Research Mobility and Local Problem Solving*

Caroline Fry[1] and Gauri Subramani[2]

[1]Shidler College of Business, University of Hawai'i at Manoa

[2]College of Business, Lehigh University

Abstract

Mobile researchers have the potential to facilitate knowledge transfer and innovation targeting local challenges in the Global South. While prior research has largely examined South-North migration, this study investigates how South-South fellowships shape the research agendas of non-migrating scientists at home institutions. Specifically, we analyze the disease orientation of the non-migrant colleagues' research output before and after the fellowship. We document that non-migrants with a successful fellow in their organization publish more after the fellowship, and that, on average, their relative rate of publication in diseases that are highly prevalent in both their home country and the fellow's host country increases. However, we find a smaller observable increase in research on locally prevalent diseases among non-migrants exposed to fellows who travel to less similar Global South countries, including those with higher income levels and in different continents. Our findings highlight that mobile researchers can facilitate access to relevant knowledge and collaborations back at home, influencing the rate of research into local problems. They also reveal the importance of host destinations, suggesting that similarity between home and host countries shapes how international mobility shapes research on local challenges.

*Both authors contributed equally. We are grateful to Jeff Furman for sharing data and to the NBER Science of Science Small Grants for financial support. Thank you also to Akanksha Roy, Muzammil Khan, Sanchita Shrivastava, and Ben Subramani for research assistance.

1 Introduction

Migrants play an important role in the exchange of knowledge and resources between organizations and across borders (Nanda and Khanna, 2010; Agrawal et al., 2011a; Borjas and Doran, 2012; Ganguli, 2015; Wang, 2015; Choudhury and Kim, 2019; Hernandez and Kulchina, 2020). A growing body of literature highlights the positive impact of migrants, both those who stay in their host location and those who return back home, such as those participating in international fellowship programs, on innovation outcomes for individuals remaining at home. These outcomes include increased access to knowledge and international networks (Kahn and MacGarvie, 2016b; Fry and Furman, 2023; Kang and Eklund, 2024; Choudhury and Kim, 2019; Wang, 2015) as well as enhanced productivity (Kerr, 2008; Choudhury, 2016; Fry, 2023a). Recent research has found benefits in particular for scientists in the Global South[1] (Kahn and MacGarvie, 2016b; Fry and Furman, 2023; Choudhury, 2016; Fry, 2023a), who face structural barriers—including limited funding, infrastructure, and access to global knowledge networks–that migrants can help to overcome. Local scientists in the Global South, who possess deep contextual knowledge of pressing challenges, play a vital but often overlooked role in advancing solutions to their own local problems (Fry, 2023b), many of which are grossly under-studied (Moscona and Sastry, 2022; Vakili and McGahan, 2016; Kremer and Glennerster, 2004). Understanding how these constraints influence scientific progress and, more importantly, identifying strategies to mitigate them is essential for fostering innovation that directly addresses the challenges of the Global South.

While prior literature has advanced our understanding of how skilled migration affects innovation in home regions, especially in the Global South, the vast majority of existing research explores migration pathways from South to North.[2] We have a much more limited understanding of the role that South-South migration pathways play in innova-

[1]The Global South countries referenced in this study comprise most low- and middle-income countries and are listed here: https://owsd.net/sites/default/files/OWSD%20138%20Countries%20-%20Global%20South.pdf

[2]There are some exceptions to this: for example, Fry and Furman (2023), who study the same fellowship program as this paper, show that researchers in the Global South have more international collaborations following the departure of a colleague from their institution.

tion outcomes, particular those that address local needs. On the one hand, South-South migration could provide limited benefits to scientists as the migrant no longer facilitates access to resources and frontier knowledge in more traditional centers of knowledge. But on the other hand, South-South migrants could facilitate access to connections and knowledge that are more relevant and context-specific. This could result in improvements in innovation outcomes, particularly to address local problems. Indeed, critics of South-North pathways raise concerns about brain drain, questioning whether these programs ultimately benefit home-country innovation or primarily serve to integrate researchers into foreign networks with limited effect on progress toward home country problems. Despite this theoretical ambiguity, there is limited empirical evidence on whether and how South-South migration shapes innovation directed at local problems. Furthermore, the impact of migration is likely to depend on where scientists go. Different hosts even within the Global South are likely to provide varying levels and types of access to resources, networks, and relevant expertise, which may shape the nature and scope of knowledge transfer back to home institutions. As a high proportion of international fellowships for scientists in the Global South support mobility to higher-income or more scientifically advanced countries, it is crucial to understand how host destination characteristics impact innovation back home.

In this paper we examine the relationship between South-South migration and the rate of innovation targeting local challenges in the Global South. We ask: how does a researcher traveling to another country in the Global South shape the research output directed towards local problems of non-migrating colleagues at their home organization? We conduct our analysis using data on female scientists who apply for an international fellowship administered by the Organization for Women in Science for the Developing World (OSWD). This fellowship program provides opportunities for female scientists in developing countries to study in another country in the Global South. To assess the relationship between the fellowship and the research agendas of non-migrant researchers at the fellows' originating organizations, we construct a panel dataset of non-migrants' publication outcomes. We focus on non-migrants at the OWSD fellowship applicants' home institutions and in the same scientific fields, limiting to fellows and non-migrants

who publish in disease-related topics. Combining publication data with data on country-specific disease prevalence, we analyze changes in the likelihood that a non-migrant publishes on diseases that are more prevalent in their home (and/or the fellow's host) country before and after the fellowship. We compare the publication output of non-migrants whose colleagues received an OWSD fellowship with that of non-migrants whose colleagues applied for but did not receive the fellowship, and therefore do not travel abroad. This design allows us to control for career, field, disease, and temporal changes in publication output.

Consistent with prior literature in this area, we document that non-migrants exposed to a successful OWSD fellow publish more after the fellowship, as compared to non-migrants from institutions with an unsuccessful fellowship applicant. However, we also document a novel dynamic: that non-migrants exposed to a South-South fellow tend to increase their research focus on diseases that are highly prevalent in their home countries following the fellowship, particularly those diseases that are also highly prevalent in the host country. Conversely, they experience a smaller increase in research productivity in diseases that are prevalent in only the host or home country and not the other. We also document that this effect on overlapping home and host issues is mainly driven by publications on neglected tropical diseases,[3] indicating that the mobility of researchers plays a crucial role in advancing much-needed research on global health grand challenges.

Importantly, our findings suggest that while non-migrating colleagues of successful fellows see similar levels of increased productivity following the fellowship irrespective of the host destination, their relative focus on home-country disease research increases more when the home and host countries share similarities in income level, geography, or disease prevalence. In other words, non-migrating colleagues of fellows traveling to countries in different income bands, in different continents, and with fewer overlapping diseases experience smaller increases in the rate at which they publish on local problems. This finding suggests that the location of the fellowship can have a significant influence on the focus and outcomes of research of colleagues back at home. This observation in-

[3]Neglected tropical diseases (NTDs) are diseases that are primarily present in low-income and tropical areas that affect more than 1 billion people worldwide (https://www.who.int/news-room/questions-and-answers/item/neglected-tropical-diseases last accessed 8.23.24).

troduces a trade-off in choosing the location of the fellowship. Home institutions might have to weigh the value of local relevance versus the breadth of knowledge exposure or improved access to resources when deciding on supporting such initiatives.

We investigate possible mechanisms for these effects and do not find strong evidence that direct collaboration with the fellow, or competitive effects with the departing fellow, are primary drivers of non-migrants' increased focus on locally-relevant research topics. Instead, we find suggestive evidence that non-migrants increase the rate at which they collaborate with researchers in the fellows' new host institutions and countries, as well as the rate at which they utilize knowledge from the host institution. This access to relevant foreign connections and knowledge partially drives the observed relative increase in research on locally-salient problems.

We seek to bridge two literatures to contribute to our understanding of how mobile researchers can catalyse innovation addressing problems faced by countries in the Global South. While an extensive prior literature has documented the important role of mobile researchers in global knowledge diffusion, network formation, and ultimately in innovative productivity, in particular for scientists in the Global South (Saxenian, 2002; Rosenkopf and Almeida, 2003; Saxenian, 2005; Agrawal et al., 2006; Saxenian, 2007; Kerr, 2008; Oettl and Agrawal, 2008; Agrawal et al., 2011b; Choudhury, 2016; Kahn and MacGarvie, 2016b; Kerr, 2018; Choudhury and Kim, 2019; Kolympiris et al., 2019; Fry and Furman, 2023; Kang and Eklund, 2024), this literature has tended to examine migration pathways from South to North. Our findings highlight the important role that South-South migration can play in targeting context-specific problems back at home. However, the extent and nature of this influence depend on where researchers travel to. These findings complement a literature seeking to understand the drivers of innovation addressing grand challenges in the Global South (Arslan and Tarakci, 2022; Vakili and McGahan, 2016; Kremer and Glennerster, 2004), providing a potential explanation as to why rates of innovation addressing problems in the Global South remain low. Our findings have implications for policies and programs as well as for organizations aiming to leverage scientific mobility to tackle pressing societal challenges.

2 South-South Migration and Innovation Addressing Local Problems

Migrants play a key role in both transferring knowledge and connections from their home regions to their new locations (Borjas and Doran, 2012; Ganguli, 2015; Wang, 2015; Choudhury and Kim, 2019; Fry, 2023a; Fry and Ganguli, 2023). Whether they remain in their host location or return home, migrants can act as a bridge between individuals in their home and host regions. This brokerage effect has been well-established in scientific and innovative fields; for example, migrants can facilitate the formation of new collaborations and boost the patenting and publication productivity of their non-migrating peers (Choudhury, 2016; Fry and Furman, 2023; Fry, 2023a; Xie and Freeman, 2023). This effect is particularly salient for researchers based in countries in the Global South who have limited opportunities to access resources, knowledge and connections.

However, the majority of this research has explored migration from countries in the Global South to the North–such as fellowships supporting African scientists to study in the United States (Kahn and MacGarvie, 2016a; Fry, 2023a). There is much less evidence about how and when exposure to a migrating colleague moving to a country within the Global South affects non-migrants in their home institutions. Whilst Fry and Furman (2023) demonstrate that South-South fellows can facilitate network building across countries, there is a gap in our understanding of how this kind of migration can shape the production of knowledge towards local problems. This study addresses this gap by analyzing how the research trajectories of non-migrating colleagues of a fellow evolve after the South-South fellowship.

On the one hand, losing a colleague to a fellowship in another country can temporarily disrupt organizational capacity and thus inhibit the ability for individuals to make changes to their existing portfolios, which is costly and difficult (Myers, 2020). Further, while mobility is often seen as a mechanism for knowledge diffusion, the benefits of South-South migration may be more limited compared to migration toward well-established scientific hubs in the Global North. When scientists travel to another low- or middle-income country, they may face similar constraints in funding, infrastructure, and

access to high-impact networks as they did in their home institutions, which could have limited impact on their colleagues back at home, and ultimately on progress towards local problems.

On the other hand, a fellow leaving from their institution traveling to another country in the Global South might result in their non-migrating colleagues focusing more on salient local problems. Mobile individuals play a key role in creating a bridge from their home regions to knowledge and collaborators in their new locations (Borjas and Doran, 2012; Choudhury and Kim, 2019; Fry, 2023a; Fry and Ganguli, 2023; Ganguli, 2015). Insofar as the knowledge and collaborations transferred home are relevant to local problems, this could increase the rate of research into these problems. For example, scientists from malaria-endemic countries may gain new insights or discover alternative solutions, or meet new collaborators, during a fellowship in another Global South country where malaria is also prevalent. These new learnings and connections might be shared with colleagues back home, resulting in a change in their research agenda to better address local malaria challenges.

That said, differences in regional capabilities and priorities mean that 'where you go' influences 'what you learn.' The fellow's host location, in terms of its similarity and proximity to the home country, could affect the research focus of non-migrating colleagues through shaping the type of knowledge and collaborations transferred back home. Prior research has documented that knowledge is best transferred when there is some degree of overlap between home and host locations (Kolympiris et al., 2019; Lane et al., 2021; Kang and Eklund, 2024); therefore the similarity between the home and host country could not only influence the type of knowledge accessed, but also the levels.

In this paper we take a question-driven approach to explore two central research questions:

1. How does a researcher traveling to another country on a fellowship program influence the research conducted on local problems by non-migrating colleagues at their home organization?

2. To what extent is a change in research output directed towards local problems moderated by the similarity and proximity of a fellow's host location to their home country?

3 Setting, Data and Empirical Strategy

To study these questions, we explore the relationship between exposure to a scientist who is supported by a South-South fellowship operated by the Organization for Women in Science for the Developing World (OWSD) and outcomes of their non-migrating colleagues. We compare changes in the types of research output of non-migrating colleagues of a successful OWSD PhD fellow to changes in the types of research output of non-migrating colleagues of an unsuccessful OWSD PhD fellowship applicant. The colleagues of the unsuccessful applicant allow us to identify counterfactual outcomes in the absence of the OWSD fellowship.

3.1 Organization for Women in Science for the Developing World

OWSD provides grants to female scientists to study abroad during their PhDs. Female scientists in developing countries apply to OWSD for funding to support their PhD studies after having obtained approval from their proposed host institution in another country in the Global South. Fellows can apply for a full-time fellowship, which allows them to complete their PhD degrees at a host institution for up to four years, or they can pursue a "sandwich" fellowship which enables them to study for a shorter period in another country and then return to their home institutions to complete their degrees.

Fellows are selected based on the feasibility of completing their proposed research project, the match between the proposal and the potential host institution, and the quality of the connection that the applicant has made with the host and other relevant partner organizations. OWSD also aims to achieve diversity each year in terms of the fellows' locations and scientific fields. Importantly for our study, there is no evidence that OWSD selects fellows based on the alignment of their home organization's research interests and the prevalence of these topics in their home or proposed host countries.

The unsuccessful OWSD applicants in our sample are unlikely to have opportunities to travel abroad. In the absence of the OWSD fellowship, access to international opportunities is limited for female scientists in the Global South. OWSD applicants are generally "young women who may never have left their home town before, let alone their coun-

try" (Quadrio-Curzio et al., 2020). Indeed, our conversations with OWSD administrators confirmed that the majority of unsuccessful applicants fail to visit their proposed host institutions, providing us with a reasonable control group to which we can compare the effect of exposure to a fellow.

3.2 Sample construction

3.2.1 OWSD applicants

We begin with a sample of OWSD PhD fellows who applied to the program from 1996 to 2016 and agreed to have their data used for research. We have data on 315 applicants in total, which includes data on 134 successful applicants and 181 unsuccessful applicants.[4] We have information on their home institution, proposed host institution, year of fellowship, and field of study. The home institution is where they are based at the time of applying for the OWSD fellowship. We match the names, institutions and field of study of the fellows with their full publication record, if any, using the Elsevier Scopus database.

3.2.2 Non-migrant colleagues

As we are primarily interested in the effects of being *exposed* to a fellow as opposed to the effects on the fellow directly, we generate a sample of applicants' non-migrating colleagues in their home institution.

To identify non-migrant colleagues, we select individuals from the Elsevier Scopus publication database who are affiliated with applicants' home institutions in the three years prior to fellowship application and who publish in the same broad field (such as agricultural research or health research).[5] We extract each non-migrant scientist's entire publication history from the Elsevier Scopus publication database to generate annual variables on research output as described below. We retain publication records of sample

[4]There are a total of 2,952 applicants, out of which which 654 were awarded fellowships. Our sample of 315 applicants is representative of all applicants in terms of country of origin and field of study.

[5]Many scientists in our sample are exposed to more than one OWSD applicant. We treat each non-migrant $\times$ applicant pair as a separate observation, and cluster standard errors at the level of the colleague. Appendix Table A15 shows our results limiting only to the set of non-migrants treated by a single applicant.

scientists for the five years prior and up to ten years following the fellowship application year.[6]

Measuring the extent to which researchers focus on local problems is challenging. We do so by concentrating our publications and measures on diseases (more details in the next section) and leverage geographic variations in disease burden worldwide to create clear indicators of both the local nature of issues and the focus of a scientist's research agenda. Given this focus, we restrict our sample to the set of OWSD applicants and their non-migrating colleagues who ever publish in *any* disease. This leaves us with 3,667 treated non-migrant scientists (those whose colleague received an OWSD fellowship) and 5,256 control scientists (those whose colleague applied for but did not receive an OWSD fellowship) who ever publish in diseases.[7] This corresponds to 198 OWSD applicants, of whom just over 50% applied for sandwich fellowships. Of the non-sandwich successful fellows, just over 35% of them returned home following their fellowship, but conversations with OWSD administrators and fellows suggest that a large proportion of fellows remain connected with their home countries even if they remain abroad. Just over 59% of the non-migrant disease sample is from Nigeria, with other common home countries being South Africa, Zimbabwe, Cameroon, Bangladesh, Ethiopia, and Kenya. Around 40% of the sample identified South African host institutions, with other common host institutions situated in China, Malaysia, India, Pakistan, Kenya, and the Philippines.[8]

[6]For ancillary analysis, we impute the gender of each non-migrant scientist in the sample. We use a commercial database obtained from the company Ethnic Technologies that identifies gender using first names. However, we only have data on first names for 45% of the individuals in our data because the remainder of our individuals do not have a first name listed on their publications. For the sample of people for whom we do have first names, we can identify gender unambiguously for 35%. Our gender-identified sample of non-migrants is 27% female and 73% male.

[7]In total (not limiting to those who ever publish on diseases), we have 7,831 treated non-migrant scientists and 11,639 control scientists.

[8]Appendix Table A1 lists the set of home and proposed host countries for non-migrating colleagues. In Appendix Table A9, we run our main analyses excluding non-migrants who are exposed to an applicant who selected an institution in South Africa to ensure that our results are not driven only by one host country.

3.3 Data and measurement

3.3.1 Disease publications and prevalence

In addition to annual counts of total publications, we develop several measures of research topics of focus. For all non-migrating scientists' publications, we identify whether each publication is related to diseases that are prevalent in their home country and/or the applicant's proposed host country.[9] To measure the disease focus of each publication, we use a key word search to identify whether a publication covers a specific disease topic out of all 156 diseases[10] included in the Global Burden of Disease database (i.e., for each disease we identify whether it is mentioned in the title or abstract of the publication). 30% of publications in our sample match to at least one disease.

We then identify the country-specific prevalence of each of our 156 diseases using the Global Burden of Disease database. The majority of our analysis utilizes the disease prevalence measure of the percent of 'Disability-Adjusted Life Years' in a country in the year prior to the publication year (to allow for publication lag). We consider a disease to be of high prevalence in a country if its prevalence is greater than the median disease prevalence as calculated from our full set of diseases and countries.[11] There is significant variation in highly-prevalent diseases across home and host countries. For some home countries, high-prevalence diseases include liver cancer and diabetes, whilst for others they include malaria and schistosomiasis.[12]

We leverage our data on diseases featured in publication abstracts and titles to generate several key measures that capture the local nature of research. In our primary analysis, we measure individual-level annual publication counts to evaluate the number of articles

[9] As a robustness check, we validate our main results using data on crop and pest prevalence. This is discussed in greater detail in Appendix B. In a manner identical to the disease identification, we use key word searches of the title and abstract of the publications to identify whether each publication covers a crop or pest topic. We next merge the set of all sample publications to data on country-specific crop/pest prevalence in using the Centre for Agriculture and Bioscience International's (CABI) Crop Protection Compendium (CPC). Results of these analyses are presented in Appendix Table A16

[10] We exclude diseases that are not endemic to physical environments; eg, "falls", "drowning", "fire, heat, and hot substances".

[11] Appendix Table A3 shows our results using alternative measures of disease prevalence.

[12] We are particularly interested in research on Neglected Tropic Diseases (NTDs); thus, in additional analysis, we narrow our set of diseases to specifically examine non-migrant scientists' proclivities to publish on NTDs.

a non-migrant publishes[13] on diseases that are of high prevalence in:

- The home country
- The host country[14]
- Both the home and host countries
- The home country but *not* the host country
- The host country but *not* the home country

Figure 1 provides a graphical representation of the five different types of publications we explore in this paper.

3.3.2 Home–Host Country Similarity

In order to evaluate the extent to which host country features shape the non-migrant scientists' research agenda we generate several measures of home and host country similarity. Our first measure is a dummy variable that represents whether or not the home and host countries have similar levels of income per the World Bank's country classification (high, upper middle, lower middle, or low). Next, we generate a dummy variable for whether the home and host country are in the same continent. Third, we leverage our data on disease prevalence to identify countries that have a high proportion of shared diseases. We look at home–host country pairs that are in the top 25 percent of country pairs when it comes to shared disease prevalence, as these represent countries that have a high level of problem overlap.

[13] Around 8% of our sample publications cover more than one disease. In this instance, if there is at least one disease covered in the publication in the named measurement category, the publication takes the value of 1 for that measure. If there is more than one high prevalence disease covered in the named category, the publication takes the value of the count of diseases covered. These measures therefore reflect the number of disease publications per person per year.

[14] For non-migrating colleagues, the host country is where their colleague who applies for the fellowship identifies as the country they will visit.

3.3.3 Collaborations

To explore any mechanisms driving our observed effects, we evaluate whether the fellow shapes the non-migrants' research trajectory directly by collaborating with non-migrant colleagues or indirectly via facilitating international collaborations. We generate variables that reflect the rate at which non-migrant colleagues collaborate with both the OWSD applicant and with individuals in the applicant's host country. Namely, we evaluate the author lists of each publication in our sample and identify if they include: 1) the OWSD applicant, 2) anyone in the host institution, and 3) anyone in the host country.

3.3.4 Access to new knowledge

We also explore the extent to which the fellow facilitates access to new knowledge from the host institution. We collect data from Elsevier Scopus on all host institutions' publications in applicants' areas of study. We then extract the keywords associated with each publication to construct a list of keywords that describe the research conducted in each host institution in each year.[15] For example, for a specific host institution that had a fellow visiting for a PhD in biology, we collect data on all host institution publications in biology per year.

We count the proportion of host institution keywords that the non-migrating colleague uses in their own research before and after the fellowship. Through this analysis, we are able to evaluate if, after the fellowship, non-migrating colleagues' research more closely resembles the work done at the host institution.[16]

[15] Keywords are indexed by Elsevier Scopus.

[16] We limit to pre-fellowship event host institution keywords to avoid any contamination that may occur because of the fellowship itself (that is, if the fellow introduces home country ideas or topics to the host institution).

3.4 Empirical strategy

In order to identify the relationship between exposure to a fellow and non-migrant scientists' research agendas, we leverage a differences-in-differences design. We compare publication patterns on specific disease topics before a colleague's fellowship versus after, relative to the change in non-fellow-exposed scientists' outcomes, using a scientist fixed effect specification. Our preferred analysis uses data at the person-year level and our outcome variables measure the annual count of publications in each category of interest (publications on home topics, publications on host topics, publications on topics prevalent in both home and host contexts, and publications on topics prevalent only in home or host contexts). Our estimating equation (1) relates a non-migrant scientist i's publications in year t to the timing of the fellowship. The first year of treatment is the year after fellowship application (to allow for publication lag and migration time). We estimate the following:

$$E\left[y_{it}|X_{it}\right] = exp\Big[\beta_0 + \beta_1(After_Fellowship \times Successful_Applicant)_{it} + \beta_2(After_Fellowship)_t + \gamma(publications)_{it} + \sigma_t + \tau_i + f(AGE_{it})\Big] \quad (1)$$

Where $\gamma(publications)_{it}$ is a control for an individual's annual count of publications, σ_t are year fixed effects, τ_i are individual fixed effects, and $f(AGE_{it})$ is a flexible function for scientist's career age. Standard errors are clustered at the person level.[17]

While our primary results are based on a person $\times$ year-level dataset, we also conduct analyses at the person $\times$ year $\times$ disease-level; that is, we construct a dataset that identifies the number of publications an individual has on each of the 156 diseases in our sample in each year. In these specifications, we include person $\times$ disease and disease $\times$ year fixed effects, as well as person $\times$ year fixed effects. This allows us to account for unobserved

[17] In Appendix Table A17, we estimate our main results using a Poisson Maximum Likelihood model.

propensities of individuals to publish in particular diseases overall, variation in disease research output trends over time, and person-level general research output trends over time.

In heterogeneity analysis we include a triple interaction between the $After_Fellowship$ and $Successful_Applicant$ variables with an indicator variable *Similar_Income* or, in a separate model, *Same_Continent* or *High_Disease_Overlap*.

One concern with the baseline, and heterogeneity, analysis is that in the absence of the OWSD fellowship, the research trajectory of the treated scientists, or non-migrants exposed to a fellow, would have changed even in the absence of the fellowship. This could happen if the likelihood of receiving the fellowship is correlated with home institution strength in local problem research. Another concern is that the home and host institutions were developing informal relationships and that corresponds to the success of the fellowship application. To the extent that this is more likely in cases with similar home and host countries, this would bias our results. We address this concern in two main ways.

First, we are evaluating non-migrant colleagues' publication patterns, not fellows' own publications; based on conversations with OWSD administrators, we anticipate that an applicant's fellowship outcome is unlikely to be based on their colleagues' research trajectories. However, in the event that it is, our empirical strategy accounts for underlying heterogeneity between individual non-migrants through the use of the fixed-effects specification. Therefore, our effect measures the change in outcomes pre- and post-fellowship, net of any differences between individuals in the sample. Because we are relying on the timing of the fellowship event to identify our main relationship, we test whether there were already differences in publication patterns between our treatment and control groups in the pre-treatment periods. Figures 2 and 3 present these results for publications on home-country diseases. These figures demonstrates that there is no significant pre-trend; treated and control colleagues publish on diseases present in their home coun-

tries at equivalent rates prior to the fellowship.[18]

Second, we conduct a series of tests (described in detail later in Section 4.4) to assess the extent to which pre-treatment characteristics of the fellow and the home institution, in particular – the fellow's and the home institution's proclivities to publish in local problems – is driving the observed effects. Our results are robust to accounting for these features that could potentially drive our results.

While we attempt to account for alternative explanations that could be driving our findings, in the absence of an experimental context, we are unable to claim causality. Rather, we are documenting an important relationship between mobile researchers and the rate of research into local problems back at home.

4 Results

4.1 Descriptive Statistics

In Table 1, we provide details on the treated and control samples at the year of application. Columns 1 and 2 show averages and standard errors in brackets for our control and treatment groups, respectively. Column 3 shows the raw difference in means between treated and control groups, and column 4 shows the t-statistic for significance of that difference.[19] As we see in this table, there is balance of key productivity statistics, career, gender and research direction at the time of treatment. While there are some small

[18] Treatment occurs at various times in our setting; not all fellows are treated in the same year. When treatment timing varies, our estimation approach using difference-in-differences estimation compares already-treated units and later-treated units (a "forbidden comparison" (Goodman-Bacon, 2021)), which can negatively weight units that are treated earlier and lead to inaccurate estimates. As a result, we also evaluate our results using differences-in-differences estimators that avoid comparing already-treated units to later-treated units. We identify that the estimators developed by Borusyak et al. (2022) and Gardner (2022) are best suited for our setting. Appendix B provides further detail about our empirical approach. We conduct our main analysis using these estimators and Figure B1 shows pre-trend figures using these results. Our results are very similar to what we find in our Ordinary Least Squares analyses.

[19] The t-test compares the control and treatment group means and assumes a null hypothesis that the two means are equal. If the t-statistic is greater than 1.96, we can reject the null hypothesis that there is no difference between means at a 5% level of significance.

differences in the collaboration rate between home and host country researchers between treated and control scientists, for our empirical approach, we include individual-level fixed effects in our regressions and run a number of of robustness checks to verify if it is these non-migrants driving our results.

Across the sample of treated and control scientists, on average, sample scientists publish more on home country diseases than host country diseases, and these papers on home-country topics make up a large proportion (on average, 42%) of their overall publications pre-treatment. Collaborations with outside researchers are fairly rare, and non-migrants produce just 0.09-0.11 publications with host country researchers in the five years prior to treatment.

4.2 Changes in research addressing local problems

We start by examining the effect of exposure to a fellow on non-migrants' overall productivity. Table 2 presents these results.[20] Consistent with prior research in this area, we show that non-migrants whose colleagues received an OWSD fellowship publish more after their exposure to a fellow (column 1). Non-migrants' annual publication outputs are 6.3% higher following their colleagues' initiation of a fellowship abroad.

After controlling for this overall increase in annual publications, we see that treated non-migrants increase their relative publication output on disease topics (column 2) whilst decreasing the relative focus on non-disease topics (column 3), and the majority of this increase in disease research is in neglected tropical diseases (NTDs) (column 4) which are diseases predominantly affecting communities in the Global South. In the remainder of the paper, we explore the relative focus on different categories of diseases, controlling for overall annual publications throughout which allows us to measure the relative increase in a topic focus accounting for changes in productivity.

[20] Note that the coefficients for $After_Fellowship$ and $Successful_Applicant$ are absorbed by fixed effects for years since treatment, and individual level fixed effects.

Our central question is: do non-migrants publish more on diseases that are prevalent in their home countries after exposure to a fellow from their institution? Table 3 shows results for publication counts on diseases that are of high prevalence in home (and host) countries. Figure 1 provides a graphical representation of the five different sets of publications we explore in this table. We start by evaluating the change in the count of publications on high-prevalence home country diseases. We show the raw data representing the changes in publication output on locally-prevalent diseases plotted in Figure 2 and the corresponding regression results in column 1 of Table 3. Through looking at both the total number of publications annually in high prevalence diseases, controlling for overall annual publications (top row), and looking at the proportion of publications in home country diseases (bottom row), we see that non-migrants are more likely to publish on diseases that are high prevalence in their home countries after the fellowship (column 1). We also find that they are more likely to publish on diseases that are high prevalence in their colleagues' host countries (column 2). However, these effects are mostly driven by publications on diseases that are high prevalence in *both* home and host countries (column 3).[21] with no observable increase in research on diseases prevalent only in the host country and not in the home country (column 5).[22][23][24] These diseases represent significant grand challenges to which non-migrants in the Global South are directly exposed, giving them a unique comparative advantage in studying these issues.

Figure 3 illustrates the temporal nature of these changes in research focus; non-migrants

[21] In very few cases scientists can be both treated and control. We drop those scientists in Table A15 and the results are much greater in magnitude.

[22] In appendix Table A2, we show that this increase in relative focus in home , and home and host country disease publications is not just driven by low-quality publications. We measure the sum of quality adjusted-publications, as well as the average quality of the annual publications of treated scientists in these diseases areas, demonstrating increases in both measures after the fellowship.

[23] We consider how some of our measures may contribute to the results. Appendix Table A3 shows our results using alternative measures and margins of disease prevalence. We see that even using these alternate measures, the propensity at which non-migrants increase their relative focus on local, and particularly on shared, problems increases.

[24] In Appendix Table A4, we show that this increased focus on local topics is exclusively driven by non-migrants' increased publications in high prevalence neglected tropical diseases (NTDs). In fact, there is no effect on publications on non-NTD topics.

make a relatively quick change in research focus towards home country diseases following their colleagues' fellowship (about two years after the fellowship, we see a noticeable change), and this effect is sustained (although with slightly less intensity) until 10 years after the fellowship.

4.3 Heterogeneity by Home–Host Country Similarity

Next, we examine whether the similarity between home and host country influences productivity more generally or specifically shapes the research direction of non-migrating colleagues. Table 4 shows that while non-migrating colleagues of successful fellowship applicants experience a significant increase in their total publication output after the fellowship, as indicated by the positive and significant coefficient in column 1, the effect on shared home and host country disease-related publications is amplified when the home and host countries are in the same income band (Panel A), are on the same continent (Panel B) or share high disease overlap (Panel C). This suggests that access to relevant information is a critical factor in shaping the relationship between exposure to a fellow and the relative research focus within the home country.[25]

4.4 Potential Mechanisms

We investigate several potential mechanisms that could be driving our observed effects. Namely, fellows might collaborate with researchers at home and directly transfer new knowledge or resources that support research into local problems. Or, they could indirectly facilitate access to knowledge and collaborators relevant to local problems.

[25] In Table A5, we show summary statistics for our treated and control groups on our outcomes of interest for these measures of similarity. We are interested in evaluating if our treatment and control samples differ *more* for the set of researchers exposed to similar contexts than for our full sample, which could bias our results. On the whole, we don't see that this is the case.

4.4.1 Direct effect

First, we consider that fellows could directly influence their home country colleagues' research trajectories either by collaborating with them, advising them, or acting as role models who research local topics on an international stage. Non-migrating colleagues could also be influenced by a departing fellow through competitive dynamics, such as the freeing up of resources, particularly as it pertains to specific topic areas, allowing them to move into those topic areas (Borjas and Doran, 2015; Azoulay et al., 2019).

To understand if these direct effects are driving our results, we evaluate our results separately for non-migrants based on whether or not their fellow publishes in home country diseases.[26] In Table 5, we start by looking at whether successful fellows themselves orient their research trajectories towards home country topics (columns 1 and 2). Interestingly, we see that treated fellows are not significantly more likely to research home country diseases after they migrate. Next, we explore any heterogeneity of results according to the fellows' orientation. Column 3 shows results for colleagues of applicants who never publish on home country diseases, and column 4 shows results for colleagues of applicants who do publish on home country diseases. In both of these samples, we see a similar magnitude of increase in the non-migrating colleague's propensity to publish on home country diseases, which further suggests that collaboration (or competition) with the fellow, or specific role model or advising channels, are not likely to be driving the publication outcomes we observe. Note also that collaboration with a fellow is extremely unlikely; just 0.6% of sample scientists in this sample ever have a publication with a fellow. In addition to a low rate of collaboration with the fellow themselves, we see that our sample of non-migrant scientists are not more likely to publish with fellowship recipients following the fellowship (column 5). In column 6, we drop non-migrants who publish with the fellow from our sample and our results remain robust.

[26] We identify whether fellowship applicants publish in any home country disease at any point in the 5 years before or 10 years after applying for the fellowship.

Together, we fail to find strong evidence that indicates that non-migrants' increased research investment in local diseases is driven by the fellow directly. This also serves as a robust test: insofar as we are concerned that a fellow is being sent to a similar host institution *in order to* strengthen home institution research into local problems, we would expect the fellows' local research focus to be correlated with non-migrant outcomes. We fail to find evidence for this channel, which is comforting.

4.4.2 Indirect effect

Another way in which exposure to a fellow can affect a home country colleague is if they facilitate connections or knowledge flows between their host country and home country colleagues. As shown in Fry and Furman (2023), fellowships of this nature can increase the rate at which home country colleagues collaborate with the migrant's new colleagues in the host country. Given this, we first examine the coauthorship patterns with host country researchers to understand the extent to which access to collaborators could be driving the changes we see in non-migrants' research agendas.

Our results are presented in Table 6. In columns 1 and 2, and consistent with prior research in this area, we see that after being exposed to a fellow, non-migrants are more likely to coauthor on any subject with a researcher from the host institution and host country as compared to the control non-migrants. To further investigate this collaboration mechanism, we examine the content of publications coauthored by non-migrants and host country scientists. We evaluate whether these publications are on high prevalence home country diseases (column 3), non-home country diseases (column 4), shared home and host diseases (column 5), or diseases that are prevalent only in the home country and not host country (column 6). We see that coauthored publications focus predominantly on diseases that are prevalent in *both* home and host countries, and that there is no effect on collaborative research on diseases that are present in one environment but not the other. These results imply that the fellowship creates opportunities for home and host

colleagues to invest in areas of common interest and relevance.[27]

Lastly, fellows can have an impact by facilitating the flow of knowledge between countries. Several studies have documented that migrants play this role (Wang, 2015; Kahn and MacGarvie, 2016b). It is plausible that an increase in access to new knowledge, particularly pertaining to problems shared between the home and host countries, could influence a non-migrants' research orientation.

Specifically, we are interested in evaluating if after treatment, non-migrating colleagues' research begins to more closely resemble the work being done in the host institution. We investigate this by testing whether treated non-migrants' are more likely to use key words associated with the host institution after the fellowship. Table 7 shows these results. In columns 1-4, we see that treated non-migrants use a higher proportion of host institution keywords in all of their publications after the fellowship. Then, we focus specifically on publications on home country diseases (column 5) and both home and host country diseases (column 6). We see that the increase in use of host country keywords is both greater in magnitude and more significant in these contexts. That is, non-migrating colleagues are more likely to use host institution keywords in their publications on both local and shared problems after exposure to a fellow.[28]

Together, this analysis provides suggestive evidence that non-migrants' increased investment in local research following exposure to a fellow leaving from their institution is driven at least in part by gaining access to connections and knowledge in the context of shared problems.

[27] In Appendix Table A6, we conduct this analysis with triple-differences to evaluate if collaboration behavior changes based on home and host similarity. Interestingly, we see that for individuals exposed to peers traveling to countries with similar income and high disease overlap, there is no difference in the volume of collaborations.

[28] Note that for this analysis, it is difficult to explore the use of keywords in non-home country publications given the relative dearth of those publications and our use of a single post-treatment measure which drastically reduces the sample size available for analysis.

4.5 Alternative Explanations

To assess the extent to which the increase we observe in non-migrants colleagues' relative focus on problems shared by the home and the host country could be driven by alternative explanations, we conduct several tests.

First, to account for the possibility that there are person-disease trends or disease-specific trends in research that might be driving our results, we conduct our analysis at the person $\times$ year $\times$ disease level in appendix Table A7. Note that this analysis includes scientist $\times$ disease fixed effects as well as disease $\times$ year fixed-effects and scientist $\times$ year fixed effects. The scientist $\times$ disease fixed effects remove time-invariant differences in the extent to which scientists research on particular diseases. Disease $\times$ year fixed effects remove the possibility that any trends observed are driven by general changes in the level of research into given diseases over the observed time period. Scientist $\times$ year fixed effects account for any changes in productivity of scientists over years which could generate changes in levels of research in given diseases. As we see in Table A7, our results are robust to this alternative data structure. In fact, the coefficient on the diseases prevalent at home and not in the host country is close to zero using this specification, suggesting that the small effect picked up in the baseline specification was due to one of these alternative channels (such as an increase in research more generally into those diseases).

Second, we evaluate the possibility that the observed effect is driven by OWSD fellows who are traveling from institutions with a greater focus on shared home and host diseases. In other words: are fellows traveling to countries with similar problems more likely to be from institutions with a focus on those problems, and is that driving our effects? We test this in a few ways in Appendix Table A8. First, we drop the minority of non-migrants who had coauthored publications with anyone in the host country prior to the fellowship, and second, we split the sample of non-migrants into those in an institution with above median relative publication output prior to the treatment on home and

host diseases, and the remainder. We fail to find any evidence that institutions with prior relationships with the host country or that non-migrants in institutions with high home and host disease focus are driving the observed effects. Further, we run the full analysis excluding fellows who travel to South Africa, the most common host country in our sample, in Appendix Table A9. Our results remain robust.

Third, we compare colleagues of OWSD fellows who had "sandwich" fellowships for just to those who had longer, full-time PhD fellowships. Of our sample, 36% apply for full-time fellowships. Presumably, the sandwich fellows had less time in which to build relationships and connect their home and host country colleagues. In Table A10, we see that our results are meaningfully smaller for sandwich fellows. This suggests that the intensity of the connection to the host institution does matter for the outcomes we find. While not the focus of the paper, in a reduced sample of non-migrants for which we could obtain gender data, we show that the female non-migrants are more likely to increase their local research focus following the fellowship (Appendix Table A11). Given the potential for homophily, this is likely to be the group of non-migrants most affected by the departure of a female fellow from their institution, thus providing further evidence that the observed effects are less likely to be driven by institution-specific characteristics, and seem to be driven by the fellowship itself.

Fourth, it could be that non-migrating colleagues are more likely to study problems that are of high prevalence globally (for example, HIV) and we are inaccurately characterizing these as shared home and host problems as opposed to problems that are shared across a variety of countries because of their global prevalence. In addition to dropping NTDs with higher global prevalence and attention (malaria, HIV and TB) in Appendix Table A4, in columns 2-4 of Table A12, we limit only to diseases with high global prevalence. Namely, we want to test whether non-migrants are orienting their efforts towards diseases that are high prevalence in general, not necessarily specific to the home and host country. We see that non-migrants are actually no more likely to focus on high global

prevalence diseases that are also high prevalence in home not host countries, host only, and diseases that are not high prevalence in the home country, after the fellowship. This result suggests that we are observing an increase in focus towards diseases specifically prevalent in the home and host countries, as opposed to towards diseases that are important globally.

We also test whether our effects are mechanical by randomly assigning treated colleagues to a placebo host country. Appendix Table A13 presents these results. The results of this counterfactual exercise are directionally similar but meaningfully smaller (2% of the size) of the effects we find in our main analysis.[29].

Lastly, we are proposing that the effects of the fellowship are indirect, but what if the non-migrant also migrates and that is driving our observed effects? We evaluate this concern in Appendix Table A14 by excluding migrating non-migrants from the sample (i.e. any non-migrant who is affiliated with another institution in another country at any point in the post-fellowship period). The results remain similar using this alternative sample.

Overall, these analyses show that even accounting for additional institution and disease-level observable characteristics, the effects we find on non-migrating colleagues' research trajectories are robust.

4.6 Crop and Pest Results

To complement our analysis of disease publications and to check whether the results we find are unique to the disease setting, we evaluate publication patterns on crops and pests.[30] This analysis is less granular as our data on crop and pest prevalence is binary and time invariant; there is a dummy variable that takes the value of 1 if a specific crop

[29]Of note, many of the placebo host countries disease prevalence overlaps with actual home countries and so in column 2 we drop those observations, and find a null effect of the fellowship to this placebo host on home and host disease research focus

[30]This analysis is discussed in greater detail in Appendix B.

or pest has appeared a country and 0 otherwise, but we do not have data on when this measure was collected or how prevalent in a continuous manner the crop or pest is.

We conduct our baseline analysis with this data in an identical manner as we do with disease outcomes and explore if there are changes in the frequency with which agriculture research focused non-migrants publish on local or global crops and pests. In Appendix Table A16, we see that non-migrants are more likely to publish on crops and pests that are present in their home countries and on those that are present in their colleague's host countries after the fellowship. However, this effect is not significant when we focus only on crops and pests that are present in home or host countries (columns 4 and 5). These findings mirror our disease results.

5 Discussion

In this paper, we explore the extent to which there exists a relationship between exposure to a fellow traveling from and to a country in the Global South and the research agendas of their non-migrating colleagues. Specifically, we ask whether non-migrant scientists at the fellow's home institution change their propensity to research more locally-relevant problems after the departure of the fellow in the context of disease research. We document that following the departure of the fellow, the non-migrating colleagues of the fellow are more productive, and this effect is driven by an increase in non-migrants' propensities to publish papers on diseases with high prevalence in their home countries. This result appears to be driven specifically by diseases shared by the home and host countries and is stronger when the fellow travels to similar host countries in terms of geography, income levels and with a shared set of high-prevalence problems. We explore mechanisms and find suggestive evidence that an increase in collaborations with host country scientists and an increase in access to knowledge from the host institution drive our observed changes in research focus.

This work is subject to several limitations. First, we study the implications of South-to-South migration of scientists, and as a result, we cannot be sure that our results will generalize to fellowships supporting migration to developed countries. Second, insofar as the organizations with successful fellows were those also more likely to collaborate with the host countries and contribute to local research, we could be overestimating the effects. While we attempt to address potential biases through robustness tests and careful consideration of the mechanisms, our results should be considered as descriptive rather than causal. Third, we observe only female fellows and thus our results may not hold if migrating fellows are both male and female. While these limitations may affect the generalizability of our analysis, we believe that given that our work is one of the first explorations into the relationship between research mobility and progress towards local, often grand, challenges, our results provide insights that are relevant for both theory and practice.

We contribute to two main literatures. First, we contribute to a growing literature on the role of migration in organizational innovation outcomes (Saxenian, 2002, 2005, 2007; Agrawal et al., 2006; Kerr, 2008; Oettl and Agrawal, 2008; Agrawal et al., 2011b; Kerr, 2018; Kahn and MacGarvie, 2016b; Fry and Furman, 2023; Choudhury and Kim, 2019; Choudhury, 2016). Our results extend this research by demonstrating that by facilitating the flow of knowledge, connections and resources back to their home regions —especially from countries facing similar issues— South-South mobile researchers can shape the research orientation of their colleagues back at home. By highlighting the important role of shared problems in shaping innovative outcomes, we propose that the influence of researcher mobility depends on the characteristics of the host locations.

Our findings also contribute to a growing literature the drivers of innovation addressing grand challenges, particularly problems affecting the Global South (Arslan and Tarakci, 2022; Vakili and McGahan, 2016; Kremer and Glennerster, 2004). We highlight the role of mobile individuals and document how and when exposure to fellows spend-

ing time in another country in the Global South can accelerate innovation addressing local problems, many of which are grand challenges. But, by documenting the important role of the host location features in shaping innovation towards local problems, we surface important insights on some of the barriers to innovation towards problems in the Global South.

This research has implications for entrepreneurs, managers, organizations, and policymakers seeking to enhance and direct innovation. Organizations, broadly defined, particularly those aiming to make progress on specific topics, can effectively leverage mobile individuals to provide a strategic advantage in innovation. For example, by creating opportunities for international immersion for employees, organizations and managers can stimulate cross-country collaborations and facilitate access to knowledge that helps others within the organization contribute to targeted innovative solutions. Further, this research suggests that to direct innovation, particular attention should be paid to mobility across locations with shared problems. In other words, organizations and managers can select contexts for relationships based on the kinds of knowledge they seek to integrate. For policymakers, this research can provide evidence of the value of opportunities for international research exchange and inform the development of frameworks to harness the potential of mobile scientists and innovators.

References

Agrawal, A., Cockburn, I., and McHale, J. (2006). Gone but not forgotten: knowledge flows, labor mobility, and enduring social relationships. *Journal of Economic Geography*, 6(5):571–591.

Agrawal, A., Kapur, D., McHale, J., and Oettl, A. (2011a). Brain drain or brain bank? the impact of skilled emigration on poor-country innovation. *Journal of Urban Economics*, 69(1):43–55.

Agrawal, A., Kapur, D., McHale, J., and Oettl, A. (2011b). Brain drain or brain bank? The impact of skilled emigration on poor-country innovation. *Journal of Urban Economics*, 69(1):43–55.

Arslan, B. and Tarakci, M. (2022). Negative spillovers across partnerships for responsible innovation: Evidence from the 2014 ebola outbreak. *Journal of Management Studies*, 59(1):126–162.

Azoulay, P., Fons-Rosen, C., and Graff Zivin, J. S. (2019). Does Science Advance One Funeral at a Time? *American Economic Review*, 109(8):2889–2920.

Borjas, G. J. and Doran, K. B. (2012). The Collapse of the Soviet Union and the Productivity of American Mathematicians*. *The Quarterly Journal of Economics*, 127(3):1143–1203.

Borjas, G. J. and Doran, K. B. (2015). Which Peers Matter? The Relative Impacts of Collaborators, Colleagues, and Competitors. *The Review of Economics and Statistics*, 97(5):1104–1117.

Borusyak, K., Jaravel, X., and Spiess, J. (2022). Revisiting Event Study Designs: Robust and Efficient Estimation. Technical Report 2108.12419, arXiv.org. Publication Title: Papers.

Butts, K. and Gardner, J. (2022). did2s: Two-Stage Difference-in-Differences. *The R Journal*, 14(3):162–173.

Callaway, B. and Sant'Anna, P. H. C. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Choudhury, P. (2016). Return migration and geography of innovation in MNEs: a natural experiment of knowledge production by local workers reporting to return migrants. *Journal of Economic Geography*, 16(3):585–610.

Choudhury, P. and Kim, D. Y. (2019). The ethnic migrant inventor effect: Codification and recombination of knowledge across borders. *Strategic Management Journal*, 40(2):203–229. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/smj.2977.

de Chaisemartin, C. and D'Haultfœuille, X. (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review*, 110(9):2964–2996.

Fry, C. and Furman, J. L. (2023). Migration and Global Network Formation: Evidence from Female Scientists in Developing Countries. *Organization Science*. Publisher: INFORMS.

Fry, C. and Ganguli, I. (2023). Return on Returns: Building Scientific Capacity in AIDS Endemic Countries.

Fry, C. V. (2023a). Bridging the Gap: Evidence from the Return Migration of African Scientists. *Organization Science*, 34(1):404–432. Publisher: INFORMS.

Fry, C. V. (2023b). Crisis and the trajectory of science: Evidence from the 2014 ebola outbreak. *Review of Economics and Statistics*, 105(4):1028–1038.

Ganguli, I. (2015). Immigration and Ideas: What Did Russian Scientists "Bring" to the United States? *Journal of Labor Economics*, 33(S1):S257–S288. Publisher: The University of Chicago Press.

Gardner, J. (2022). Two-stage differences in differences. arXiv:2207.05943 [econ].

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277.

Hernandez, E. and Kulchina, E. (2020). Immigrants and foreign firm performance. *Organization Science*, 31(4):797–820.

Kahn, S. and MacGarvie, M. (2016a). Do return requirements increase international knowledge diffusion? evidence from the fulbright program. *Research Policy*, 45(6):1304–1322.

Kahn, S. and MacGarvie, M. J. (2016b). How Important Is U.S. Location for Research in Science? *The Review of Economics and Statistics*, 98(2):397–414.

Kang, H. and Eklund, J. C. (2024). Fast Friends: The Impact of Short-Term Visits on Firms' Invention Outcomes. *Organization Science*. Publisher: INFORMS.

Kerr, W. R. (2008). Ethnic Scientific Communities and International Technology Diffusion. *The Review of Economics and Statistics*, 90(3):518–537.

Kerr, W. R. (2018). *The Gift of Global Talent: How Migration Shapes Business, Economy & Society*. Stanford University Press, Stanford.

Kolympiris, C., Hoenen, S., and Klein, P. G. (2019). Learning by Seconding: Evidence from National Science Foundation Rotators. *Organization Science*. Publisher: INFORMS.

Kremer, M. and Glennerster, R. (2004). *Strong Medicine: Creating Incentives for Pharmaceutical Research on Neglected Diseases*. Princeton University Press.

Lane, J. N., Ganguli, I., Gaule, P., Guinan, E., and Lakhani, K. R. (2021). Engineering serendipity: When does knowledge sharing lead to knowledge production? *Strategic Management Journal*, 42(6):1215–1244. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/smj.3256.

Moscona, J. and Sastry, K. (2022). Inappropriate Technology: Evidence from Global Agriculture.

Myers, K. (2020). The Elasticity of Science. *American Economic Journal: Applied Economics*, 12(4):103–134.

Nanda, R. and Khanna, T. (2010). Diasporas and domestic entrepreneurs: Evidence from the indian software industry. *Journal of Economics & Management Strategy*, 19(4):991–1012.

Oettl, A. and Agrawal, A. (2008). International labor mobility and knowledge flow externalities. *Journal of International Business Studies*, 39(8):1242–1260. Publisher: Palgrave Macmillan.

Quadrio-Curzio, A., Blowers, T., and Thomson, J. (2020). Women, science and development: The leading role of OWSD. *Economia Politica*, 37(1):1–12.

Rambachan, A. and Roth, J. (2023). A More Credible Approach to Parallel Trends. *The Review of Economic Studies*, page rdad018.

Rosenkopf, L. and Almeida, P. (2003). Overcoming Local Search Through Alliances and Mobility. *Management Science*, 49(6):751–766. Publisher: INFORMS.

Roth, J., Sant'Anna, P. H., Bilinski, A., and Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics*, 235(2):2218–2244.

Saxenian, A. (2002). Silicon Valley's New Immigrant High-Growth Entrepreneurs. *Economic Development Quarterly*, 16(1):20–31. Publisher: SAGE Publications Inc.

Saxenian, A. (2005). From Brain Drain to Brain Circulation: Transnational Communities and Regional Upgrading in India and China. *Studies in Comparative International Development*, 40(2):35–61.

Saxenian, A. (2007). *The New Argonauts: Regional Advantage in a Global Economy*. Harvard University Press, Cambridge, Mass., first thus edition edition.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.

Vakili, K. and McGahan, A. M. (2016). Health Care's Grand Challenge: Stimulating Basic Science on Diseases That Primarily Afflict the Poor. *The Academy of Management Journal*, 59(6):1917–1939. Publisher: Academy of Management.

Wang, D. (2015). Activating Cross-border Brokerage: Interorganizational Knowledge Transfer through Skilled Return Migration. *Administrative Science Quarterly*, 60(1):133–176.

Xie, Q. and Freeman, R. B. (2023). Creating and Connecting US and China Science: Chinese Diaspora and Returnee Researchers.

Figures and Tables

Figure 1: Home and host country overlap

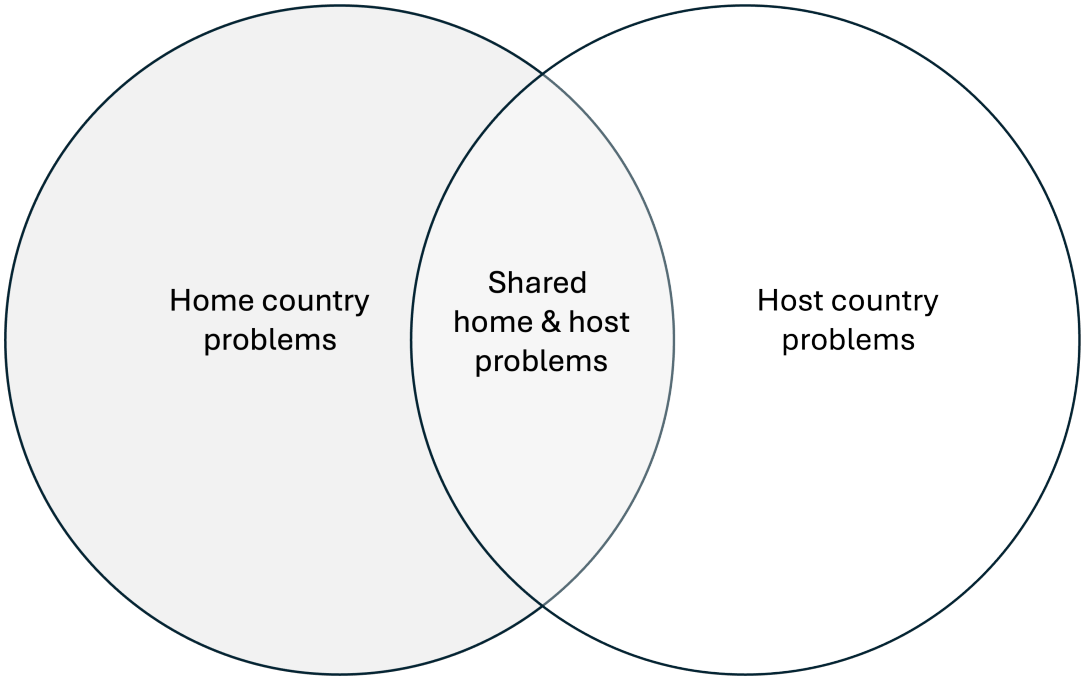
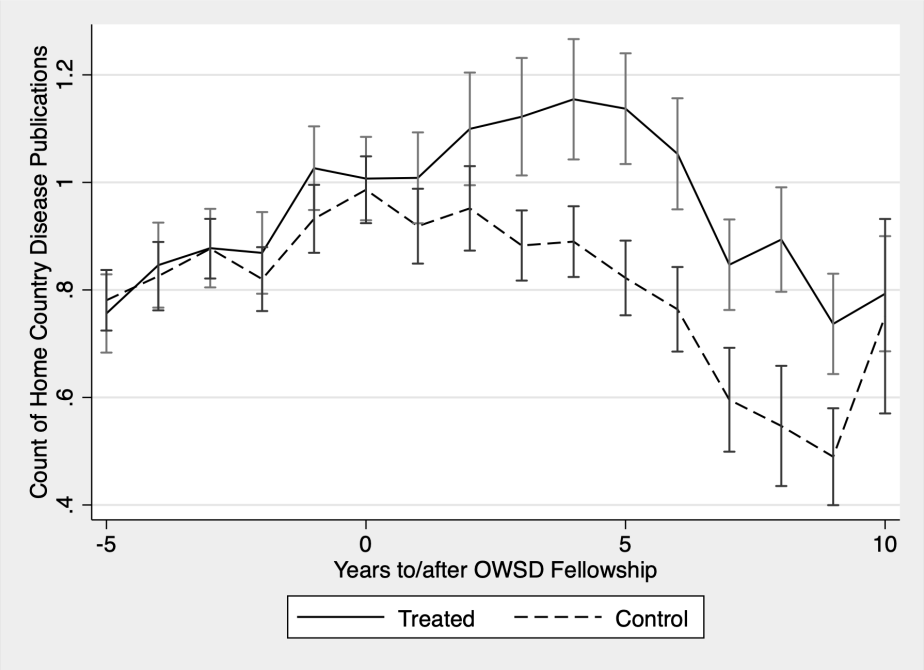
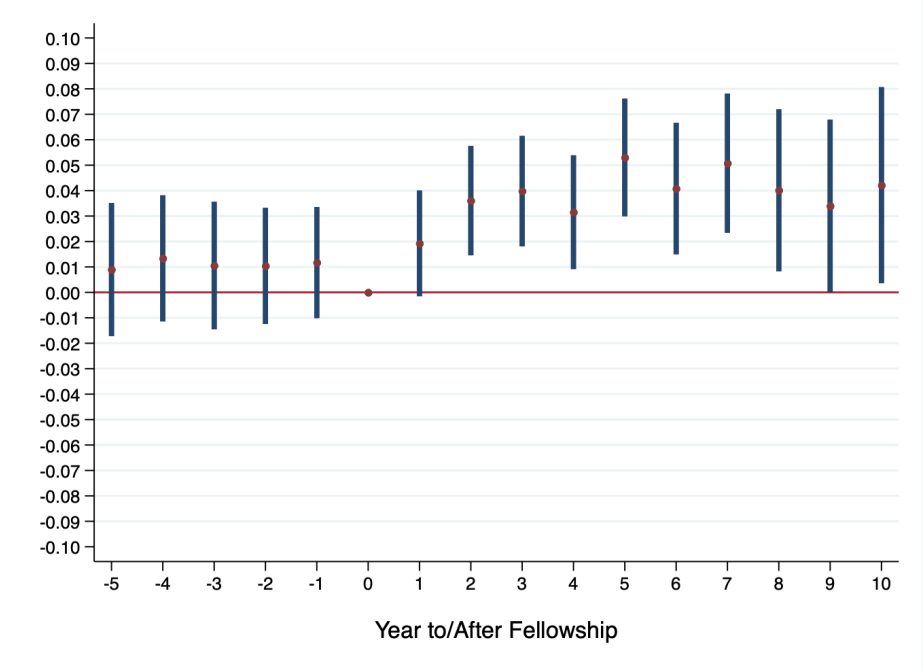


Figure 2: Changes in Home Country Disease Publication Rates



Notes: This figure plots raw trends of average publication rates on home-country diseases (per non-migrant peer scientist) for the five years before and ten years after the fellowship.

Figure 3: Differences in publication rate into home country diseases between treated and untreated non-migrant peer scientists



Notes: This figure shows coefficient estimates stemming from conditional (scientist) fixed-effects OLS specifications. The inverse hyperbolic sine outcomes are regressed onto year effects and scientist age effects, and include interaction terms between treatment status and the number of years before/after the OWSD fellowship, and a control variable representing the annual overall publication count for each scientist. The 90% confidence interval robust standard errors clustered around the scientist is plotted with solid bars.

Table 1: Summary statistics for non-migrant peer scientists (at year of application)

	(1) Disease Control Sample mean/sd	(2) Disease Treatment Sample mean/sd	(3) Difference Between Groups b	t
Female	0.30	0.25	0.05	3.19
	(0.46)	(0.43)		
Career Age	10.40	9.15	1.25	8.33
	(7.24)	(6.74)		
Total publications	11.92	11.06	0.86	2.79
	(14.88)	(14.05)		
Any disease publications	0.77	0.71	0.06	6.66
	(0.42)	(0.46)		
Total disease publications	5.12	5.12	-0.00	-0.01
	(10.18)	(10.83)		
Any home disease publications	0.73	0.67	0.06	5.73
	(0.44)	(0.47)		
Total home disease publications	4.59	4.60	-0.01	-0.02
	(9.43)	(9.97)		
Any host disease publications	0.62	0.54	0.08	7.41
	(0.49)	(0.50)		
Total host disease publications	3.36	3.38	-0.02	-0.10
	(7.79)	(8.70)		
Any home and host disease publications	0.61	0.53	0.07	6.64
	(0.49)	(0.50)		
Total home and host disease publications	3.25	3.29	-0.04	-0.24
	(7.57)	(8.50)		
Total home only disease publications	1.34	1.31	0.04	0.44
	(4.21)	(3.57)		
Total host only disease publications	0.11	0.09	0.02	1.55
	(0.58)	(0.77)		
Any collaborative publications with host country researchers	0.05	0.07	-0.02	-4.03
	(0.21)	(0.25)		
Total count of collaborative publications with host country researchers	0.09	0.11	-0.03	-2.20
	(0.54)	(0.59)		
High disease overlap between home and host country	0.27	0.09	0.18	23.82
	(0.44)	(0.28)		
Home and host country in same continent	0.43	0.63	-0.21	-19.56
	(0.50)	(0.48)		
Home and host country in same income band	0.32	0.11	0.21	25.65
	(0.47)	(0.32)		
Observations	5256	3667	8923	

Notes: This table depicts summary statistics for our treatment and control samples. All productivity statistics are generated for the 5 years prior to treatment. Columns 1 and 2 show averages and standard errors for our control and treatment groups, respectively. Column 3 shows the raw difference in means between control and treatment groups, and column 4 shows the t-value resulting from a t-test for the treatment and control samples. The t-test compares the control and treatment group means and assumes a null hypothesis that the two means are equal. If the t-statistic is greater than 1.96, we can reject the null hypothesis that there is no difference between means at a 5% level of significance.

Table 2: Changes in overall publication output of non-migrating peers after OWSD fellowship

	Number of Publications	Number of Disease Publications	Number of Non-Disease Publications	Number of NTD Disease Publications	Number of Non-NTD Disease Publications
	(1)	(2)	(3)	(4)	(5)
After Fellowship × Successful Applicant	0.0634***	0.0223***	-0.0125**	0.0212***	0.0058
	(0.015)	(0.007)	(0.006)	(0.006)	(0.006)
Total Observations	107576	107576	107576	107576	107576
Mean of Dep. Variable	2.2171	0.7643	1.4528	0.4317	0.3326
Author FE	X	X	X	X	X
Year FE	X	X	X	X	X
Career Age FE	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X
Annual Pubs		X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed-effects OLS regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts. [b] Heteroskedastic robust standard errors, clustered at the level of the scientist, are given in parentheses.

Table 3: Changes in disease topic publications of non-migrating peers after OWSD fellowship

	Home Country Diseases	Host Country Diseases	Home and Host Country Diseases	Home not Host Country Diseases	Host not Home Country Diseases
	(1)	(2)	(3)	(4)	(5)
DV: Number of publications					
After Fellowship × Successful Applicant	0.0285***	0.0217***	0.0210***	0.0150***	-0.0004
	(0.008)	(0.007)	(0.007)	(0.005)	(0.002)
Total Observations	107576	107576	107576	107576	107576
Mean of Dep. Variable	0.9111	0.6835	0.6636	0.2474	0.0198
Author FE	X	X	X	X	X
Year FE	X	X	X	X	X
Career Age FE	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X
Annual Pubs	X	X	X	X	X
DV: Proportion of publications					
After Fellowship × Successful Applicant	0.0201***	0.0136***	0.0132***	0.0078***	0.0001
	(0.005)	(0.004)	(0.004)	(0.003)	(0.001)
Total Observations	107576	107576	107576	107576	107576
Mean of Dep. Variable	0.2595	0.1863	0.1810	0.0786	0.0053
Author FE	X	X	X	X	X
Year FE	X	X	X	X	X
Career Age FE	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed-effects OLS regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts.[b] Heteroskedastic robust standard errors, clustered at the level of the scientist, are given in parentheses.

Table 4: Disease topic publications of non-migrants

	Number of Publications	Number of Home Country Publications	Number of Home and Host Country Publications
Panel A: Similar Home and Host Income	(1)	(2)	(3)
After Fellowship x	0.0857***	0.0288***	0.0200**
Successful Applicant	(0.017)	(0.009)	(0.008)
After Fellowship x	-0.0246	0.0193	0.0356*
Successful Applicant x Similar Income	(0.039)	(0.021)	(0.021)
Total Observations	107576	107576	107576
Panel B: Shared Home and Host Continent	(1)	(2)	(3)
After Fellowship x	0.0723***	0.0118	0.0059
Successful Applicant	(0.020)	(0.011)	(0.011)
After Fellowship x	0.0083	0.0264*	0.0285*
Successful Applicant x Same Continent	(0.029)	(0.015)	(0.015)
Total Observations	107576	107576	107576
Panel C: High Home and Host Disease Overlap	(1)	(2)	(3)
After Fellowship x	0.0854***	0.0264***	0.0144*
Successful Applicant	(0.016)	(0.008)	(0.008)
After Fellowship x	0.0088	0.0193	0.0544**
Successful Applicant x High Disease Overlap	(0.043)	(0.023)	(0.024)
Total Observations	107576	107576	107576
Mean of Dep. Variable	1.0710	0.4681	0.3513
Author FE	X	X	X
Year FE	X	X	X
Career Age FE	X	X	X
Time Since Grant FE	X	X	X
Annual Pubs		X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed-effects OLS regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts. [b] Heteroskedastic robust standard errors, clustered at the level of the scientist, are given in parentheses.

Table 5: Impact of fellow's research focus on non-migrating peer's research trajectory

	Number of Home Country Disease Publications (fellow sample)		Number of Home Country Disease Publications (non-migrant peer sample fellow not home country disease researcher)	Number of Home Country Disease Publications (non-migrant peer sample fellow home country disease researcher)	Number of publications with fellow (non-migrant peer sample)	Number of Home Country Disease Publications (non-migrant peer sample excluding those who ever coauthor with fellow)
	(1)	(2)	(3)	(4)	(5)	(6)
After Fellowship × Successful Applicant	0.1051	0.0086	0.0307***	0.0318***	0.0004	0.0289***
	(0.108)	(0.080)	(0.011)	(0.012)	(0.000)	(0.008)
Total Observations	240	240	69606	37970	107576	107178
Mean of Dep. Variable	0.4344	0.4344	0.9379	0.8620	0.0005	0.9064
Author FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Career Age FE	X	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X	X
Annual Pubs		X	X	X		X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed-effects OLS regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts. [b] Heteroskedastic robust standard errors, clustered at the level of the scientist, are given in parentheses.

Table 6: Changes in host country collaborations of non-migrating peers after OWSD fellowship

	Number of Publications Coauthored with Host Institution	Number of Publications Coauthored with Host Country	Number of Home Country Disease Publications Coauthored with Host Country	Number of Non-Home Country Disease Publications Coauthored with Host Country	Number of Home and Host Country Disease Publications Coauthored with Host Country	Number of Home not Host Country Disease Publications Coauthored with Host Country
	(1)	(2)	(3)	(4)	(5)	(6)
After Fellowship × Successful Applicant	0.0034**	0.0119***	0.0117***	0.0022	0.0116***	0.0004
	(0.001)	(0.004)	(0.003)	(0.003)	(0.003)	(0.000)
Total Observations	107576	107576	107576	107576	107576	107576
Mean of Dep. Variable	0.0058	0.0311	0.0307	0.0298	0.0307	0.0307
Author FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Career Age FE	X	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X	X
Annual Pubs	X	X	X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed-effects OLS regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts. [b] Heteroskedastic robust standard errors, clustered at the level of the scientist, are given in parentheses.

Table 7: Changes in use of host institution keywords of non-migrating peers after OWSD fellowship

	Proportion of Host Institution Keywords Used				Proportion of Host Institution Keywords Used in Home Country Disease Publications	Proportion of Host Institution Keywords Used in Home and Host Country Disease Publications
	(1)	(2)	(3)	(4)	(5)	(6)
After Fellowship × Successful Applicant	0.0062**	0.0061**	0.0055*	0.0056*	0.0068**	0.0066**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Total Observations	12254	12254	12254	12254	11872	11588
Mean of Dep. Variable	0.2779	0.2779	0.2779	0.2779	0.2759	0.2771
Author FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Career Age FE	X	X	X	X	X	X
Annual Pubs		X		X	X	X
Annual Keywords			X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed-effects OLS regression specifications in which dependent variables are inverse hyperbolic sine transformed counts of the number of host affiliation keywords in publications pre-migration event in the migrant's field of study that are used by non-migrating peers before and after the migration event, divided by the overall number of keywords in the non-migrating peers publications before and after the migration event. There is one observation per sample scientist (pre and post migration event) which combines all keywords used in publications in these two time periods. The sample includes only non-migrating peers for whom host institution publications could be obtained. [b] Heteroskedastic robust standard errors are given in parentheses.

Appendix A: Appendix Figures and Tables

Table A1: Home and Host Countries

	Home Country Statistics				Host Country Statistics			
	Successful applicants	Unsuccessful applicants	Treated peers	Control peers	Successful applicants	Unsuccessful applicants	Treated peers	Control peers
Argentina	0	0	0	0	0	1	0	97
Bangladesh	6	2	117	97	0	0	0	0
Benin	3	0	112	0	0	3	0	175
Botswana	0	0	0	0	2	1	16	10
Brazil	0	0	0	0	3	4	51	44
China	0	0	0	0	9	21	1018	959
Cameroon	11	19	805	568	0	0	0	0
Egypt	0	0	0	0	0	3	0	33
Eswatini	0	2	0	2	0	0	0	0
Ethiopia	6	2	252	57	0	0	0	0
Ghana	6	1	104	25	2	1	59	58
India	0	0	0	0	6	7	129	523
Kenya	7	7	48	386	2	2	79	93
Madagascar	1	1	30	1	0	0	0	0
Malawi	1	1	2	1	0	0	0	0
Malaysia	0	0	0	0	6	9	47	781
Morocco	0	0	0	0	0	1	0	2
Nepal	1	1	2	5	0	0	0	0
Nigeria	24	59	1673	3606	1	1	8	59
Pakistan	0	0	0	0	2	1	27	495
Philippines	0	0	0	0	1	1	84	44
Rwanda	0	1	0	1	0	0	0	0
Senegal	2	0	29	0	2	0	71	0
Singapore	0	0	0	0	0	2	0	9
South Africa	0	1	0	404	48	50	2054	1785
Sudan	5	5	34	48	0	0	0	0
Tanzania	2	1	54	3	1	1	5	1
Thailand	0	0	0	0	2	0	17	0
Uganda	3	4	88	33	1	1	2	88
Yemen	1	1	5	1	0	0	0	0
Zambia	2	0	9	0	0	0	0	0
Zimbabwe	7	2	303	18	0	0	0	0

Notes: This table shows the frequency of OWSD applicants and peers from our set of home and host countries. The columns under 'Home Country Statistics' describe the number of successful applicants (fellows), unsuccessful applicants, treated peers (scientists from fellows' home institutions) and control peers (scientists from unsuccessful applicants' home institutions) from each listed country. The columns under 'Host Country Statistics' describe the number of successful applicants (fellows), unsuccessful applicants, treated peers (scientists from fellows' home institutions) and control peers (scientists from unsuccessful applicants' home institutions) associated with each host country. Zero observations indicate there was no individual in this category with this host or home country.

Table A2: Controlling for Publication Quality

	Home Country Publications	Host Country Publications	Home and Host Country Publications	Home not Host Country Publications	Host not Home Country Publications
	(1)	(2)	(3)	(4)	(5)
DV: sum source normalized impact factor					
After Fellowship × Successful Applicant	0.0538***	0.0467***	0.0461***	0.0183***	0.0007
	(0.014)	(0.014)	(0.014)	(0.005)	(0.001)
Total Observations	107576	107576	107576	107576	107576
Mean of Dep. Variable	1.9799	1.7937	1.7850	0.1949	0.1949
DV: mean source normalized impact factor					
After Fellowship × Successful Applicant	0.0175***	0.0165**	0.0159**	0.0102***	0.0008
	(0.006)	(0.007)	(0.007)	(0.003)	(0.001)
Total Observations	107576	107576	107576	107576	107576
Mean of Dep. Variable	0.5258	0.4949	0.4943	0.1001	0.1001
Author FE	X	X	X	X	X
Year FE	X	X	X	X	X
Career Age FE	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed-effects OLS regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts. [b] Heteroskedastic robust standard errors, clustered at the level of the scientist, are given in parentheses. We measure publication quality in two ways: 1) In the first panel, we use the sum of SNIP (Source Normalized Impact per Paper) impact factor of the individual's annual publications. 2) In panel 2, we use the mean impact factor for an individual's annual publications.

Table A3: Alternative High Home Prevalence Measures

	Number of Publications on Diseases Above Median Prevalence	Number of Publications on Diseases With Any Prevalence	Number of Publications on Diseases with Above 1 Percent Prevalence	Number of Publications on Diseases with Above 0.1 Percent Prevalence
	(1)	(2)	(3)	(4)
After Fellowship × Successful Applicant	0.0285***	0.0233***	0.0176**	0.0283***
	(0.008)	(0.008)	(0.007)	(0.008)
Total Observations	107576	107576	107576	107576
Mean of Dep. Variable	0.9111	1.0070	0.5166	0.8986
Author FE	X	X	X	X
Year FE	X	X	X	X
Career Age FE	X	X	X	X
Time Since Grant FE	X	X	X	X
Annual Pubs	X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed-effects OLS regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts. [b] Heteroskedastic robust standard errors, clustered at the level of the scientist, are given in parentheses.

Table A4: Changes in NTD disease topic publications of non-migrating peers after OWSD fellowship

	Number of High Prevalence Home Country NTD Disease Publications	Number of Any Prevalence Home Country NTD Disease Publications	Number of Publications on High Prevalence Home Country Non-NTD Diseases	Number of High Prevalence Home Country NTD Disease Publications Excl. HIV and TB	Number of High Prevalence Home Country NTD Disease Publications Excl. malaria
	(1)	(2)	(3)	(4)	(5)
After Fellowship × Successful Applicant	0.0280***	0.0251***	0.0089	0.0197***	0.0210***
	(0.007)	(0.007)	(0.007)	(0.005)	(0.006)
Total Observations	107576	107576	107576	107576	107576
Mean of Dep. Variable	0.4537	0.4881	0.4573	0.2421	0.2801
Author FE	X	X	X	X	X
Year FE	X	X	X	X	X
Career Age FE	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X
Annual Pubs	X	X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed-effects OLS regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts. [b] Heteroskedastic robust standard errors, clustered at the level of the scientist, are given in parentheses. As described in text, NTDs are Neglected Tropical Diseases that are primarily present in low-income and tropical areas that affect more than 1 billion people worldwide (https://www.who.int/news-room/questions-and-answers/item/neglected-tropical-diseases last accessed 8.23.24).

Table A5: Summary statistics for non-migrant peer scientists (at year of application)

Panel A: Similar Home and Host Income	(1) Disease Control Sample mean/sd	(2) Disease Treatment Sample mean/sd	(3) Difference Between Groups b	t
Total publications	13.93	8.20	5.73	6.49
	(16.52)	(9.88)		
Total disease publications	4.69	4.00	0.69	1.12
	(9.99)	(7.24)		
Total home disease publications	4.06	3.55	0.50	0.93
	(8.94)	(6.37)		
Total home and host disease publications	3.69	3.17	0.51	0.99
	(8.50)	(6.11)		
Any collaborative publications with host country researchers	0.03	0.01	0.03	3.37
	(0.17)	(0.07)		
Total count of collaborative publications with host country researchers	0.05	0.02	0.04	1.77
	(0.38)	(0.22)		
Total count of host institution publications on home & host topics	5114.97	629.89	4485.08	34.62
	(4219.40)	(190.49)		
Observations	1073	186	1259	

Panel B: Shared Home and Host Continent	(1) Disease Control Sample mean/sd	(2) Disease Treatment Sample mean/sd	(3) Difference Between Groups b	t
Total publications	7.48	8.11	-0.63	-0.52
	(10.43)	(13.68)		
Total disease publications	1.62	2.82	-1.20	-2.54
	(3.01)	(6.29)		
Total home disease publications	1.51	2.44	-0.93	-2.22
	(2.84)	(5.46)		
Total home and host disease publications	1.37	2.15	-0.79	-1.98
	(2.66)	(5.20)		
Any collaborative publications with host country researchers	0.02	0.02	-0.00	-0.09
	(0.14)	(0.14)		
Total count of collaborative publications with host country researchers	0.02	0.08	-0.06	-1.31
	(0.14)	(0.73)		
Total count of host institution publications on home & host topics	961.61	584.43	377.19	16.40
	(239.23)	(192.36)		
Observations	153	239	392	

Panel C: High Home and Host Disease Overlap	(1) Disease Control Sample mean/sd	(2) Disease Treatment Sample mean/sd	(3) Difference Between Groups b	t
Total publications	13.04	8.53	4.50	4.93
	(15.42)	(14.40)		
Total disease publications	5.13	2.96	2.17	4.92
	(10.12)	(6.13)		
Total home disease publications	4.46	2.63	1.83	4.67
	(9.09)	(5.43)		
Total home and host disease publications	3.91	2.37	1.54	4.13
	(8.51)	(5.22)		
Any collaborative publications with host country researchers	0.02	0.02	0.01	0.64
	(0.16)	(0.14)		
Total count of collaborative publications with host country researchers	0.04	0.07	-0.03	-0.86
	(0.34)	(0.66)		
Total count of host institution publications on home & host topics	4611.59	562.50	4049.09	39.29
	(3836.17)	(212.52)		
Observations	1405	312	1717	

Notes: This table depicts summary statistics for our treatment and control samples. All productivity statistics are generated for the 5 years prior to treatment. Columns 1 and 2 show averages and standard errors for our control and treatment groups, respectively. Column 3 shows the raw difference in means between control and treatment groups, and column 4 shows the t-value resulting from a t-test for the treatment and control samples. The t-test compares the control and treatment group means and assumes a null hypothesis that the two means are equal. If the t-statistic is greater than 1.96, we can reject the null hypothesis that there is no difference between means at a 5% level of significance.

Table A6: Heterogeneous changes in host country collaborations of non-migrating peers after OWSD fellowship

	Number of Publications Coauthored with Host Institution	Number of Publications Coauthored with Host Country	Number of Home Country Publications Coauthored with Host Country	Number of Non-Home Country Publications Coauthored with Host Country	Number of Home and Host Country Publications Coauthored with Host Country	Number of Home not Host Country Publications Coauthored with Host Country
Panel A: Similar Home and Host Income	(1)	(2)	(3)	(4)	(5)	(6)
After Fellowship x	0.0033**	0.0113***	0.0113***	0.0032	0.0114***	0.0002
Successful Applicant	(0.002)	(0.004)	(0.004)	(0.003)	(0.004)	(0.001)
After Fellowship x	-0.0023	-0.0084*	-0.0099*	-0.0153***	-0.0098*	-0.0002
Successful Applicant x Similar Income Host	(0.003)	(0.005)	(0.005)	(0.005)	(0.005)	(0.001)
Total Observations	107576	107576	107576	107576	107576	107576
Panel B: Shared Home and Host Continent	(1)	(2)	(3)	(4)	(5)	(6)
After Fellowship x	-0.0017*	0.0023	0.0009	-0.0045*	0.0006	0.0002
Successful Applicant	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.000)
After Fellowship x	0.0080***	0.0136**	0.0158***	0.0092*	0.0163***	0.0001
Successful Applicant x Same Continent	(0.003)	(0.006)	(0.006)	(0.005)	(0.006)	(0.001)
Total Observations	107576	107576	107576	107576	107576	107576
Panel C: High Home and Host Disease Overlap	(1)	(2)	(3)	(4)	(5)	(6)
After Fellowship x	0.0029**	0.0111***	0.0111***	0.0013	0.0110***	0.0003
Successful Applicant	(0.001)	(0.004)	(0.004)	(0.003)	(0.004)	(0.001)
After Fellowship x	0.0128	-0.0032	-0.0042	-0.0000	-0.0044	-0.0001
Successful Applicant x High Disease Overlap	(0.011)	(0.006)	(0.006)	(0.007)	(0.006)	(0.001)
Total Observations	107576	107576	107576	107576	107576	107576
Mean of Dep. Variable	0.0058	0.0311	0.0307	0.0014	0.0293	0.0014
Author FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Career Age FE	X	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X	X
Annual Pubs		X	X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed-effects OLS regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts. [b] Heteroskedastic robust standard errors, clustered at the level of the scientist, are given in parentheses.

Table A7: Changes in disease topic publications of non-migrating peers after OWSD fellowship at the peer, year disease level

	Number of publications				
	(1)	(2)	(3)	(4)	(5)
After Fellowship × Successful Applicant × High Home Prevalence Disease	0.0011***				
	(0.0002)				
After Fellowship × Successful Applicant × High Host Prevalence Disease		0.0008***			
		(0.0002)			
After Fellowship × Successful Applicant × High Prevalence Home and Host Disease			0.0010***		
			(0.0003)		
After Fellowship × Successful Applicant × High Prevalence Home Not Host Disease				0.0003	
				(0.0004)	
After Fellowship × Successful Applicant × High Prevalence Host Not Home Disease					-0.0005**
					(0.0002)
Total Observations	16781856	16781856	16781856	16781856	16781856
Mean of Dep. Variable	0.0065	0.0065	0.0065	0.0065	0.0065
Author Year FE	X	X	X	X	X
Disease Year FE	X	X	X	X	X
Author Disease FE	X	X	X	X	X
Career Age FE	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed-effects OLS regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts of publications in a given disease of the specific type (e.g. high home prevalence) in a given year for the focal scientist. [b] Heteroskedastic robust standard errors, clustered at the level of the scientist, are given in parentheses.

Table A8: Accounting for Existing Home and Host Similarity

	Number of Home and Host Country Publications (excl prior host country connections)	Number of Home and Host Country Publications (low home institution prior home and host pubs)	Number of Home and Host Country Publications (high home institution prior home and host pubs)
	(1)	(2)	(3)
After Fellowship × Successful Applicant	0.0211***	0.0395***	-0.0023
	(0.007)	(0.011)	(0.010)
Total Observations	101686	53563	54013
Mean of Dep. Variable	0.6147	0.4842	0.8416
Author FE	X	X	X
Year FE	X	X	X
Career Age FE	X	X	X
Time Since Grant FE	X	X	X
Annual Pubs	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed-effects OLS regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts. [b] Heteroskedastic robust standard errors, clustered at the level of the scientist, are given in parentheses.

Table A9: Robust to Excluding South African Host Countries

	Number of Home Country Publications	Number of Host Country Publications	Number of Home and Host Country Publications	Number of Home not Host Country Publications	Number of Host not Home Country Publications
	(1)	(2)	(3)	(4)	(5)
After Fellowship × Successful Applicant	0.0252**	0.0243**	0.0212**	0.0190**	0.0031
	(0.011)	(0.010)	(0.010)	(0.008)	(0.003)
Total Observations	59743	59743	59743	59743	59743
Mean of Dep. Variable	0.9544	0.7632	0.7314	0.2229	0.0317
Author FE	X	X	X	X	X
Year FE	X	X	X	X	X
Career Age FE	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X
Annual Pubs	X	X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed-effects OLS regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts. [b] Heteroskedastic robust standard errors, clustered at the level of the scientist, are given in parentheses. Because 40% of our OWSD applicants identified South African host institutions, we re-run our main analyses excluding non-migrants whose fellow colleague selected an institution in South Africa to ensure that our results are not driven only by one host country.

Table A10: Sandwich fellows versus non-sandwich fellows

	Number of Publications (not sandwich fellowship)	Number of Home Country Publications (not sandwich fellowship)	Number of Home and Host Country Publications (not sandwich fellowship)	Number of Publications (sandwich fellowship)	Number of Home Country Publications (sandwich fellowship)	Number of Home and Host Country Publications (sandwich fellowship)
	(1)	(2)	(3)	(4)	(5)	(6)
After Fellowship × Successful Applicant	0.0922***	0.0422***	0.0299**	0.0537***	0.0162*	0.0102
	(0.028)	(0.015)	(0.014)	(0.018)	(0.010)	(0.009)
Total Observations	42740	42740	42740	64836	64836	64836
Mean of Dep. Variable	2.1810	0.8789	0.3336	2.2409	0.9323	0.3629
Author FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Career Age FE	X	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X	X
Annual Pubs		X	X		X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed-effects OLS regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts. [b] Heteroskedastic robust standard errors, clustered at the level of the scientist, are given in parentheses.

Table A11: Effect of non-migrating peer's gender on changes in research trajectory

	Number of Publications (female non-migrant)	Number of Home Country Publications (female non-migrant)	Number of Home and Host Country Publications (female non-migrant)	Number of Publications (male non-migrant)	Number of Home Country Publications (male non-migrant)	Number of Home and Host Country Publications (male non-migrant)
	(1)	(2)	(3)	(4)	(5)	(6)
After Fellowship × Successful Applicant	0.0510	0.0703***	0.0708***	0.0514*	0.0244	0.0165
	(0.041)	(0.026)	(0.027)	(0.029)	(0.016)	(0.015)
Total Observations	10534	10534	10534	28255	28255	28255
Mean of Dep. Variable	1.9713	0.9684	0.7033	2.5306	1.0041	0.7418
Author FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Career Age FE	X	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X	X
Annual Pubs		X	X		X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed-effects OLS regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts. [b] Heteroskedastic robust standard errors, clustered at the level of the scientist, are given in parentheses.

Table A12: Just High Prevalence Diseases

	Number of publications				
	(1)	(2)	(3)	(4)	(5)
After Fellowship × Successful Applicant × High Prevalence Home and Host Disease	0.0010***				
	(0.000)				
After Fellowship × Successful Applicant × High Prevalence Home and Host Disease (high global prevalence disease)		0.0009***			
		(0.000)			
After Fellowship × Successful Applicant × High Prevalence Home not Host Disease (high global prevalence disease)			0.0004		
			(0.000)		
After Fellowship × Successful Applicant × High Prevalence Host not Home Disease (high global prevalence disease)				-0.0005*	
				(0.000)	
After Fellowship × Successful Applicant × High Prevalence Not Home Disease (high global prevalence disease)					-0.0008***
					(0.000)
Total Observations	16781856	16781856	16781856	16781856	16781856
Mean of Dep. Variable	0.0065	0.0065	0.0065	0.0065	0.0065
Author Year FE	X	X	X	X	X
Disease Year FE	X	X	X	X	X
Author Disease FE	X	X	X	X	X
Career Age FE	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed-effects OLS regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts. High global prevalence diseases are those with above median global prevalence. [b] Heteroskedastic robust standard errors, clustered at the level of the scientist, are given in parentheses.

Table A13: Placebo Host Country

	Number of publications			
	(1)	(2)	(3)	(4)
After Fellowship × Successful Applicant × High Prevalence Home and Host Disease	0.0006**	-0.0002		
	(0.000)	(0.001)		
After Fellowship × Successful Applicant × High Prevalence Home not Host Disease			0.0009***	
			(0.000)	
After Fellowship × Successful Applicant × High Prevalence Host not Home				-0.0004***
				(0.000)
Total Observations	16781856	2295540	16781856	16781856
Mean of Dep. Variable	0.0065	0.0061	0.0065	0.0065
Author Year FE	X	X	X	X
Disease Year FE	X	X	X	X
Author Disease FE	X	X	X	X
Career Age FE	X	X	X	X
Time Since Grant FE	X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed-effects OLS regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts. Each sample scientist is randomly assigned a placebo host country and home and host diseases are identified using the placebo host. In column 2 we drop person, disease, year observations for whom the placebo host disease prevalence is identical to the actual host disease prevalence. [b] Heteroskedastic robust standard errors, clustered at the level of the scientist, are given in parentheses.

Table A14: Excluding Migrating 'Non-Migrants'

	Number of Home Country Publications	Number of Host Country Publications	Number of Home and Host Country Publications	Number of Home not Host Country Publications	Number of Host not Home Country Publications
	(1)	(2)	(3)	(4)	(5)
After Fellowship × Successful Applicant	0.0220**	0.0147*	0.0141*	0.0113*	-0.0006
	(0.009)	(0.008)	(0.008)	(0.006)	(0.002)
Total Observations	81791	81791	81791	81791	81791
Mean of Dep. Variable	0.8643	0.6567	0.6362	0.2281	0.0205
Author FE	X	X	X	X	X
Year FE	X	X	X	X	X
Career Age FE	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X
Annual Pubs	X	X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed-effects OLS regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts. [b] Heteroskedastic robust standard errors, clustered at the level of the scientist, are given in parentheses.

Table A15: Exploring Strength of Exposure

	Excl Treatment/ Control Overlap	Single Exposure	Two Exposures	Three or more Exposures	More than Five Exposures	First Country Pair	Second Country Pair	Three or more Country Pair	More than Five Country Pair
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
After Fellowship × Successful Applicant	0.0512***	0.0319**	0.0335*	0.0174	0.0305**	0.0556***	0.0443**	0.0108	-0.0079
	(0.011)	(0.013)	(0.018)	(0.011)	(0.014)	(0.019)	(0.022)	(0.010)	(0.011)
Total Observations	79997	39366	20124	48085	23008	22235	18065	67276	47685
Mean of Dep. Variable	0.8908	0.8795	0.7701	0.9959	1.1009	0.6791	1.0508	0.9502	0.9923
Author FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X	X	X
Career Age FE	X	X	X	X	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X	X	X	X	X
Annual Pubs	X	X	X	X	X	X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed-effects OLS regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts. [b] Heteroskedastic robust standard errors, clustered at the level of the scientist, are given in parentheses. Columns 1-5 of this table show our results excluding non-migrants who have been exposed to more than one OWSD applicant. Columns 6-9 evaluate if our effects vary on based on the strength of the connection between a host and home country pair as measured by the number of fellows who have traveled from a given home to a given host.

Table A16: Changes in crop and pest topic publications of non-migrating peers after OWSD fellowship

	Number of Home Country Publications	Number of Host Country Publications	Number of Home and Host Country Publications	Number of Home not Host Country Publications	Number of Host not Home Country Publications
	(1)	(2)	(3)	(4)	(5)
After Fellowship × Successful Applicant	0.0169**	0.0137**	0.0125**	0.0062	0.0012
	(0.007)	(0.007)	(0.006)	(0.004)	(0.004)
Total Observations	135979	135979	135979	135979	135979
Mean of Dep. Variable	0.3960	0.3821	0.2889	0.1071	0.0932
Author FE	X	X	X	X	X
Year FE	X	X	X	X	X
Career Age FE	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X
Annual Pubs	X	X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed-effects OLS regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts. [b] Heteroskedastic robust standard errors, clustered at the level of the scientist, are given in parentheses.

Table A17: Alternative Specification - Poisson Maximum Likelihood model

	Number of Home Country Publications		Number of Host Country Publications		Number of Home and Host Country Publications	Number of Home not Host Country Publications	Number of Host not Home Country Publications
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
After Fellowship × Successful Applicant	0.1788***	0.0207	0.1472***	-0.0015	-0.0002	0.0698**	-0.0834
	(0.036)	(0.019)	(0.040)	(0.025)	(0.025)	(0.033)	(0.106)
Total Observations	97326	97326	84932	84932	83855	57588	11665
Mean of Dep. Variable	0.9111	0.9111	0.6835	0.6835	0.6636	0.2474	0.0198
Author FE	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X
Career Age FE	X	X	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X	X	X
Annual Pubs		X		X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from poisson maximum likelihood regression specifications in which dependent variables are annual counts. [b] Heteroskedastic robust standard errors, clustered at the level of the scientist, are given in parentheses.

Table A18: Alternative Specification - no annual controls

	Number of Home Country Publications	Number of Host Country Publications	Number of Home and Host Country Publications	Number of Home not Host Country Publications	Number of Host not Home Country Publications
	(1)	(2)	(3)	(4)	(5)
After Fellowship × Successful Applicant	0.0574***	0.0441***	0.0429***	0.0255***	0.0007
	(0.010)	(0.009)	(0.009)	(0.006)	(0.002)
Total Observations	107576	107576	107576	107576	107576
Mean of Dep. Variable	0.9111	0.6835	0.6636	0.2474	0.0198
Author FE	X	X	X	X	X
Year FE	X	X	X	X	X
Career Age FE	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed-effects OLS regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts. [b] Heteroskedastic robust standard errors, clustered at the level of the scientist, are given in parentheses.

Appendix B: Alternative Measures and Empirical Approaches

B1 Alternative Measures: Crop and Pest Analysis

We validate our main results using data on crop and pest prevalence. In a manner identical to the disease identification, we use key word searches of the title and abstract of the publications to identify whether each publication covers a crop or pest topic. We next merge the set of all sample publications to data on country-specific crop/pest prevalence in using the Centre for Agriculture and Bioscience International's (CABI) Crop Protection Compendium (CPC). This database provides comprehensive information on crops, pests, and their interactions, and provides detailed information on the geographic distribution of all crops and pests globally. CABI provides a prevalence measure that takes the value of 1 if a crop or pest is present in a country and 0 otherwise, and so we extract that measure for the non-migrant's home/host country. In total, we have a sample of 7,551 crops and pests, 585 of which are published on by at least one individual in our sample. 22.63% of publications match to at least one of these crops or pests.

Because our focus is on identifying research productivity in local/global diseases and crops/pests, we constrict our sample to the set of peers who we can plausibly expect to publish in these topics. We limit our sample to peers who ever publish on *any* crop/pest (we describe the construction of our disease and crop/pest sample in the following section). This leaves us with 7,831 treated non-migrant scientists (those whose peers received an OWSD fellowship) and 11,639 control scientists (those whose peers applied for but did not receive an OWSD fellowship), of which 4,164 treated scientists and 6,383 control scientists ever publish in crops. This corresponds to 211 OWSD applicants, with a median OWSD applicant treating 35 stayers in their home institution. Our main analysis is conducted at the applicant $\times$ peer level.

We leverage our data on publication matches to diseases and crops/pests to generate

several key measures of local/global nature of the research. In our empirical analyses, we measure individual-level annual publication counts on our diseases and organisms of interest. In our primary analysis, we measure individual-level annual publication counts to evaluate the number of articles an individual publishes on crops/pest that are present in:

- The home country
- The host country[31]
- Both the home and host countries
- The home country but *not* the host country
- The host country but *not* the home country

We first focus on the full set of crops and pests that are present in our home and host countries. Table A16 shows these results. We see that stayers are more likely to publish on crops and pests that are present in their home countries and on those that are present in their peer's host countries. However, this effect is not significant when we focus only on crops and pests that are present in home or host countries (columns 6 and 7). Treated researchers are thus increase their focus on organisms that appear in *both* their home and their peers' host environments, but not those that only appear either at home or in the host country.

B1.1 Alternative DiD estimation

The basic setup of differences-in-differences analysis relies on comparing treated and untreated units with the assumption that their outcomes would have been similar in the absence of treatment. The standard approach to this analysis is a two-way fixed effects regression, which regresses the outcome of interest on a unit's treatment status and includes

[31] For migrant peers, the host country is where their peer who applies for the fellowship identifies as the country they will visit.

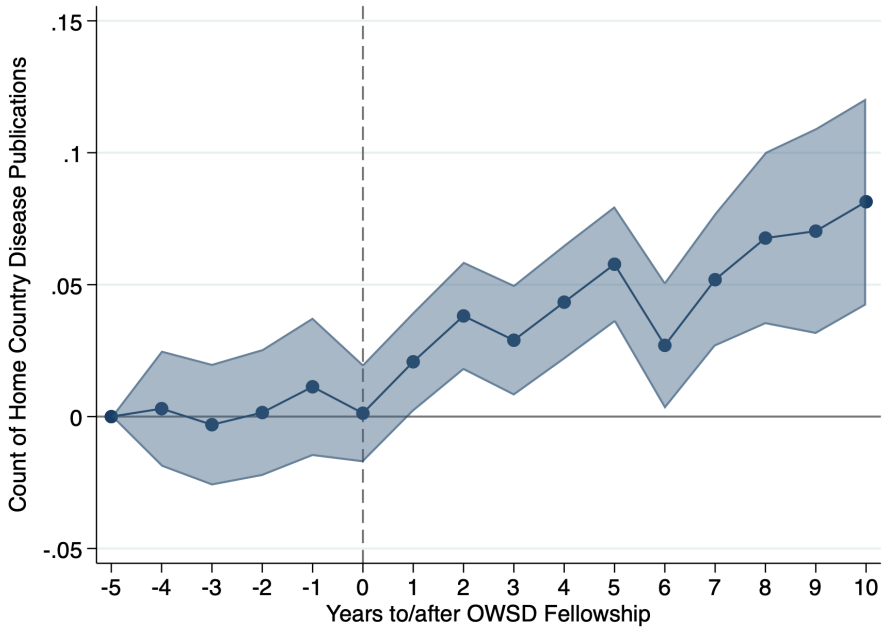
unit and time fixed effects. However, when treatment timing varies, this estimation approach compares already-treated units and later-treated units (a "forbidden comparison" (Goodman-Bacon, 2021)), which can negatively weight units that are treated earlier and lead to inaccurate estimates.

There are several novel estimation approaches that seek to address this issue and estimate causal parameters of interest in contexts with variation in treatment timing (de Chaisemartin and D'Haultfœuille, 2020; Callaway and Sant'Anna, 2021; Sun and Abraham, 2021; Borusyak et al., 2022; Gardner, 2022). These estimators vary in how they construct comparison groups and what assumptions they impose on parallel trends. The estimator proposed by Callaway and Sant'Anna (2021) does not require strongly balanced data (that is, for all units to be observed in all periods), but when estimating each group-time average treatment effect (ATTGT), only observations that are balanced within a specific 2x2 (two-period and two-group) are used for the estimator. Thus, when groups are not observed in all periods, estimation is conceptually more complicated. The estimator proposed by Sun and Abraham (2021) requires a strongly balanced panel, which we do not have.

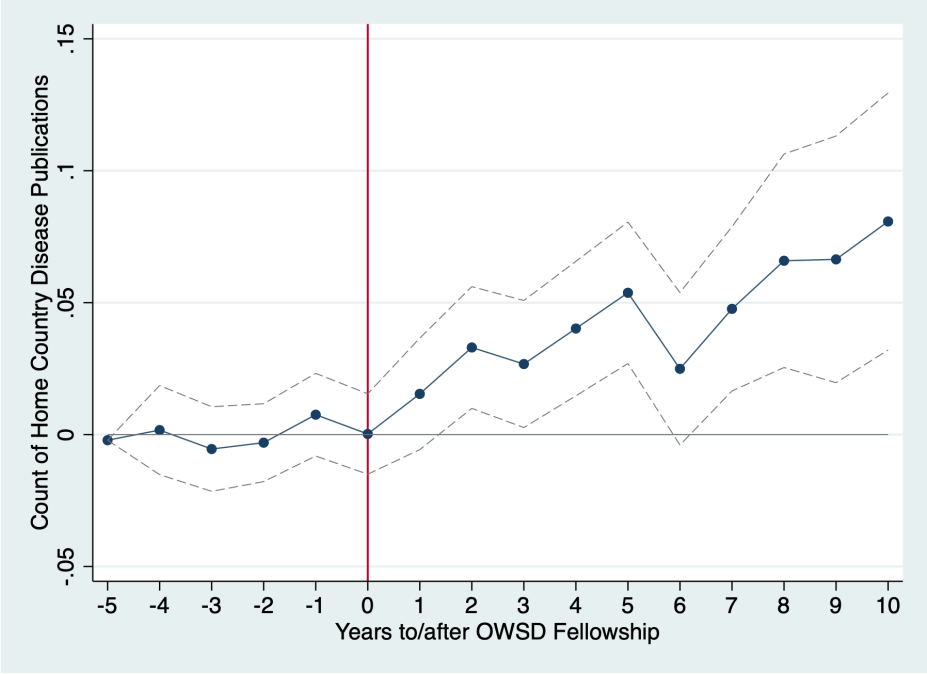
The imputation-based estimators developed in Borusyak et al. (2022) and Gardner (2022) fit a TWFE regression using observations only for units and time periods that are not-yet-treated. They then infer the never-treated potential outcome for each treated unit using the predicted value from this regression. This provides an estimate of the treatment effect for each treated unit, and these individual-level estimates are aggregated to create estimates of the average treatment effect on treated units (ATT) (Roth et al., 2023). Both estimators return identical point estimates but differ in their standard errors (Butts and Gardner, 2022). These approaches accommodate panels that are not strongly balanced and allow for event-study estimation; thus, these methods are best suited for our setting. Figures B1 and B2 below shows our pretend results for home country publications and home and host country publications using both Boryusak and Gardner estimation.

Figure B1: Alternative DiD Estimation Pretrends: Publications on Home Diseases

Panel A: Boryusak Estimation



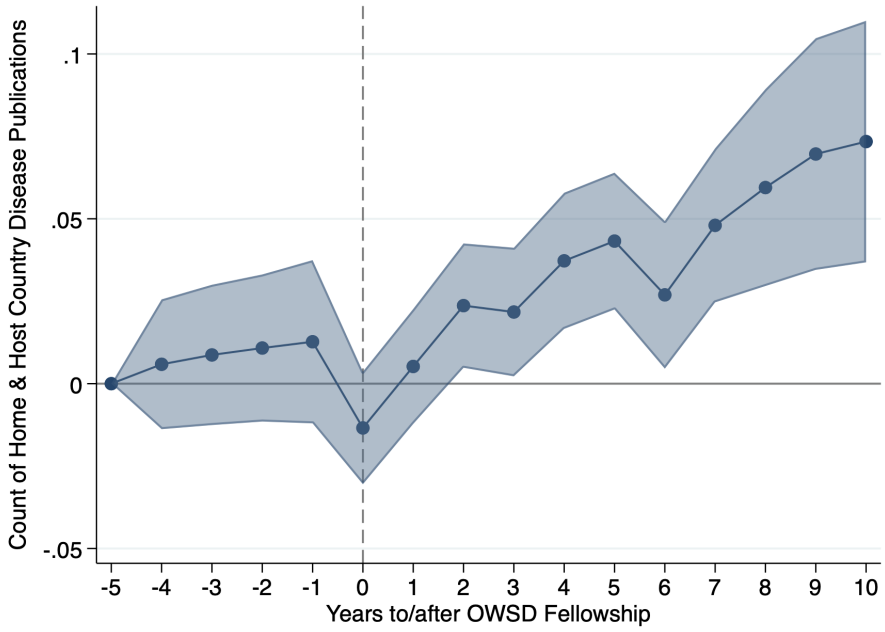
Panel B: Gardner Estimation



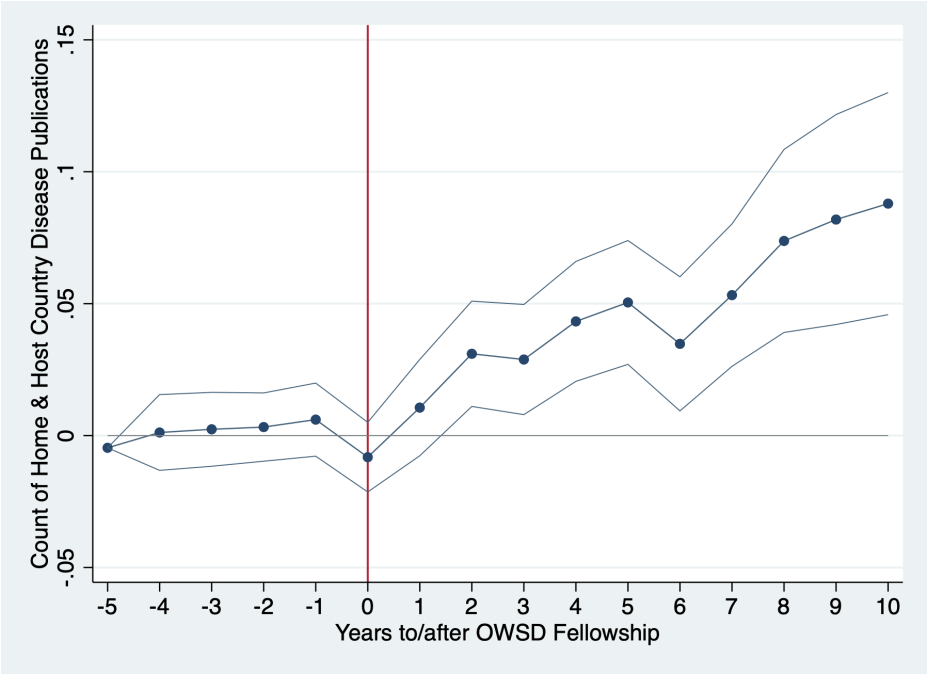
Notes: These figures shows coefficient estimates stemming from conditional (scientist) fixed-effects OLS specifications using the Boryusak estimation (Panel A) and Gardner estimation (Panel B) methods. The inverse hyperbolic sine outcomes are regressed onto year effects and scientist age effects, and include interaction terms between treatment status and the number of years before/after the OWSD fellowship, and a control variable representing the annual overall publication count for each scientist. The 95% confidence interval robust standard errors clustered around the scientist is plotted with solid bars.

Figure B2: Alternative DiD Estimation Pretrends: Publications on Home and Host Diseases

Panel A: Boryusak Estimation



Panel B: Gardner Estimation



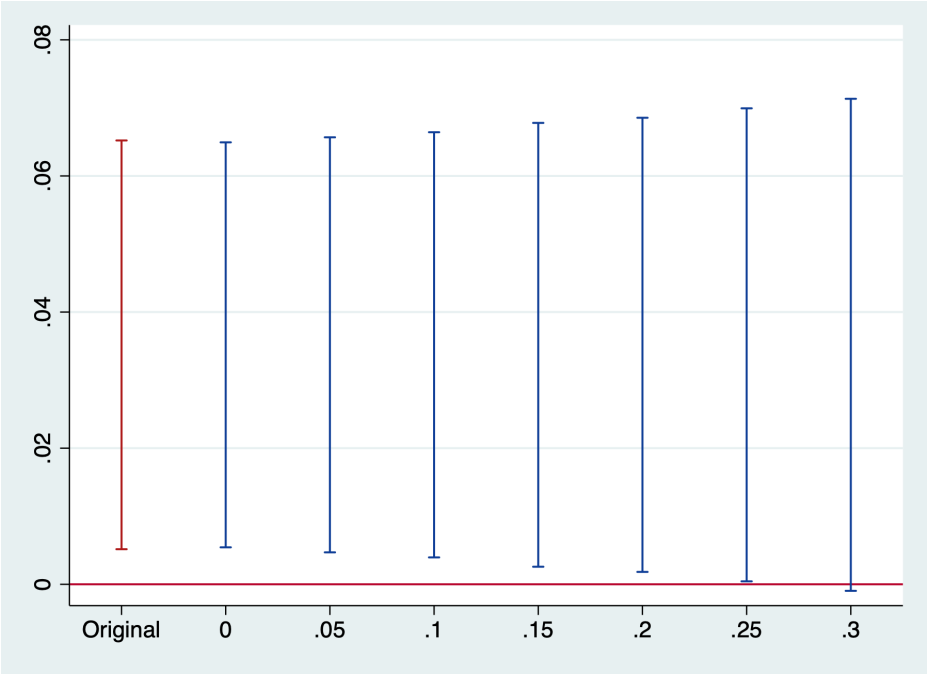
Notes: These figures shows coefficient estimates stemming from conditional (scientist) fixed-effects OLS specifications using the Boryusak estimation (Panel A) and Gardner estimation (Panel B) methods. The inverse hyperbolic sine outcomes are regressed onto year effects and scientist age effects, and include interaction terms between treatment status and the number of years before/after the OWSD fellowship, and a control variable representing the annual overall publication count for each scientist. The 95% confidence interval robust standard errors clustered around the scientist is plotted with solid bars.

B1.2 Parallel trend violations

A key assumption with differences-in-differences methods is the parallel trends assumption, which assumes that the average outcome for treated and untreated samples will evolve along parallel paths. Instead of requiring that parallel trends hold, Rambachan and Roth (2023) propose a robust inference approach. This approach measures the sensitivity of pre-trends by examining how results change under different assumptions about potential violations of parallel trends.

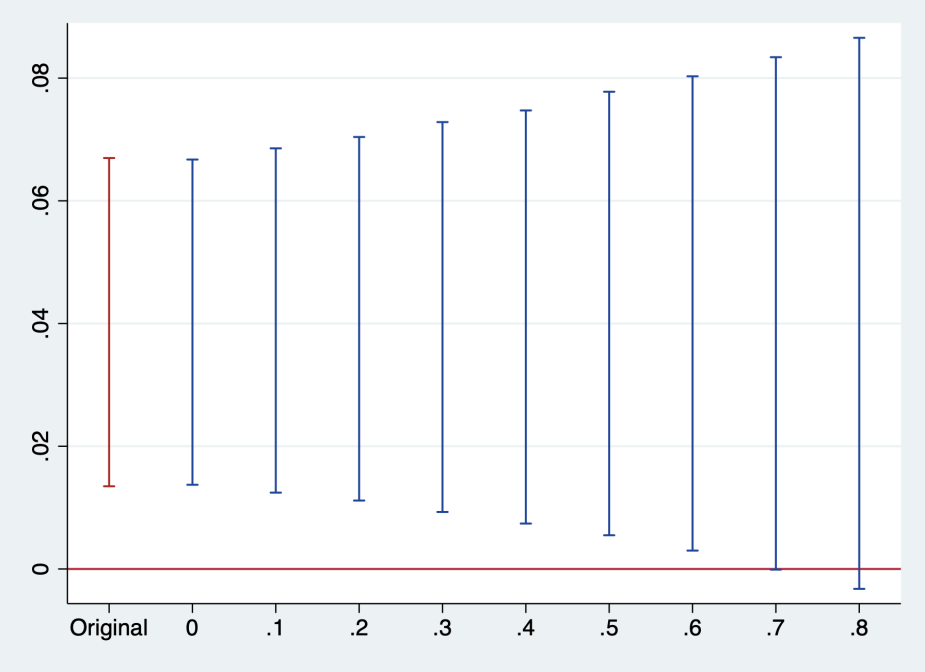
To confirm that our results are robust to potential differences in pre-trends, we adopt this inference approach and identify by how much the slope of pre-trends can change without leading us to observe a null treatment effect. Figure B3 shows the results of this sensitivity analysis for home country disease publications, and Figure B4 shows this for home and host country disease publications. We see that the critical value of M is 0.25 in the former figure, which means we can reject a null effect unless we are willing to allow for the linear extrapolation across consecutive periods to be off by more than 0.25 percentage points. For home and host country publications, the critical value of M is 0.7, which means , which means we can reject a null effect unless we are willing to allow for the linear extrapolation across consecutive periods to be off by more than 0.7 percentage points.

Figure B3: Sensitivity Analyses: Home Disease Publications (Rambachan and Roth, 2023)



Notes: This figure shows the results of sensitivity analysis based on Rambachan and Roth (2023). The original estimate (in red) is our OLS estimate. This approach formalizes the intuition that pre-trends are informative about violations of parallel trends by allowing the slope of the pre-trends to change by no more than some constant 'M' across consecutive periods. M = 0 corresponds with allowing only for linear violations of parallel trends, and larger values of M allow for larger deviations from linearity. We see that the critical value for a significant effect is M = 0.25. That is, we can reject a null effect (the treatment effect in a particular period is zero) unless we are willing to allow for the linear extrapolation across consecutive periods to be off by more than 0.25 percentage points.

Figure B4: Sensitivity Analyses: Home and Host Disease Publications (Rambachan and Roth, 2023)



Notes: This figure shows the results of sensitivity analysis based on Rambachan and Roth (2023). The original estimate (in red) is our OLS estimate. This approach formalizes the intuition that pre-trends are informative about violations of parallel trends by allowing the slope of the pre-trends to change by no more than some constant 'M' across consecutive periods. M = 0 corresponds with allowing only for linear violations of parallel trends, and larger values of M allow for larger deviations from linearity. We see that the critical value for a significant effect is M = 0.25. That is, we can reject a null effect (the treatment effect in a particular period is zero) unless we are willing to allow for the linear extrapolation across consecutive periods to be off by more than 0.25 percentage points.