of being abroad on citation measures is shown in Figure 4. Here too, the negative impact of foreign location falls as the propensity to be abroad rises. The far left of this graph for citations to first authored and high impact articles is the only exception to this pattern, since the greatest negative impact is observed at propensities to be abroad between 45% and 50%. This result seems anomalous primarily because the citations have a different pattern than do the publications themselves. Insofar as there are very few Fulbrights below a propensity score of 0.45 in our data, it is likely to be of little import.

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

[23] We separately estimated the first-stage F-statistic in each decile of the propensity score, and found that it was below the Stock-Yogo rule-of-thumb value below the 30th percentile and above the 70th percentile of the propensity score.

in the U.S. A naïve comparison of post-Ph.D. publication records for a sample of such students suggests 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 causal negative effect of location on research output is restricted to scientists located in poorer countries. Those in richer countries are just as likely to publish and be cited as those remaining in the U.S.

We did, however, identify two differences between those in the US and those in other rich countries. first, a scarcity of last authorship among those abroad, presumably due to the lower inability to obtain funding for large research labs, and less collaboration with American scientists. Accounting for heterogeneity in terms of the propensity to be abroad – thus simultaneously allowing these scientists to differ not just on the dimension of country's wealth but also on the dimension of pre-graduation quality measures, field, etc. – the negative effect on publications of being abroad is largest for those with the lowest propensity of being abroad, those who, for visa or other reasons, left the U.S. despite observable characteristics suggesting that they would remain in the US.

Overall, our findings suggest that research is carried on as much and as successfully in the countries in the top income distribution as it is in the US. This finding is fairly surprising in light of the high degree of concentration of top scientists at U.S. universities. It may reflect the dual factors of increasing numbers of research centers around the world in both the academic and private sectors and increasingly easy international collaboration and communication via the internet. However, our findings suggest that propinquity does favor collaboration – as evidenced by the fewer number of American collaborations – and that the funding Americans receive does seem to allow researchers in the US more ability to manage large labs and to conduct science that needs such labs. Given these intriguing results, in future stages of our research we intend to study the effects of location on international collaboration networks in science.

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.

Finn, M. G. (2007), "Stay Rates of Foreign Doctorate Recipients from U.S. Universities, 2005", Working paper, Oak Ridge Institute for Science and Education.

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.

Goldberger, Marvin L., Brendan A. Maher, and Pamela Ebert Flattau, Editors; (1995), *Research Doctorate Programs in the United States: Continuity and Change*, Committee for the Study of Research-Doctorate Programs in the United States, National Research Council. NATIONAL RESEARCH COUNCIL, NATIONAL ACADEMY OF SCIENCES, Washington, D.C. 1995

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

Stephan, Paula (2007),"Wrapping It Up in a Person: The Location Decision of New PhDs Going to Industry." *Innovation Policy and the Economy,* MIT Press, edited by Adam Jaffe, Josh Lerner and Scott Stern.

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: Publications per year, by location

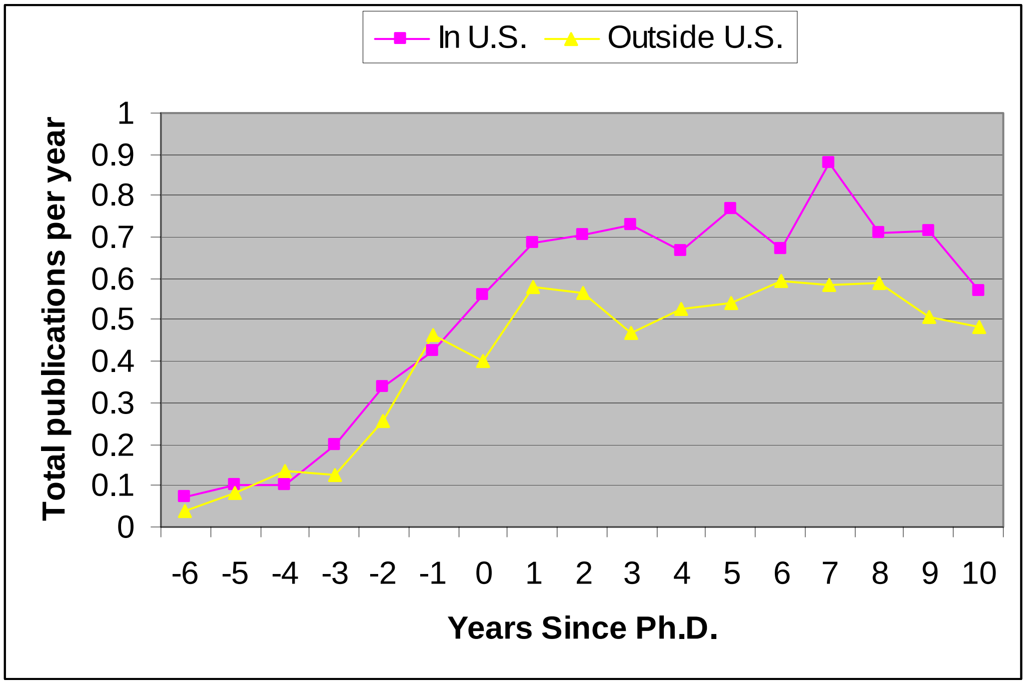
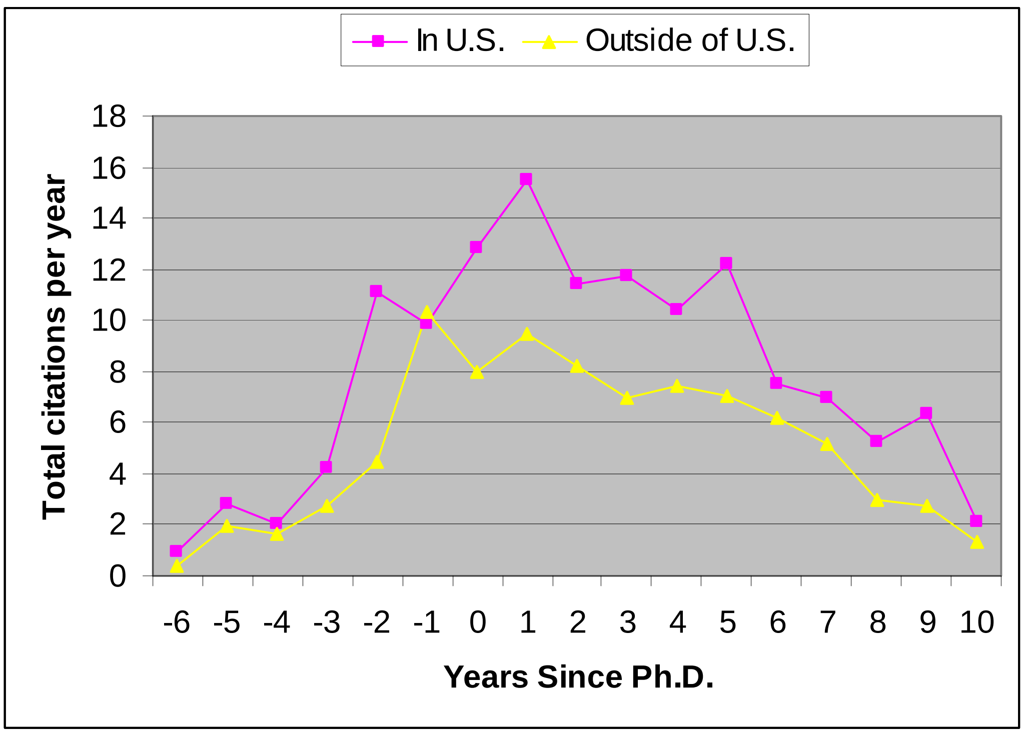


Figure 2: Forward citations per year by year of publication and location



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

Figure 3
The Impact of Being Abroad on Publications, as a Function of Propensity to be Abroad

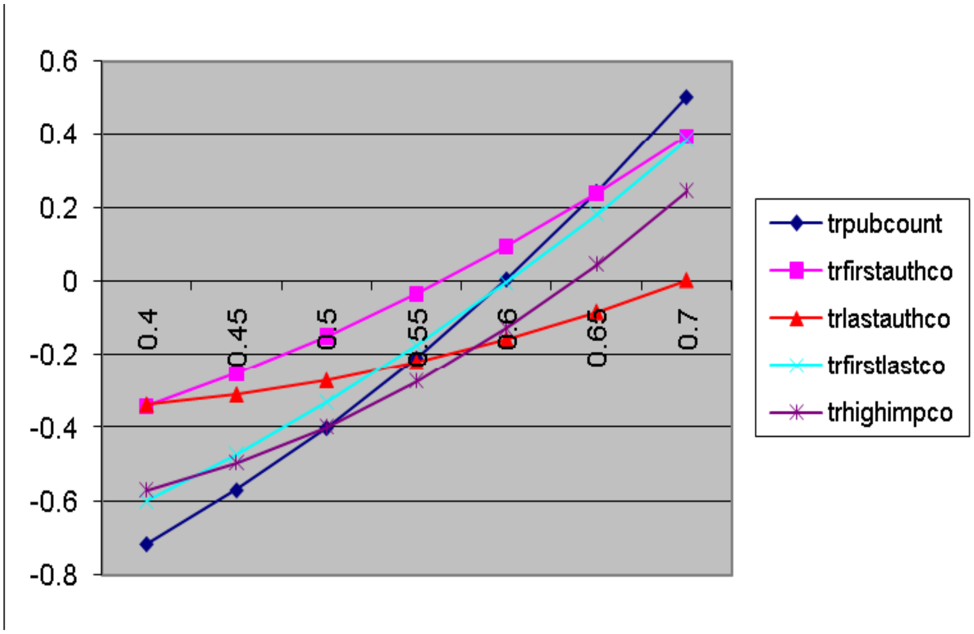


Figure 4
The Impact of Being Abroad on Citations, as a Function of Propensity to be Abroad

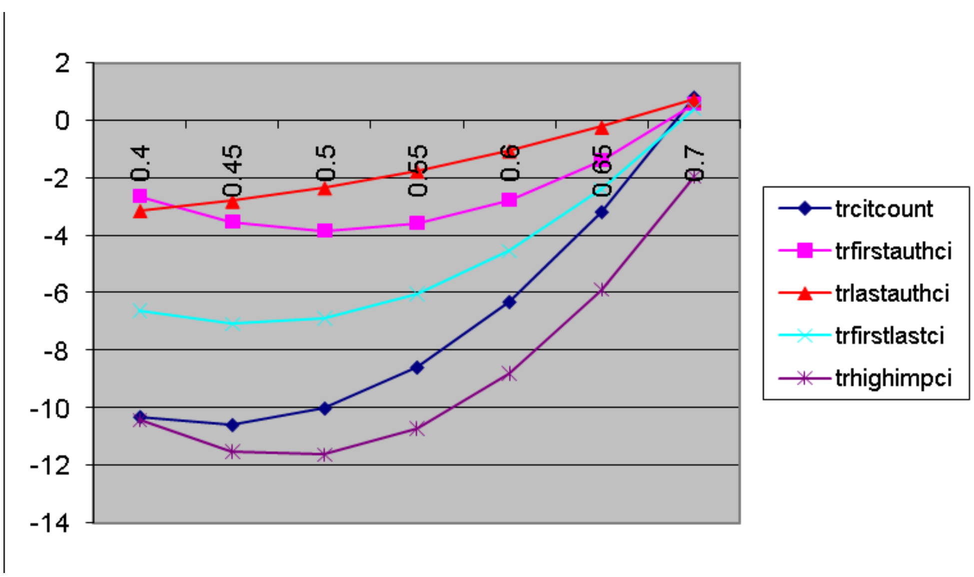


Table 1: Share of post-Ph.D. years spent outside U.S.

	Mean	Std. Dev.	Count
Overall			
223 control scientists	0.333	0.472	2116
223 Fulbright scientists	0.779	0.415	2053

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

Region/Country of origin	Control	Fulbright	Total	Region/Country of origin	Control	Fulbright	Total
Asia	**79**	**6**	**85**	**Latin America**	**35**	**118**	**153**
Bangladesh	2	0	2	Argentina	3	4	7
China	18	0	18	Bolivia	0	1	1
India	23	0	23	Brazil	7	0	7
Indonesia	4	0	4	Chile	3	0	3
Japan	5	0	5	Colombia	5	8	13
Korea	8	0	8	Costa Rica	0	3	3
Malaysia	1	0	1	Ecuador	1	0	1
Pakistan	1	0	1	Guatemala	1	2	3
Philippines	3	1	4	Haiti	0	1	1
Singapore	1	0	1	Mexico	9	94	103
Sri Lanka	1	0	1	Panama	1	1	2
Taiwan	7	0	7	Peru	2	2	4
Thailand	5	5	10	Trinidad & Tobago	1	1	2
Europe/Canada/Aust/NZ	**79**	**69**	**148**	Venezuela	2	1	3
Australia	0	4	4	**Middle East/Africa**	**30**	**30**	**60**
Austria	2	1	3	Armenia	1	0	1
Bulgaria	1	0	1	Botswana	0	1	1
Canada	8	0	8	Cote D'Ivoire	1	3	4
Croatia	1	1	2	Cyprus	1	0	1
Czech Republic	3	1	4	Egypt	3	0	3
Denmark	2	1	3	Ethiopia	2	2	4
Finland	2	3	5	Ghana	0	2	2
France	1	0	1	Iran	1	0	1
Germany	9	0	9	Iraq	1	0	1
Greece	4	7	11	Israel	3	2	5
Hungary	2	1	3	Jordan	1	0	1
Iceland	2	7	9	Kenya	0	2	2
Ireland	2	1	3	Lesotho	0	1	1
Italy	3	3	6	Malawi	1	1	2
Lithuania	0	1	1	Morocco	0	2	2
Macedonia	1	0	1	Nigeria	2	0	2
Netherlands	3	5	8	Solomon Islands	0	1	1
Norway	1	4	5	South Africa	0	7	7
Poland	1	1	2	Swaziland	1	0	1
Portugal	1	12	13	Tanzania	1	1	2
Romania	4	1	5	Togo	0	2	2
Russia	8	0	8	Turkey	9	1	10
Spain	5	7	12	Uganda	1	2	3
Sweden	1	3	4	Zimbabwe	1	0	1
Switzerland	2	1	3				
UK	2	4	6				
Ukraine	5	0	5				
Yugoslavia	3	0	3				

Table 3: Publications, Citations and U.S. Collaboration, by post-Ph.D. location and Fulbright status

Variable	Obs.	Mean	Std. Dev.	Min	Max
Scientists located in USA					
Total publications count	1864	0.8798	1.5461	0	7
Total fwd citations count	1864	10.7677	28.7773	0	153
First-authored publications	1864	0.3514	0.7076	0	3
First-authored fwd citations	1864	4.0606	13.1498	0	77
Last-authored publications	1864	0.1921	0.5830	0	3
Last-authored fwd citations	1864	1.1856	5.3192	0	35
Fist-or-last authored publications	1864	0.5016	0.9524	0	4
First-or-last authored fwd citations	1864	5.3997	16.3136	0	91
High-impact publications	1864	0.5075	1.1043	0	5
Fwd citations to high-impact publications	1864	8.3777	25.5719	0	140
Median citation lag	698	7.2264	3.8890	0	34.5
Share of publications with advisor	652	0.3073	0.4401	0	1
Share of publications with U.S. co-authors	723	0.5421	0.4413	0	1
Scientists located outside USA					
Total publications count	2305	0.5132	1.1063	0	7
Total fwd citations count	2305	3.9046	15.3748	0	153
First-authored publications	2305	0.2174	0.5591	0	3
First-authored fwd citations	2305	1.6594	7.7266	0	77
Last-authored publications	2305	0.1293	0.4485	0	3
Last-authored fwd citations	2305	0.7076	4.0346	0	35
Fist-or-last authored publications	2305	0.3106	0.7210	0	4
First-or-last authored fwd citations	2305	2.3640	9.9063	0	91
High-impact publications	2305	0.1796	0.6193	0	5
Fwd citations to high-impact publications	2305	2.1844	11.6254	0	140
Median citation lag	603	8.7670	4.5030	1	33
Share of publications with advisor	578	0.2087	0.3853	0	1
Share of publications with U.S. co-authors	620	0.4327	0.4560	0	1
Control Scientists					
Total publications count	2249	0.7443	1.3834	0	7
Total fwd citations count	2249	8.0614	24.2516	0	153
First-authored publications	2249	0.3006	0.6578	0	3
First-authored fwd citations	2249	3.0707	11.3838	0	77
Last-authored publications	2249	0.1859	0.5605	0	3
Last-authored fwd citations	2249	1.0765	4.9061	0	35
Fist-or-last authored publications	2249	0.4429	0.8770	0	4
First-or-last authored fwd citations	2249	4.2797	14.3309	0	91
High-impact publications	2249	0.3766	0.9290	0	5
Fwd citations to high-impact publications	2249	5.9022	21.1050	0	140
Median citation lag	769	7.7815	4.2140	1	34.5
Share of publications with advisor	715	0.2912	0.4324	0	1
Share of publications with U.S. co-authors	868	0.5232	0.4503	0	1
Fulbright Scientists					
Total publications count	2242	0.5361	1.1971	0	7
Total fwd citations count	2242	5.1035	19.2044	0	153
First-authored publications	2242	0.2297	0.5788	0	3
First-authored fwd citations	2242	2.1280	8.9947	0	77
Last-authored publications	2242	0.1088	0.4222	0	3
Last-authored fwd citations	2242	0.6463	4.0408	0	35
Fist-or-last authored publications	2242	0.3091	0.7471	0	4
First-or-last authored fwd citations	2242	2.7645	11.1095	0	91
High-impact publications	2242	0.2413	0.7814	0	5
Fwd citations to high-impact publications	2242	3.4286	16.0208	0	140
Median citation lag	573	8.3307	4.4679	0	33
Share of publications with advisor	555	0.2441	0.4114	0	1
Share of publications with U.S. co-authors	636	0.4404	0.4518	0	1

Table 4: Number of Fulbrights and Controls, by Field of Study

Field of Study	Controls	Fulbrights	Total
Biological/Biomedical/Health Sciences	33	36	69
Engineering/ Computer & Information Sciences	74	70	144
Ocean/Marine/Environmental Science / Geological & Earth Sciences/Agricultural	76	76	152
Mathematics/ Chemistry/ Physics	40	41	81
Total	223	223	446

Table 5: 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	7	5	12
1994	11	15	26
1995	12	21	33
1996	27	24	51
1997	38	30	68
1998	39	39	78
1999	32	32	64
2000	23	18	41
2001	12	21	33
2002	8	10	18
2003	7	6	13
2004	2	1	3
2005	2	1	3
Total	223	223	446

Table 6: Regression Results on Publications

	(1) Total publications count	(2) First-authored publications	(3) Last-authored publications	(4) First-or last-authored publications	(5) High-impact publications
Impact of Being Abroad, Un-instrumented (Poisson ML)					
Average impact of being abroad	-0.359	-0.270	-0.277	-0.313	-0.776
	(0.110)***	(0.107)**	(0.164)*	(0.107)***	(0.159)***
Impact of being abroad in a poor country	-0.583	-0.474	-0.459	-0.531	-0.899
	(0.125)***	(0.127)***	(0.231)**	(0.131)***	(0.181)***
Impact of being abroad in a rich country	-0.178	-0.100	-0.156	-0.140	-0.670
	(0.139)	(0.134)	(0.177)	(0.128)	(0.213)***
Foreign location in a country with GDP at the:					
25% percentile	-0.741***	-0.739***	-0.690*	-0.765***	-0.776***
	(p = 0.0050)	(p = 0.0012)	(p = 0.0868)	(p = 0.0010)	(p = 0.0000)
50% percentile	-0.536***	-0.483***	-0.478*	-0.522***	-0.776***
	(p = 0.0008)	(p = 0.0003)	(p = 0.0538)	(p = 0.0003)	(0.0016)
75% percentile	-0.352***	-0.252**	-0.287*	-0.303***	-0.776***
	(p = 0.0014)	(p = 0.0176)	(p = 0.0811)	(p = 0.0043)	(p = 0.0000)
90% percentile	-0.188	-0.047	-0.117	-0.109	-0.775***
	(p = 0.1999)	(p = 0.7408)	(p = 0.5644)	(p = 0.4455)	(p = 0.0000)
Impact of Being Abroad, Instrumented (GMM Mullahy model)					
Average impact of being abroad	-0.784**	-0.844*	-1.837**	-1.122**	-2.438**
	(0.395)	(0.470)	(0.809)	(0.445)	(1.251)
Impact of being abroad in a poor country	-0.702	-0.738	-1.138	-0.985	-1.968
	(0.294)**	(0.393)*	(0.533)	(0.356)***	(0.791)**
Impact of being abroad in a rich country	-0.001	-0.302	-0.961	-0.554	-1.661
	(0.515)	(0.589)	(0..622)	(0.525)	(1.113)
Foreign location in a country with GDP at the:					
25% percentile	-1.372***	-1.269***	-1.895***	-1.506***	-2.168***
	(p=0.0013)	(p=0.0094)	(p=0.0086)	(p=0.0005)	(p=0.0050)
50% percentile	-0.834***	-0.962**	-1.702***	-1.102***	-1.587***
	(p=0.0086)	(p=0.0165)	(p=0.0079)	(p=0.0023)	(p=0.0003)
75% percentile	-0.350	-0.686	-1.529**	-0.738*	-0.853
	(p=0.3324)	(p=0.1400)	(p=0.0219)	(p=0.0855)	(p=0.1124)
90% percentile	-0.080	-0.440	-1.374*	-0.415	-0.200
	(p=0.8702)	(p=0.4655)	(p=0.0715)	(p=0.4604)	(p=0.7651)

Robust standard errors clustered by scientist in parentheses: * significant at 10%; ** significant at 5%; *** significant at 1%

Table 7: Regression Results on Citations

	(1) Total citations count	(2) First-authored cites	(3) Last-authored cites	(4) First-or last-authored cites	(5) High-impact cites
Impact of Being Abroad, Un-instrumented (Poisson ML)					
Average impact of being abroad	-0.495	-0.386	-0.314	-0.391	-0.684
	(0.123)***	(0.125)***	(0.180)*	(0.131)***	(0.153)***
Impact of being abroad in poor country	-0.819	-0.584	-0.431	-0.599	-0.898
	(0.146)***	(0.152)***	(0.242)*	(0.155)***	(0.191)***
Impact of being abroad in rich country	-0.260	-0.218	-0.233	-0.224	-0.516
	(0.146)*	(0.154)	(0.199)	(0.157)	(0.193)***
Foreign location in country with GDP at the:					
25% percentile	-1.064***	-0.766***	-0.875***	-0.906***	-1.062***
	(p = 0.0000)	(p = 0.0018)	(p = 0.0301)	(p = 0.0002)	(p = 0.0019)
50% percentile	-0.757***	-0.552***	-0.586**	-0.623***	-0.857***
	(p = 0.0000)	(p = 0.0003)	(p = 0.0214)	(p = 0.0000)	(p = 0.0001)
75% percentile	-0.481***	-0.359***	-0.326*	-0.369***	-0.673***
	(p = 0.0001)	(p = 0.0037)	(p = 0.0672)	(p = 0.0039)	(p = 0.0000)
90% percentile	-0.235	-0.188	-0.094	-0.142	-0.509***
	(p = 0.1424)	(p = 0.2667)	(p = 0.6602)	(p = 0.4103)	(p = 0.0086)
Impact of Being Abroad, Instrumented (GMM Mullahy model)					
Average impact of being abroad	-1.894**	-1.443**	-1.865***	-2.663*	-1.095**
	(0.784)	(0.619)	(0.576)	(1.398)	(0.499)
Impact of being abroad in a poor country	-1.687	-1.546	-2.214	-2.600	-0.938
	(0.457)***	(0.447)***	(0.543)***	(1.147)**	(0.347)***
Impact of being abroad in a rich country	-0.687	-0.707	-1.367	-2.130	-0.266
	(0.715)	(0.583)	(1.027)	(1.276)*	(0.492)
Foreign location in a country with GDP at the:					
25% percentile	-2.403***	-2.167***	-3.658***	-2.326***	-2.157***
	(p = 0.0000)	(p = 0.0001)	(p = 0.0046)	(p = 0.0000)	(p = 0.0044)
50% percentile	-1.589***	-1.609***	-2.954**	-1.690***	-1.477***
	(p = 0.0003)	(p = 0.0007)	(p = 0.0188)	(p = 0.0001)	(p = 0.0010)
75% percentile	-0.853	-1.106**	-2.320*	-1.118**	-0.864*
	(p = 0.1124)	(p = 0.0283)	(p = 0.0725)	(p = 0.0264)	(p = 0.0961)
90% percentile	-0.200	-0.660	-1.756	-0.609	-0.320
	(p = 0.7651)	(p = 0.2811)	(p = 0.1997)	(p = 0.3220)	(p = 0.6911)

Robust standard errors clustered by scientist in parentheses: * significant at 10%; ** significant at 5%; *** significant at 1%

Table 8: Characteristics of Publications

	(1)	(2)	(3)
	Median citation lag	Share of publications with advisors	Share of publications with US co-authors
Impact of Being Abroad, Un-instrumented (OLS)			
Average impact of being abroad	1.157 **	-0.038	-0.062
	(0.384)	(0.034)	(0.038)
Impact of being abroad in poor country	1.696 ***	-0.043	-0.071
	(0.488)	(0.039)	(0.048)
Impact of being abroad in rich country	0.715	-0.034	-0.054
	(0.446)	(0.040)	(0.044)
Foreign location in country with GDP at the:			
25% percentile	2.139***	-0.030	0.035
	(p = 0.0054)	(p = 0.6297)	(p = 0.6269)
50% percentile	1.610***	-0.035	-0.018
	(p = 0.0010)	(p = 0.4017)	(p = 0.7064)
75% percentile	1.134***	-0.039	-0.065
	(p = 0.0033)	(p = 0.2579)	(p = 0.0875)
90% percentile	0.712	-0.043	-0.106**
	(p = 0.1515)	(p = 0.3266)	(p = 0.0255)
Impact of Being Abroad, Instrumented (LIML)			
Average impact of being abroad	1.869	-0.123	-0.262
	(1.204)	(0.106)	(0.105)**
Impact of being abroad in a poor country	1.767	-0.147	-0.296
	(1.178)	(0.095)	(0.103)***
Impact of being abroad in a rich country	0.936	-0.117	-0.245
	(1.577)	(0.139)	(0.152)
Foreign location in a country with GDP at the:			
25% percentile	2.846**	-0.145	-0.305**
	(p = 0.0499)	(p = 0.3382)	(p = 0.0313)
50% percentile	2.152*	-0.152	-0.303***
	(p = 0.0685)	(p = 0.1314)	(p = 0.0033)
75% percentile	1.528	-0.158	-0.302
	(p = 0.2231)	(p = 0.1321)	(p = 0.0093)
90% percentile	0.972	-0.163	-0.300*
	(p = 0.5285)	(p = 0.2663)	(p = 0.0572)

Robust standard errors clustered by scientist in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Appendix A: 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."[24] 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 US, there will inevitably be a high share of journals in this index based in the US. 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 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*,

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

Acta Pharmacalogica Sinica, etc.) were all included in the ISI index. While it is possible that ISI data is less comprehensive for articles published in non-Roman alphabets, it should be noted that only a very small number of scientists in our sample are located in Asian countries (0.36% of our observations are on scientists located in China, 0.55% in Japan, 0.87% in Korea, 1.03% in Taiwan and 1.5% in Thailand). Furthermore, these are scientists who began their careers in the United States and are thus likely to continue publishing in English-language journals.

To verify more rigorously that our sample of publications is not biased towards finding articles by U.S.-based researchers, we performed the following test. We had a research assistant collect data on the number of articles listed on scientists' C.V.s and the number of articles we obtained from ISI. We computed the share of a scientist's articles from the C.V. that were listed in the ISI database, and performed a t-test of difference in means between scientists outside the U.S. and those inside the U.S. The average share of articles found on Web of Science was 0.705 for those in the U.S. and 0.651 for those outside the U.S. We cannot reject the hypothesis of no difference in means (with a t-statistic of 0.788 and p-value of 0.433 for a two-tailed test).[25] We thus do not feel that a systematic U.S. bias is introduced by restricting our attention to journals included in 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.

[25] We also tested the hypothesis that this depended on the number of years abroad by regressing the share of articles on ISI on the number of years abroad, and the coefficient on this latter variable was -0.001 with a standard error of 0.006 (insignificantly different from zero).

Table A1: 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 A2: 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 A1

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 A2

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.* 48th *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.* 48th *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.* 48th *Midwest Symposium on Circuit and System*, vol.1, pp.187–190, August 2005.

vii

Appendix Table B1: Complete set of regression coefficients for un-instrumented regression
Dependent variables: Publication measures
Estimation method: Poisson

	(1)	(2)	(3)	(4)	(5)
	Total Publications	First-authored pubs	Last-authored pubs	First-or last-authored pubs	High-impact pubs
Lagforloc	-0.359	-0.270	-0.277	-0.313	-0.776
	(0.110)***	(0.107)**	(0.164)*	(0.107)***	(0.159)***
Year	-0.039	-0.055	-0.114	-0.065	-0.038
	(0.022)*	(0.019)***	(0.032)***	(0.020)***	(0.027)
Yrs since graduation	0.016	-0.052	0.141	0.014	0.022
	(0.022)	(0.025)**	(0.032)***	(0.022)	(0.028)
Ln(Pregrad pubs)	0.410	-0.122	0.326	0.050	0.768
	(0.162)**	(0.152)	(0.269)	(0.166)	(0.188)***
Ln(First-authored pregrad pubs)	0.517	1.002	1.028	0.917	0.377
	(0.206)**	(0.196)***	(0.376)***	(0.223)***	(0.238)
Dummy for no pregrad pubs	-0.081	-0.428	-0.228	-0.342	0.178
	(0.333)	(0.313)	(0.514)	(0.337)	(0.462)
Dummy for no first-authored pregrad pubs	0.312	0.401	0.681	0.427	0.193
	(0.316)	(0.305)	(0.535)	(0.333)	(0.403)
Ln(rank of Ph.D. institution)	-0.048	-0.061	-0.158	-0.079	-0.009
	(0.042)	(0.036)*	(0.066)**	(0.040)**	(0.058)
D(Female)	-0.093	-0.132	-0.500	-0.197	-0.015
	(0.131)	(0.120)	(0.213)**	(0.124)	(0.168)
Engineering/ Computer & Information Sciences	-0.244	-0.157	0.616	-0.007	-0.842
	(0.178)	(0.171)	(0.297)**	(0.177)	(0.275)***
Ocean/Marine/Environmental Science / Geological & Earth Sciences/Agricultural	-0.109	0.013	0.303	0.078	-0.423
	(0.151)	(0.157)	(0.268)	(0.155)	(0.194)**
Mathematics/ Chemistry/ Physics	-0.094	-0.021	0.672	0.062	-0.306
	(0.173)	(0.170)	(0.278)**	(0.175)	(0.182)*
Employed in government	-0.589	-0.419	-1.141	-0.583	-0.522
	(0.208)***	(0.211)**	(0.365)***	(0.209)***	(0.242)**
Employed in Private Sector	-0.730	-0.621	-0.900	-0.703	-0.969
	(0.172)***	(0.166)***	(0.242)***	(0.171)***	(0.227)***
D(gender unknown)	-0.018	0.591	-0.823	0.277	-0.237
	(0.159)	(0.177)***	(0.303)***	(0.172)	(0.287)
Constant	76.958	109.574	225.708	129.304	73.754
	(42.991)*	(38.021)***	(64.236)***	(40.157)***	(54.587)
Observations	4169	4169	4169	4169	4169

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Appendix Table B2: Complete set of regression coefficients for un-instrumented regression
Dependent variables: Citation measures
Estimation method: Poisson

	(1)	(2)	(3)	(4)	(5)
	Total citations count	First-authored cites	Last-authored cites	First-or last-authored cites	High-impact cites
lagforloc	-0.495	-0.386	-0.314	-0.391	-0.684
	(0.123)***	(0.125)***	(0.180)*	(0.131)***	(0.153)***
Year	-0.137	-0.134	-0.178	-0.154	-0.113
	(0.022)***	(0.020)***	(0.038)***	(0.023)***	(0.025)***
Yrs since graduation	-0.020	-0.093	0.146	-0.025	-0.026
	(0.024)	(0.028)***	(0.037)***	(0.027)	(0.026)
Ln(Pregrad pubs)	0.622	0.067	0.508	0.256	0.755
	(0.158)***	(0.161)	(0.265)*	(0.187)	(0.175)***
Ln(First-authored pregrad pubs)	0.188	0.763	0.721	0.608	0.165
	(0.201)	(0.210)***	(0.364)**	(0.246)**	(0.227)
Dummy for no pregrad pubs	-0.092	-0.216	-0.244	-0.129	-0.233
	(0.373)	(0.358)	(0.609)	(0.468)	(0.434)
Dummy for no first-authored pregrad pubs	-0.020	-0.049	0.478	-0.063	0.083
	(0.341)	(0.336)	(0.605)	(0.446)	(0.379)
Ln(rank of Ph.D. institution)	-0.090	-0.073	-0.135	-0.107	-0.038
	(0.048)*	(0.047)	(0.067)**	(0.050)**	(0.056)
D(Female)	-0.120	-0.097	-0.371	-0.171	-0.127
	(0.126)	(0.120)	(0.204)*	(0.130)	(0.152)
Engineering/ Computer & Information Sciences	-0.830	-0.628	0.452	-0.595	-1.159
	(0.212)***	(0.187)***	(0.309)	(0.226)***	(0.299)***
Ocean/Marine/Environmental Science / Geological & Earth Sciences/Agricultural	-0.201	-0.074	0.351	-0.043	-0.370
	(0.137)	(0.143)	(0.260)	(0.150)	(0.165)**
Mathematics/ Chemistry/ Physics	-0.269	-0.255	0.701	-0.171	-0.282
	(0.165)	(0.188)	(0.274)**	(0.195)	(0.177)
Employed in government	-0.388	-0.516	-1.103	-0.561	-0.422
	(0.188)**	(0.227)**	(0.427)***	(0.220)**	(0.226)*
Employed in Private Sector	-0.441	-0.516	-0.706	-0.507	-0.657
	(0.200)**	(0.183)***	(0.276)**	(0.218)**	(0.231)***
D(gender unknown)	-0.089	0.255	-1.500	-0.002	-0.175
	(0.257)	(0.252)	(0.337)***	(0.288)	(0.361)
Constant	276.242	269.781	352.978	309.474	227.162
	(43.032)***	(40.809)***	(75.054)***	(46.582)***	(49.900)***
Observations	4169	4169	4169	4169	4169

Robust standard errors, clustered by scientist, in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Appendix Table B3: First-stage regressions of endogenous variables on control variables and instruments (Estimation method: LIML)

Dependent variable:	Lagforloc		Lagforloc X GDP	Rich Foreign Region	Poor Foreign Region
Fulbright Dummy	0.422***	0.361***	3.362***	0.220***	0.141***
	(0.039)	(0.043)	(0.409)	(0.036)	(0.033)
Latin American country of origin		0.150***	1.134***	-0.464***	0.614***
		(0.048)	(0.443)	(0.037)	(0.042)
Middle East/African country of origin		0.166***	1.350***	-0.229***	0.395***
		(0.071)	(0.632)	(0.053)	(0.066)
Home country's lagged real GDP per capita		0.106***	1.243***	0.098***	0.007
		(0.027)	(0.252)	(0.025)	(0.018)
Year	-0.025***	-0.030***	-0.272	-0.023***	-0.007
	(0.008)	(0.008)	(0.071)	(0.006)	(0.005)
Yrs since graduation	0.039***	0.039***	0.371	0.024***	0.015***
	(0.008)	(0.008)	(0.071)	(0.006)	(0.005)
Ln(Pregrad pubs)	-0.110	-0.126**	-1.210	-0.073	-0.053
	(0.061)	(0.059)	(0.574)	(0.057)	(0.045)
Ln(First-authored pregrad pubs)	0.0489	0.077	0.785	0.040	0.037
	(0.095)	(0.095)	(0.917)	(0.087)	(0.066)
Dummy for no pregrad pubs	0.0908	0.062	0.451	-0.007	0.069
	(0.087)	(0.082)	(0.784)	(0.081)	(0.071)
Dummy for no first-authored pregrad pubs	-0.0954	-0.058	-0.388	0.011	-0.069
	(0.111)	(0.107)	(1.030)	(0.099)	(0.082)
Ln(rank of Ph.D. institution)	-0.0169	-0.012	-0.157	-0.039***	0.026**
	(0.014)	(0.014)	(0.134)	(0.013)	(0.010)
D(Female)	0.0538	0.048	0.467	0.027	0.021
	(0.046)	(0.044)	(0.425)	(0.040)	(0.031)
Engineering/ Computer & Information Sciences	0.0656	0.081	0.771	0.045	0.036
	(0.064)	(0.062)	(0.582)	(0.051)	(0.045)
Ocean/Marine/Environmental Science / Geological & Earth Sciences/Agricultural	0.1968***	0.188***	1.720	0.120***	0.068
	(0.058)	(0.057)	(0.528)	(0.045)	(0.042)
Mathematics/ Chemistry/ Physics	0.0808	0.078	0.720	0.038	0.040
	(0.069)	(0.068)	(0.654)	(0.057)	(0.046)
Employed in government	0.0463	0.038	0.482	0.079	-0.041
	(0.064)	(0.064)	(0.625)	(0.064)	(0.041)
Employed in Private Sector	-0.0869	-0.064	-0.465	-0.008	-0.057*
	(0.048)	(0.048)	(0.463)	(0.046)	(0.032)
D(gender unknown)	-0.1178	-0.028	-0.607	-0.116	0.088
	(0.128)	(0.127)	(1.165)	(0.132)	(0.075)
Constant	49.3189***	58.778***	533.630	44.489***	14.289
	(16.241)	(15.238)	(142.121)	(12.926)	(10.849)
First-stage F-statistic	119.3	42.36	44.610	58.460	131.62
Chi-sq stat for test of over-identification (p-val in parentheses)		3.590 (0.166)	3.590 (0.166)	0.444 (0.801)	0.444 (0.801)

Robust standard errors, clustered by scientist, in parentheses; significant at 10%; ** significant at 5%; *** significant

Appendix C: Alternative matching criteria

Tables 6 and 7 are based on a dataset that matches Fulbrights to control students in the same broad field of study who have (where possible) the same advisor and who graduated from the same institution within 3 years of the Fulbright. The text of the paper describes this procedure in greater detail. Here we consider alternative matching criteria and compare the results of un-instrumented OLS estimates of publication and citation variables based on our method of matching Fulbrights and controls described in the text and based on alternative matching criteria.

The top row of Tables A1 and A2 contains results from an OLS regression of the dependent variable on the foreign location dummy and the control variables used in our analysis. The second set of estimates are obtained via nearest-neighbor matching on these covariates, restricted to exact matches on the student's region of origin. Stata's "nnmatch" command was used to obtain these estimates. The third set of estimates is also nearest-neighbor matches, but require exact matching on region and field of study. Finally, the bottommost estimates are obtained after matching on the propensity score (using Stata's "attk" function), where the variables included in the propensity score are the same control variables used to obtain the other three sets of estimates.

Table C1 Coefficient on *LAGFORLOC* in un-instrumented publications equations using alternative matching criteria

	(1)	(2)	(3)	(4)	(5)
	Total Publications	First-authored pubs	Last-authored pubs	First-or last-authored pubs	High-impact pubs
OLS (Poisson)	-0.150	-0.050	-0.031	-0.093	-0.193
	(0.041)***	(0.020)**	(0.017)*	(0.027)***	(0.027)***
NN matching exactly on region	-0.192	-0.062	-0.014	-0.092	-0.205
	(0.056)***	(0.026)**	(0.019)	(0.033)***	(0.034)***
NN matching exactly on region & field	-0.192	-0.072	-0.008	-0.101	-0.197
	(0.059)***	(0.029)**	(0.022)	(0.037)***	(0.037)***
Matching on Propensity Score	-0.224	-0.084	-0.032	-0.129	-0.229
	(0.041)***	(0.026)***	(0.017)*	(0.034)***	(0.035)***

Table C2 Coefficient on *LAGFORLOC* in un-instrumented forward-citations equations using alternative matching criteria

	(1)	(2)	(3)	(4)	(5)
	Total citations	First-authored cites	Last-authored cites	First-or last-authored cites	High-impact cites
OLS	-3.624	-1.124	-0.253	-1.658	-3.611
	(0.698)***	(0.333)***	(0.153)*	(0.415)***	(0.605)***
NN matching exactly on region	-3.277	-1.094	-0.075	-1.596	-3.107
	(0.834)***	(0.389	(0.185)	(0.496)***	(0.704)***
NN matching exactly on region and field	-3.076	-1.103	-0.008	-1.591	-2.798
	(0.956)***	(0.428)***	(0.022)	(0.553)***	(0.816)***
Matching on Propensity Score	-3.434	-1.074	-0.222	-1.557	-3.222
	(0.598)***	(0.234)***	(0.144)*	(0.426)***	(0.545)***

Standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%