Extended Abstract
Finding Nemo, Abroad[*]

Ratra, Vastav[†]
London School of Economics (LSE)

January, 2026

Abstract

In recent years, international students have become an increasingly important source of both revenue and research workforce at universities in developed countries. Yet little is known about how this expansion of educational migration has shaped the direction of scientific research conducted at these institutions. To study this question, I exploit the staggered introduction of English-taught master's programs (ETMPs) in biological and agricultural science departments at highly ranked universities in non-Anglophone developed countries. By changing the language of instruction, these programs lower barriers to entry for a broad pool of international students whose primary language is English. I measure the geographic orientation of research by linking the plant and animal species studied in each publication to their native habitats, using newly assembled habitat-level data from the ecological sciences. The analysis reveals large effects on both the composition of researchers and the geographic relevance of research output, indicating that changes in who enters universities can substantially redirect what science they produce.

[*]**Acknowledgements:** This paper is being written to meet the requirements for EC599 MRes research Paper. I thank my thesis advisors, Robin Burgess and Oriana Bandiera for their support and feedback through the process.

[†]**Correspondence:** E: v.ratra@lse.ac.uk

1 Introduction

Despite addressing global phenomena, modern science remains disproportionately focused on WEIRD settings (Western, Educated, Industrialized, Rich, Democratic), raising concerns about applicability of its findings. Across a wide range of domains, from agriculture (Suri and Udry (2022); Moscona and Sastry (2025)) and tropical diseases (Association et al. (2004); Dimitris et al. (2021)) to social and economic outcomes (Das et al. (2013)), scientific knowledge about developing countries remains strikingly limited, with potentially important consequences for their ability to confront pressing development challenges (De Souza (2022); Moscona and Sastry (2025)). These gaps, in part, reflect the limited research infrastructure and training capacity in many developing-country institutions, where small budgets constrain the scale and continuity of scientific work (Nchinda (2002); Lederman and Maloney (2003)).

This paper examines whether increased educational migration to better-funded universities in developed countries could redirect scientific attention toward previously under-studied topics and regions. I study this using the staggered introduction of English-taught master's programs (ETMPs) at highly ranked universities in 20 non-Anglophone developed countries, such as France and Germany. I argue that language of instruction constituted a major barrier to access for a large pool of prospective students, and eliminating this constraint altered the composition of researchers entering these institutions. By bringing in students from new countries, these programs may have reshaped the geographies and questions studied in these universities.

To trace shifts in the direction of scientific attention, I focus on departments in the biological and agricultural sciences, where research topics can be credibly linked to geographic contexts through the plants and animals being studied and the locations used to source samples for the paper. I construct a unique dataset which tracks the openings of hundreds of English-taught master's programs in biology-adjacent fields at 118 well-ranked universities across 20 Asian and European countries which attract significant numbers of international migrants. I restrict the data to years between 1990 and 2019. In the first year of the analysis, only four of these universities offered an ETMP. By 2010, this had risen to 70 universities, and by 2019, all but seven universities had at least one ETMP. While this expansion is rapid, it remains modest relative to the growth of English-taught programs in fields such as computer science and business administration. Each treated university is matched to a control university located in one of the "Big Four" destination countries—the United States, United Kingdom, Canada, and Australia—which have historically attracted the largest international student populations.

A novel index of geographical relevance constructed using this data for over 1.1 million papers published in biology-adjacent journals, suggests huge gaps in knowledge base across countries. For every 10 published papers of relevance to high-income countries, only 4.1 papers are relevant to upper-middle income countries, 2.4 papers to lower-middle income countries, and 0.6 papers to low-income countries. These gaps are larger

in the 236 universities that make up our sample.

If language of instruction constitutes a binding barrier to entry, the introduction of English-taught programs should first manifest as a shift in the composition of research talent at treated universities. Using an event study design estimated with robust methods following Callaway and Sant'Anna (2021), I find a statistically significant increase of 14 papers per year authored by researchers coming from countries that do not share the university's primary language. The estimates rise to 39 papers/year by the tenth year. At the same time, while there is an upward drift in the number of authors from linguistically similar countries, the point estimate is lower (7.7 papers/year) and is not statistically distinguishable from zero.

These changes in researcher composition translate into shifts in the geographic orientation of research output. Prior to ETMP introduction, universities in non-Anglophone countries in the sample produced on average 0.03 papers per 10 million population on countries that did not share their primary language. Following program introduction, research on linguistically different countries rises sharply, with event-study estimates indicating a large increase in output along this margin.

The empirical setting also allows me to examine mechanisms shaping research direction by observing how newly recruited researchers select research topics. Several factors may influence these choices, including the ecological and logistical requirements of studying particular species, prior exposure to specific plants or animals through cultural or personal background, and academic incentives related to publication and citation outcomes. Advisor guidance may further shape early-career research trajectories. I explore these channels in additional analysis, including tests of whether researchers with exposure to distinct climatic or biodiversity domains disproportionately contribute to research on those domains.

Finally, I return to the broader concern motivating the project: the production of policy-relevant knowledge across countries. I introduce new measures of policy relevance by parsing the full text of more than one million papers and focusing on publications in top-quartile journals. I document stark differences in policy-oriented discussion between papers focused on developed versus developing countries, even when authored by individuals from the same home country. Within-author comparisons show that papers on developing countries exhibit similar scientific rigor but receive fewer citations and contain substantially less policy-relevant content. These patterns are more pronounced in academically strong laboratories, proxied by advisors' success in placing students into tenure-track positions. Together, these findings suggest that while educational migration can redirect scientific attention toward under-studied regions, academic incentives may limit the extent to which such shifts translate into policy-central knowledge.

1.1 Related literature

This paper contributes to a long literature on the role of migrants in scientific production and innovation in developed countries. Much of this work focuses on the United States and primarily studies established scientists (for example, Moser et al. (2014); Ganguli (2015)). As a result, we know comparatively little about how the entry of new cohorts of researchers shapes the direction of scientific inquiry itself. Two closely related exceptions are Truffa and Wong (2025) and Subramani and Fry (2025). The former studies the effects of expanding women's access to institutions of higher education through the roll-out of co-education across universities in the United States and documents substantial increases in research on women-related topics. Importantly, however, the mechanism in that setting operates through changes in the research agendas of incumbent faculty, rather than through the skills, experiences, preferences or topic choices of students themselves. In contrast, this paper isolates a mechanism through which field and topic choices made by early-career researchers directly reshape the geography and content of scientific research. In this regard, the latter paper is closer to our work as it focuses on postdoctoral scientists, but it focuses on South-to-South migration, where the sample of migrant researchers might look really different from those in our study.

The paper also contributes to the literature on the internationalization of higher education in Europe and Asia through the expansion of English-taught master's programs. While a large body of economic research documents the role of language and medium of instruction in shaping trade flows, migration, and capital movements, comparatively little is known about the consequences of linguistic convergence and anglicization within developed-countries. So far, i could only find one study from Italy which exploits the introduction of English-taught programs in specific national contexts (Nocito (2021)), although some do examine the expansion of Erasmus programs across countries (Parey and Waldinger (2011); Parey et al. (2017)). More indirectly, Avdeev (2025) studies how cross-border educational migration between Netherlands and neighboring countries affects long-run outcomes such as international social ties and residential mobility. This paper extends that literature by introducing a new outcome—namely, the direction and geographic relevance of scientific research—and by providing causal evidence on how educational internationalization shapes scientific production. In doing so, it contributes empirical grounding to discussions of scientific diplomacy and international knowledge production that have largely remained rhetorical and outside the core economics literature.

The remainder of the extended abstract proceeds as follows. Section 2 describes the institutional background underlying the expansion of English-taught master's programs. Section 4 outlines the research design and identifying variation. Section 3 describes the data and measurement strategy. Section 5 presents the main empirical findings and discusses their implications. The abstract concludes with a timeline outlining the remaining steps required to complete the paper by the time of the conference pre-

sentation.

2 Background

2.1 English-Taught Master's Programs

The last few decades have witnessed a sharp increase in educational migration to non-traditional destination countries—that is, countries outside the United States, the United Kingdom, Canada, and Australia. A growing share of these flows has been absorbed by non-Anglophone, high- and upper-middle-income countries in Europe and Asia, which constitute the focus of this paper (Figure 1). A central mechanism through which these countries expanded access to international students is the introduction of English-taught master's programs (ETMPs): one- to three-year professional degrees offered in fields with broad international demand. By adopting English as the medium of instruction, these programs lowered language barriers that had previously restricted access to elite universities in non-Anglophone countries, thereby broadening the pool of prospective international students.

Several institutional forces shaped the timing and location of ETMP adoption. First, the Bologna Process, initiated in 1999, sought to harmonize higher-education systems across Europe by standardizing degree structures and introducing the European Credit Transfer and Accumulation System (ECTS). Prior to Bologna, limited compatibility of degrees across countries constrained both educational and labor mobility, as credits were difficult to transfer and employers faced challenges in interpreting foreign credentials. The Bologna reforms increased cross-border comparability and encouraged universities to redesign curricula. Many institutions used this transition as an opportunity to introduce English-taught programs aimed at attracting a broader international student base. Together with the Erasmus programs, Bologna represents a major phase of higher-education internationalization among universities in the sample.

Financial pressures also played an important role. Following the 2008 global financial crisis and the subsequent eurozone crisis, constraints on public funding increased universities' reliance on own-source revenues, particularly international student fees. English-taught professional master's programs provided a scalable means of attracting fee-paying international students, contributing to a shift away from exclusive dependence on government transfers.

Finally, two longer-run trends reinforced internationalization pressures. First, demographic shifts in several countries—aging populations and declining domestic student cohorts—reduced the supply of national students and early-career researchers, particularly in countries such as Japan. Second, global university rankings increasingly incorporated indicators of internationalization, often measured as the share of international students enrolled. In response, governments introduced funding schemes that rewarded universities for meeting enrollment targets for foreign students. These policies were largely discipline-neutral, leading to rapid expansion of English-taught pro-

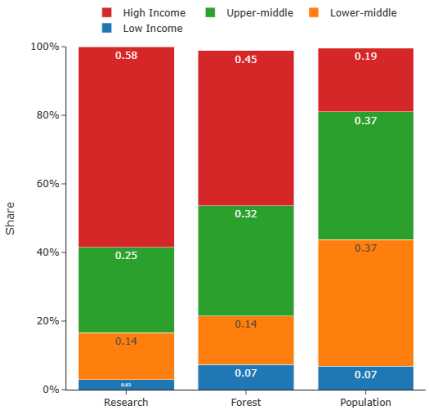


Figure 1: Research effort and need mismatch

grams in fields with large international markets, such as computer science and business administration. Biology and related disciplines, while present, remained a secondary focus of these internationalization efforts.

2.2 Geography of Biological Knowledge

We condense the information on the geography of relevance of biological knowledge into Figure 1. The first and second bars in the figure depicts the share of research in biology of relevance to different countries. As expected the proportion of research relevant to developing countries is much lower than developed countries. Then, we use two metrics to quantify mismatch in research. The gap in relevance is even starker when benchmarked against the global share of forest and population. Low-income countries receive half the research attention they should (3% versus 7%), given their forest area (which is also strongly correlated with biodiversity indicators). All but high-income countries receive lower focus than their share of the global forest area, with upper-middle income countries being particularly crowded out.

3 Data: Sources & Descriptive Statistics

This project relies on four key streams of data, which are discussed in turn.

3.1 Outcome Variables

3.1.1 Relevance of research output

To calculate the relevance of research performed at a university *i* for country *c*, we need to measure the geographical relevance of papers published at the university in a year to all countries. The metadata for all papers is obtained from Openalex.org, an open-source dataset containing information on millions of published papers and widely used in meta-science research (for example, Arts et al. (2025)). I restrict the sample to Q1-Q3 journals in biological and agricultural sciences, with first or last author based at sample universities. Roughly a million papers make the cut.

The geographical relevance of the papers is computed by parsing the title and abstract of these papers for three indicators - names of countries, names of GAUL-1 units (sub-national administrative units or states), and plants and animal names. For the last indicator, I also assemble a unique and globally representative data on habitats of over a million species to link each organism to countries where it is found. This is done using the Global Biodiversity Information Facility's (GBIF) backbone dataset, but several gaps in coverage and inconsistency in reported geographical units across species need to plugged manually. Using these indicators, I can assign each paper to all countries it could plausibly be relevant for. This is aggregated at the university level. Thus, for each university-country pair and for each year, we have a measure comprising of the number of papers of relevance to country written at the university in a year.

3.1.2 Researcher composition

A key intermediate outcome in this project is the effect of EMTP on composition of young researchers at these universities. To calculate this, I put together the names of all first and last authors of papers published at these universities. I cannot directly link these individuals to their nationalities, but I use a prediction algorithm to assign probability of belonging to a given nationality based on the names of these individual. The algorithm is far from perfect, but provides reasonable predictions useful for our sniff test. For example, it assigns most Anglo-Saxon names to United Kingdom, and does not assign a single author to Canada, Australia or the United States. Importantly, though, it is a good predictor of the linguistic origin of the author even though its ability to match nationalities might be severely limited. Thus, for university-year, I compute the number of authors coming from linguistically similar and dissimilar countries as the primary first-stage outcome variables.

3.2 Identifying Variation

The variation in my project comes from the staggered opening of English-taught programs across universities in my sample. There is no consolidated dataset of launch dates of ETMPs. So, I artisinally collect and verify this data through multiple sources. Most

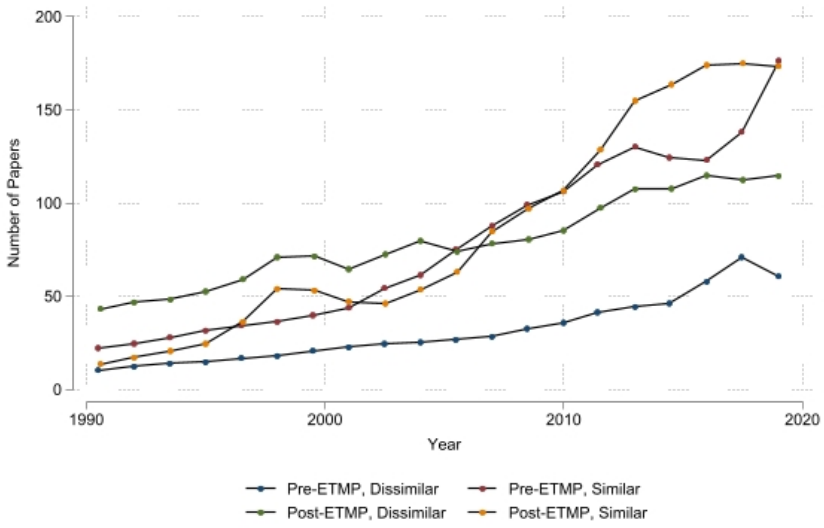


Figure 2: Caption

importantly, I go through the internet-archive websites of each of the universities over several years and identify the announcements associated with opening of different programs at the university. Among the newly opened programs, I single out those run at the biology or agriculture departments, taught in English, and open to international students without constraints. For the sake of treatment definition, I retain the first of these programs launched at the university.

3.3 Covariates

The paper relies on a series of covariates from geographical and socio-economic data-sources. A bunch of economic variables like country-level per-capita GDP and Gini indices are sourced from World Bank's World Development Indicators (WDI), alongside remotely sensed and globally harmonized climate variables from standard satellite sources.

3.4 Authorship Composition Over Time

We briefly discuss the evidence in shifts in linguistic diversity of authors over time at the treated universities. Figure **??** shows my key descriptive. For this figure, I split the sample of universities into treatment and control in a dynamic sense - in any given year, a university might be in the period before the launch of their ETMP or after. Thus, each point in time in this figure captures different set of universities. I look at the evolution of authors from linguistically similar and dissimilar countries over time. Three observations are notable. First, universities with and without ETMP are similar in the level and trend of linguistically similar authors over time. In addition to serving as a placebo,

it provides some intuition for balance across universities on their appeal among the local populations. Second, our key metric of interest, the gap in the number of papers by authors from linguistically dissimilar authors is large and stable, at roughly 50 papers/year. Third, over time, publications with authors from linguistically similar countries (including one's own) evolve more steeply than those from dissimilar countries.

4 Research Design

To identify the effects of changes in the diversity of the research workforce on the direction of scientific research at the university level, I exploit a plausibly exogenous shock to the composition of early-career researchers induced by the staggered introduction of English-taught master's programs (ETMPs) at universities in non-Anglophone countries. The rollout of ETMPs lowered language barriers to entry for prospective international students and researchers, generating variation over time and across institutions in the linguistic and geographic origins of research talent. Treated universities are matched to a panel of never-treated universities located in Anglophone countries, which historically attracted large shares of international students but did not experience comparable changes in language of instruction during the study period.

I estimate dynamic treatment effects using an event-study framework following Callaway and Sant'Anna (2021) that compares outcomes at treated universities before and after ETMP introduction to contemporaneous outcomes at matched control universities. The primary outcomes include measures of researcher composition, which is proxied by the linguistic similarity between first authors' predicted countries of origin and the host university's primary language, as well as measures of the geographic relevance of research output. Estimation follows modern difference-in-differences methods that accommodate staggered treatment timing and heterogeneous treatment effects across cohorts.

Identification relies on a parallel trends assumption: in the absence of ETMP introduction, outcomes at treated and control universities would have evolved similarly over time. For specifications based on linguistic similarity, this assumption implies that research conducted by non-Anglophone universities on linguistically dissimilar (Anglophone countries) countries would have followed trends comparable to research conducted by Anglophone universities on the territories dissimilar to them (non-Anglophone countries), absent the policy change. Thus, we are assuming that research across different contexts evolve parallely in the absence of the treatment.

This design has both advantages and limitations. On the one hand, comparing trends across distinct geographic contexts helps mitigate concerns about spillovers or strategic interactions among universities operating in closely related research fields or national systems. On the other hand, the design may be sensitive to global "attention" shocks—periods in which certain regions or topics receive increased scientific attention across institutions for reasons unrelated to ETMP adoption. Such shocks could generate differential trends in geographic research focus that are correlated with treatment

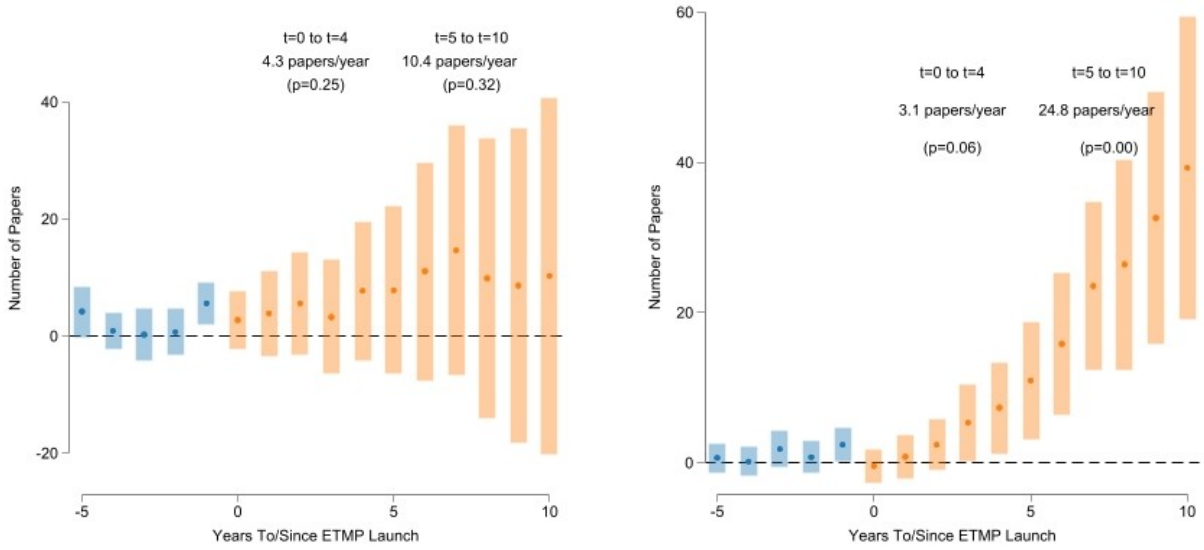
timing but not caused by changes in researcher composition.

I address these concerns in several ways. First, the event-study specification allows for visual and statistical inspection of pre-treatment trends, and robustness checks following Rambachan and Roth (2023) explicitly assess sensitivity to deviations from parallel trends. Second, the inclusion of university and year fixed effects absorbs time-invariant institutional characteristics and common shocks affecting all universities. Third, in subsequent analysis, I extend the design to a triple-difference framework that exploits within-university variation across linguistic similarity and research geography, further isolating the role of researcher composition from broader shifts in global scientific attention.

Together, this research design leverages institutional policy variation in graduate education to credibly identify how changes in the composition of early-career researchers shape the direction and content of scientific research.

5 Results

5.1 Researcher Diversity



(a) Linguistically Similar Authors **(b)** Linguistically Dissimilar Authors

Notes: This figure shows event study estimates from specification **??** separately for count of papers where first authors are from linguistically similar (a) and dissimilar (b) countries.
Source: *Author's calculations*

Figure 3: Estimates for Effect of ETMP on Geography of Researchers

Figure 3 presents the main first-stage result. Event-study estimates indicate a large, rising, and statistically significant increase in research output authored by individu-

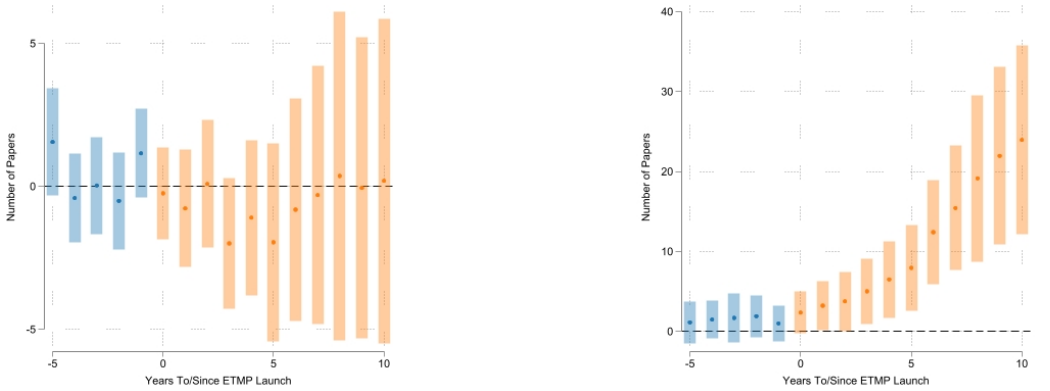
als from linguistically dissimilar countries following the introduction of English-taught master's programs. The pooled effect over the ten years after program introduction is an increase of 14 papers per year. Estimates are smaller in the initial years, consistent with expected lags in research production as new entrants integrate into university research environments. I assess the sensitivity of these estimates to potential violations of the parallel trends assumption by implementing the honest difference-in-differences approach proposed by Rambachan and Roth (2023). The pooled estimate over years five through ten remains robust to violations up to 1.5 times the magnitude of pre-treatment trend differences, suggesting that the observed effects are unlikely to be driven by differential long-run trends between treated and control universities.

In contrast, estimates for papers authored by individuals from linguistically similar countries are positive but statistically indistinguishable from zero, both at individual event times and when pooled over the decade following program introduction. This pattern is not mechanically implied. Evidence from the United States suggests that international student enrollment can cross-subsidize domestic participation in higher education (Bound et al. (2020)). Several features of the present setting may explain the absence of such effects here. First, cross-subsidization in the United States primarily operates at the undergraduate level rather than in graduate education. Second, biology-adjacent fields might attract lower international demand than fields such as computer science or business administration. Third, unlike the United States, international students in master's programs across Europe and Asia were not subject to steep tuition discrimination during the period studied. In many sample countries, international STEM students were eligible for substantial public scholarships, including France's Charpak program, Germany's DAAD, Japan's MEXT, China's government scholarships, and Switzerland's Swissnex.

Two caveats are worth noting. First, the results imply an expansion of the research workforce following program introduction, which may affect research output through scale effects independent of diversity-driven mechanisms. So, establishing a tighter link between researchers' country of origin and the content of produced research is essential. Second, while the analysis identifies level effects of the policy, previously presented evidence (Figure **??**) suggests differential trends by linguistic similarity, with locally sourced researchers exhibiting steeper growth over time. I address these issues in subsequent analysis.

5.2 Research Diversity

Figure 4 presents the main results on changes in the geographic relevance of research following the introduction of English-taught master's programs. Consistent with earlier findings, I find no statistically significant effect on research output relevant to linguistically similar countries. In contrast, universities produce, on average, ten additional papers per year that are relevant to linguistically different countries after program introduction. Results disaggregated by continent, omitted here for brevity, show a strong



(a) Linguistically Similar Authors **(b)** Linguistically Dissimilar Authors

Notes: This figure shows event study estimates from specification **??** separately for count of papers with geographical relevance to linguistically similar (a) and dissimilar (b) countries.
Source: *Author's calculations*

Figure 4: Estimates for Effect of ETMP on Geography of Research

correspondence between increases in authorship from a given continent and contemporaneous increases in research relevant to that region.

Despite these shifts, several frictions may weaken the link between a researcher's country of origin and the geographic focus of their research. Advisor expertise and preferences may constrain topic choice, particularly when advisors lack prior experience working on certain organisms or regions. In addition, early-career researchers from developing countries face asymmetric incentives. Specializing in research on their home country may generate expertise that is less valued in international academic labor markets, while domestic research opportunities in those fields may be limited. These effects suggest that the frictions, while severe, might not be insurmountable.

5.3 Timeline

The core empirical analysis linking the introduction of English-taught master's programs to changes in researcher composition and research output is complete and robust to alternative specifications, sample definitions, and outcome measures. By the time of the conference, I will extend the analysis using full-text data from publications in Q1 journals to refine measurement of policy relevance and scientific contribution. These data have already been collected, and I am currently finalizing text-based indicators that distinguish policy-oriented discussion from scientific content. This extension will allow a more precise assessment of the policy appropriateness of research produced on developing-country contexts.

In parallel, I will incorporate a novel measure of policy demand at the organism level, constructed using a newly assembled dataset of policy documents and country

submissions to international biodiversity fora. These documents have already been downloaded and are currently being processed. Combining measures of policy demand with research output will make it possible to assess whether shifts in scientific attention align with areas of greater policy need.

Finally, I will expand the analysis of mechanisms underlying the observed effects, focusing on how researcher diversity interacts with characteristics of the organisms studied, such as ecological domain and fieldwork requirements. The paper will also more explicitly situate the findings within the literature on migration, diversity, and scientific productivity. I expect the paper to be fully conference-ready by March 2026.

References

Arts, Sam, Nicola Melluso, and Reinhilde Veugelers (2025) "Beyond citations: Measuring novel scientific ideas and their impact in publication text," *Review of Economics and Statistics*, 1–33.

Association, Canadian Medical et al. (2004) "Western medical journals and the 10/90 problem."

Avdeev, Stanislav (2025) "University as a Melting Pot: Long-term Effects of Internationalization."

Bound, John, Breno Braga, Gaurav Khanna, and Sarah Turner (2020) "A passage to America: University funding and international students," *American Economic Journal: Economic Policy*, 12 (1), 97–126.

Callaway, Brantly and Pedro HC Sant'Anna (2021) "Difference-in-differences with multiple time periods," *Journal of econometrics*, 225 (2), 200–230.

Das, Jishnu, Quy-Toan Do, Karen Shaines, and Sowmya Srikant (2013) "US and them: The geography of academic research," *Journal of Development Economics*, 105, 112–130.

De Souza, Gustavo (2022) "The Labor Market Consequences of Appropriate Technology,"Technical report, Federal Reserve Bank of Chicago.

Dimitris, Michelle C, Matthew Gittings, and Nicholas B King (2021) "How global is global health research? A large-scale analysis of trends in authorship," *BMJ global health*, 6 (1).

Ganguli, Ina (2015) "Immigration and ideas: what did Russian scientists "bring" to the United States?" *Journal of Labor Economics*, 33 (S1), S257–S288.

Lederman, Daniel and William F Maloney (2003) "R&D and development," *Available at SSRN 402480*.

Moscona, Jacob and Karthik A Sastry (2025) "Inappropriate technology: Evidence from global agriculture,"Technical report, National Bureau of Economic Research.

Moser, Petra, Alessandra Voena, and Fabian Waldinger (2014) "German Jewish émigrés and US invention," *American Economic Review*, 104 (10), 3222–3255.

Nchinda, Thomas C (2002) "Research capacity strengthening in the South," *Social science & medicine*, 54 (11), 1699–1711.

Nocito, Samuel (2021) "The effect of a university degree in English on international labor mobility," *Labour Economics*, 68, 101943.

Parey, Matthias, Jens Ruhose, Fabian Waldinger, and Nicolai Netz (2017) "The selection of high-skilled emigrants," *Review of Economics and Statistics*, 99 (5), 776–792.

Parey, Matthias and Fabian Waldinger (2011) "Studying abroad and the effect on international labour market mobility: Evidence from the introduction of ERASMUS," *The economic journal*, 121 (551), 194–222.

Rambachan, Ashesh and Jonathan Roth (2023) "A more credible approach to parallel trends," *Review of Economic Studies*, 90 (5), 2555–2591.

Subramani, Gauri and Caroline Fry (2025) "Migration and Local Problem-Solving," in *Academy of Management Proceedings*, 2025, 11212, Academy of Management Valhalla, NY 10595.

Suri, Tavneet and Christopher Udry (2022) "Agricultural technology in Africa," *Journal of Economic Perspectives*, 36 (1), 33–56.

Truffa, Francesca and Ashley Wong (2025) "Undergraduate Gender Diversity and the Direction of Scientific Research," *American Economic Review*, 115 (7), 2414–2448.