Economic
Research
Working Paper
No. 88/2025

Digital Access to Knowledge and Women in Science

Elodie Carpentier, Alexander Cuntz,
Alessio Muscarnera, Julio Raffo

WIPO

Digital Access to Knowledge and Women in Science

Elodie Carpentier* Alexander Cuntz† Alessio Muscarnera†

Julio Raffo†

May, 2025

Abstract

Scientific progress relies on access to prior knowledge, yet costly access to academic literature can hinder researchers, particularly in marginalized positions of academia and developing economies. This paper examines the impact of free or lower-cost access to scientific literature on gender representation in research. Leveraging the staggered adoption of the Hinari program, which provides digital access to health science research, we analyze its effects on women's participation in research production and academic publishing across more than 600 institutions in 80 countries. Using a triple difference approach, we find that improved digital access to knowledge increases the share of women scientists in publishing faculty and enhances their research output. The program's effects are most pronounced in countries with lower gender balance in educational attainment, where it appears to help overcome attainment gaps and activate women's potential in academic labor markets. Our study contributes to the literature on digitization, access to knowledge and gender disparities in academia, while also helping to inform science and innovation policy and human capital development.

Keywords: Open science, Development, Health Science, Gender studies, Triple difference, Impact evaluation.

JEL Codes: J16, L17, 031, O33

---

*European Commission, Joint Research Centre (JRC), Seville, Spain.

†World Intellectual Property Organization, Department for Economics and Data Analytics, Geneva, Switzerland.

Corresponding author: `alessio.muscarnera@wipo.int`. The authors gratefully acknowledge Kimberley Parker and the World Health Organization for sharing the data on the Research4Life program. The authors would like to thank Mercedes Delgado, Valeria Esquivel, and the participants at the European Policy for Intellectual Property Association 2024 Annual Conference for valuable feedback on the research. The views expressed are those of the authors, and do not necessarily reflect the views of the European Commission or the World Intellectual Property Organization or their member states.

1 Introduction

Scientific discoveries are a key driver of innovation, economic growth, and development. The cumulative nature of science builds on prior knowledge, making access to research inputs crucial for scientific progress (Scotchmer, 1991; Mokyr, 2011; Azoulay et al., 2015). However, costly access to academic literature can limit the ability of researchers — especially in developing economies — to engage in scientific production. Programs like the UN-led Research4Life (R4L) initiative aim to lower these barriers, promoting knowledge diffusion and reinforcing the open science paradigm (Merton, 1973; Partha and David, 1994).

While such initiatives enhance participation in global science and innovation (Müeller-Langer, Scheufen and Waelbroeck, 2020; Cuntz et al., 2024), their impact on gender inclusion remains largely unexplored. Women continue to be underrepresented in academia, facing persistent disadvantages in publishing, citations, and career progression (Larivière et al., 2013; Hofstra et al., 2020; Huang et al., 2020; Rose and Georg, 2021; Koffi and Marx, 2023; Bikard, Fernandez-Mateo and Mogra, 2025). The "leaky pipeline" effect describes the steady attrition of women in science (Alper, 1993; Buckles, 2019), often due to unequal access to critical resources (Duch et al., 2012). Existing research shows that fields with lower resource intensity have a higher proportion of women scientists (Duch et al., 2012), but whether access to knowledge follows similar patterns is an open question.

This paper examines the effects of free or low-cost access to scientific knowledge on gender composition in developing economies. Specifically, we analyse whether improved access to academic literature influences the representation and research output and quality of women scientists. Leveraging data from the Hinari program — launched in 2002 to provide digital access to health science research — we track adoption by 608 institutions across 80 countries. Using Microsoft Academic Graph (MAG) data on over 800,000 publications and assigning gender to authors via the World Gender Name Dictionary (WGND), we construct a unique dataset to assess gender dynamics in scientific production.

Our empirical strategy employs a triple difference approach, comparing gender composition and research output before and after Hinari adoption within institutions and across disciplines (Gruber, 1994; Olden and Møen, 2022). This addresses concerns about selection bias, as institutional policies on diversity and inclusion are typically uniform across scientific fields.

We find that free or lower-cost access to knowledge has significant compositional effects on local research production, increasing the share of women scientists in publishing faculty by 2 to 9 percent and their contributions to research output (share of papers co-authored by women), ranging between 3 to 30 percent depending on the model specification. While program adoption raises overall participation of women, its impact on research quality is slightly weaker, but still positive, with an approximate average increase of 2 percent. Alternative dynamic estimators accounting for the staggered program adoption across institutions and robustness checks using local usage of program resources as an alternative continuous treatment confirm our main findings but indicate that gender composition effects take several years to fully unfold and materialize. Furthermore, we provide evidence on heterogeneity, showing that the effects are most pronounced in countries with low levels of gender balance in educational attainment. In these contexts, the program appears to help overcome gender gaps in attainment and activate the potential of women scientists in academic labor markets, thereby contributing to gender equality in scientific production and the UN's Sustainable Development Goals.

This research contributes to multiple strands of literature, including digitization and science (Agrawal and Goldfarb, 2008; Butler, Butler and Rich, 2008; Kim, Morse and Zingales, 2009), gender disparities in academia (Ding et al., 2010), and the role of intellectual property and knowledge diffusion (Biasi and Moser, 2021; Kaiser, Cuntz and Peukert, 2023; Cuntz and Sahli, 2024). Prior studies suggest that digital technology can be an equalizing force, especially for researchers in marginalized positions (Sproull, Kiesler and Kiesler, 1991). We extend this line of research by demonstrating that improved digital and remote access to scientific knowledge disproportionately benefits women scientists in developing economies, fostering inclusivity and mitigating structural disadvantages in academic publishing.

From a policy perspective, research findings highlight how targeted access initiatives and international public-private partnerships such as R4L can help shape the composition of research communities and improve gender balance in knowledge production in less developed parts of the world. This also holds the potential for longer-term growth in local human capital and economic development extending the findings on program impact from prior research (Müeller-Langer, Scheufen and Waelbroeck, 2020; Cuntz et al., 2024).

The paper is structured as follows: Section 2 reviews the literature, Section 3 presents the Hinari initiative, Section 4 describes the data and variables, Section 5 outlines the empirical strategy, Section 6 presents results, Section 7 discusses the policy implications of our findings, and Section 8 concludes.

2 Related Literature

The literature identifies several important barriers hampering women's progress in science and innovation, all the way down from biological to socio-cultural and institutional factors, including implicit stereotypes, but also conscious and unconscious biases (Ginther and Kahn, 2009; Ceci, Williams and Barnett, 2009; Ceci et al., 2014; Carlana, 2019; Bikard, Fernandez-Mateo and Mogra, 2025). Most lately, a growing literature looks at small-scale experiments and provides guidance on which interventions can help effectively improve the gender balance in academia (Buckles, 2019). Still, women are systematically underrepresented in science and innovation systems, notwithstanding their important contribution to scientific novelty among other aspects (Buffington et al., 2016; Hofstra et al., 2020; Huang et al., 2020; Carpentier and Raffo, 2023). One central barrier for scientists in general, as well as women's involvement in science and career progression in particular, is research funding and resource constraints (Duch et al., 2012; Chang et al., 2019; Graddy-Reed, Lanahan and D'Agostino, 2021). What we perceive is less well understood in the resource context is the specific role of access to prior knowledge and the ability of women scientists to gain access. Similar applies to the implications for scientific production and potential changes in gender imbalance in the science system once access is granted to them. This surprising gap in the literature sits at the core of this paper.

Conceptually, access to prior research can eliminate information asymmetries and help scientists learn about future research and funding opportunities, improve teaching quality and education, and better understand the current frontiers in science and technology. In this way, improved access might avoid research duplication and improve the efficient allocation of time and resources in global science. From a gender perspective and on a more practical level, however, through digital online access provided to local researchers through the UN-led programs, the latter also grants more anonymous and remote resource access to researchers. Eventually, this can help overcome some of the conscious or unconscious gender biases introduced by human interaction (Ginther and Kahn, 2009; Ceci, Williams and Barnett, 2009; Ceci et al., 2014; Carlana, 2019; Bikard, Fernandez-Mateo and Mogra, 2025), potentially limiting the control gatekeepers in the science system may exercise over access by researchers in more marginalized positions in the hierarchy, or by researchers facing mobility constraints, in particular women and other minority groups (Marwell, Rosenfeld and Spilerman, 1979; Shauman and Xie, 1996). More broadly, access based on digital technology can help enable wider access to research equipment and materials. For example, as shown in prior research, new digital technology can allow remote access to equipment such as synchrotrons and telescopes (Stephan, 2010). Use of new technologies to access prior knowledge hence holds the potential to increase scientific output by also granting access to marginalized and mobility-constrained groups. Moreover, digital access might not only make science production more inclusive, but also can impact the quality and direction of new research. For instance, women-led research teams are more likely to address and deliver innovative solutions to problems unique to women (Koning, Samila and Ferguson, 2021).

Similar applies to science and innovation impact when knowledge is codified and disclosed locally to the public or targeted groups (Currie and Moretti, 2003; Furman and MacGarvie, 2007; Aghion et al., 2009; Furman and Stern, 2011; Iaria, Schwarz and Waldinger, 2018; Berkes and Nencka, 2019; Furman, Nagler and Watzinger, 2021). In a historical study of U.S. invention, Furman, Nagler and Watzinger (2021) show that, following the establishment of patent libraries during the pre-Internet era, local patenting activity increased by 8–20 percent relative to other similar regions, and regions with patent library access saw increases in local business formation and job creation. In the context of early twentieth-century science, Iaria, Schwarz and Waldinger (2018) find that a boycott of enemy science during World War I discouraged the creation of new science. Moreover, a related study by Berkes and Nencka (2021) finds that granting access to public libraries and codified knowledge more broadly also helped increase women's patenting activities in a given place.[1]

Knowledge also flows through informal interactions, in addition to formal information channels such as prior scientific publications and publication of other research materials. In the academic context, again, women typically have fewer academic network ties and are less well integrated in research networks, as well as have fewer international collaborations as compared to men (Larivière et al., 2013; Liu, Song and Yang, 2020). On average, women scientists travel less often to conferences and seminars, where academic networking occurs, ideas are exchanged, and collaboration opportunities may emerge. Another way of expanding one's academic network is through job mobility. However, there is evidence that the mobility of women scientists is more constrained than that of men (Marwell, Rosenfeld and Spilerman, 1979; Shauman and Xie, 1996). Such conditions likely further decrease their access to informal knowledge sources. Andrews (2019) emphasizes the role of informal social interaction in inventive activities: alcohol prohibition in the U.S. decreased the average number of patents by 8-18 percent granted in prohibiting counties, showing a relatively smaller impact on groups of inventors that less heavily can rely on informal social interactions, notably women. In turn, this also means women might have to rely more heavily on sources and access to codified knowledge, including that provided

[1]Their rigorous study compares patenting in cities with public library access (treated group) to patenting in cities eligible for public library programs, but that ultimately did not build a library (control group).

through the UN initiative.

Via its focus on online access to new knowledge, this research contributes to the literature on the effects of digitization on science (Agrawal and Goldfarb, 2008; Butler, Butler and Rich, 2008; Kim, Morse and Zingales, 2009; Ding et al., 2010; Winkler, Levin and Stephan, 2010). Previous research has shown that the benefits of using information and technology do not affect men and women equally (Ding et al., 2010). Their research suggests that digital technology is an equalizing force, providing a greater boost to productivity and more collaboration opportunities for scientists who are more marginally positioned in academia and see themselves as less well integrated into academic networks. Early works by Barley (1986) and Orlikowski (1992) show that whether technological innovation is adopted and how it is used largely depends on the type of individuals using it and the organizational environment in which the innovation is introduced. So, relevant in the context of our paper, the relative incremental effect that access to technology has on productivity and collaboration depends on how marginally positioned a scientist is in the scientific hierarchy (Sproull, Kiesler and Kiesler, 1991).[2] We offer a more nuanced view of new technology use and extend this line of research by showing that online access to scientific knowledge over-proportionally benefits publication activity and quality of papers (co-)authored by women researchers in developing economies.

Finally, the research also adds to the literature on the effects of intellectual property (Murray and Stern, 2007; Nagaraj, 2018; Giorcelli and Moser, 2020; Biasi and Moser, 2021; Kaiser, Cuntz and Peukert, 2023; Cuntz and Sahli, 2024), the local provisioning of research and knowledge resources (Currie and Moretti, 2003; Furman and MacGarvie, 2007; Aghion et al., 2009; Furman and Stern, 2011; Iaria, Schwarz and Waldinger, 2018; Berkes and Nencka, 2019; Furman, Nagler and Watzinger, 2021) and informational shocks on science and innovation (Zheng and Wang, 2020; Kong et al., 2022; Hussinger and Palladini, 2024).

[2]Scientists at the "top" can already rely on a large network and typically have access to strong colleagues, excellent graduate students, and state-of-the-art equipment. In contrast, scientists at the "bottom" have considerably less access, and their networks are relatively smaller (e.g., minority groups) or only emerging (e.g., junior scholars). Thus, the effect of information technology is likely greater for those at the margin than for those at the top.

For example, Furman and Stern (2011) show that creating local research centers, which facilitated access to biological materials, has encouraged follow-on science. Closely related to our work, research by Biasi and Moser (2021) exploits an exogenous change toward weaker copyrights due to World War II to examine the effects of copyrights on follow-on science. They observe a substantial increase in citations to books as a measure for new cumulative knowledge building, as books impacted by framework changes were becoming more affordable through lower prices and being widely accessible. As the paper concludes,

> "compared with policies that enforce open access, schemes of price discrimination [or price subsidies] can encourage follow-on science, especially when regular fees may be prohibitive. Such policies are particularly important for promoting the creation of new knowledge among researchers at institutions that are cash-constrained." (Biasi and Moser, 2021, p. 222)

With our analytical focus on scientists in developing economies, we provide further empirical testing and new evidence supporting their general claim and extending it to the gender context.

3 Background

Public policy has long sought to address these issues, in particular to improve the overall conditions of access for marginalized groups, including women, and to provide access to science for researchers with limited financial resources in less developed regions of the world. In the context of this study, a consortium of five United Nations agencies has launched an initiative called Research for Life (R4L) to provide researchers and students at non-profit/public research institutions from developing countries with free or low-cost access to academic publications.

Institutions in about a hundred developing countries have been granted access to the repository. The access model is tiered, with the least developed countries (Group A) receiving free online access and more developed countries (Group B) paying a nominal fee. The classification of countries into Group A or B is based on explicit economic development criteria, determined at the national level.[3]

The R4L initiative includes five programs, administered by five different agencies, that provide access to prior research in different research fields as follows: the Hinari for health (World Health Organization - WHO), the Agora for agriculture (Food and Agriculture Organisation - FAO), Oare for the environment (United Nations Environment Program - UNEP), Ardi for innovation (World Intellectual Property Organization - WIPO), and Goali for global justice (International Labor Organization - ILO) research materials. Funding support and participation of publishers in Research4Life have been relatively stable over time (Cuntz et al., 2024).

In 2002, the collaborative endeavour between the World Health Organization and scholarly publishing entities resulted in the establishment of the Hinari program. Hinari aims to augment the capabilities of low- and middle-income economies by providing local researchers with access to an extensive repository of biomedical and health literature. The repository encompasses a compendium of 21,000 peer-reviewed scientific journals, accompanied by a substantial collection of 69,000 e-books and an additional 115 informational resources. Currently, these scholarly assets are accessible to researchers affiliated with health research institutions, including universities and teaching hospitals, spanning more than 124 countries, regions, and territories.

As the first program launched by the World Health Organization under the Research4Life umbrella initiative, Hinari represents a cornerstone in the collaborative efforts of the various United Nations agencies. From a methodological perspective, studying Hinari is strategically advantageous because it mitigates potential cross-treatment estimation bias arising from enrollment in other programs within the R4L framework.

Moreover, and importantly, we note that access provided by the Hinari program is likely to be the most important, and in many instances, the only source of information available to researchers in developing countries as more extensively discussed in prior research (Cuntz et al., 2024). This is because Hinari (a) provides access to scientific

[3]Countries' eligibility for Research4Life is determined as follows: countries with a Gross National Income (GNI) per capita of $1,025 or less, classified as Least Developed Countries by the United Nations, or with a Human Development Index (HDI) score of 0.55 or less are assigned to Group A, receiving free access to R4L resources. Countries with a GNI per capita between $1,026 and $3,995, classified as Low-Income Countries or Lower-Middle-Income Countries, or with an HDI score between 0.56 and 0.79 are assigned to Group B, paying a nominal fee of $1,500 per year for access to R4L resources.

publications from all over the world, (b) covers a large share of the top-tier biomedical and health publications, and (c) as a dedicated program, explicitly targets researchers in lower income countries. Alternatives routes of open access for researchers in these countries such as pre-print or post-print repositories only sporadically cover prior research in biomedical and health sciences. Furthermore, larger open access initiatives such as Plan S (2018) and Projekt Deal (2019), with their limited geographical scope, were only launched several years after the pioneering Hinari program (2002). In sum, this suggests that there are likely no, or very few, alternative sources of access to prior knowledge in the biomedical and health sciences that could challenge effect attribution to the Hinari program.

4 Data sources and descriptives

4.1 Research4Life initiative

We obtain information on the institutions participating in the Hinari program from the WHO, including on the name of institutions enrolled, the date of enrollment as well as the eventual date of enrollment to any of the other programs of the Research4Life initiative, and the group to which it has been assigned (A or B). In the 80 countries of analysis, 602 out of 1,199 institutions subscribed to Hinari.[4] Figure 1 shows that enrollment occurred over time over the 2002-2018 period, particularly in the first few years after the program was introduced, and with a spike in 2018, likely as a result of a new registration system introduced by Research4Life that allowed institutions to register online and access all R4L programs (including Hinari) through a single registration process. As part of this change, institutions that registered for one program were automatically enrolled in all other R4L programs for which they were eligible.

Furthermore, we obtain data concerning the number of recorded logins ("usage") per institution and year related to the Hinari program. However, the data on actual usage

[4] Although we mentioned previously that institutions have joined Hinari in 124 countries, we focus our analysis on the 80 countries in which at least one institution enrolled in Hinari before 2019. It is worth mentioning that within this subset of countries, we identified 24 instances where countries transitioned between two distinct groups (from group A to group B or vice versa).

and access to Hinari program resources was only collected as of 2007, i.e., five years after the initial start of the program. Despite the limited observation period, this data and restricted sample analysis still allow us to validate the main results obtained from baseline models and quantify the relationship between the actual usage of Hinari program resources and the dependent variables of interest.

4.2 Scientific publications

For all institutions in the 80 countries of interest, including those that never register in R4L, we retrieve a set of bibliometrics at the scientific article level from the Microsoft Academic Graph (MAG). Although we extract only a small portion of MAG in the context of this study, this database encompasses a diverse graph comprising over 120 million publication entities, with their respective authors, institutions, venues, and fields of study. MAG represents the most comprehensive dataset within the realm of bibliometric scientific articles, which is of particular importance when studying research outputs from less-developed countries where scholars publish more frequently in non-English journals that are usually less-well covered in other databases (Visser, van Eck and Waltman, 2021).

We retrieve a total of 757,910 scientific articles published between 1990 and 2018. Overall, 70 percent of the articles have at least one author with an affiliation to an institution that at some point joined the Hinari program.[5]

As we explain further in Section 5, an important feature of this study is the classification of articles into scientific fields, in particular distinguishing between those related to health and covered by the Hinari program, and all other (non-health related) fields. To perform this allocation of articles to fields, we leverage MAG's automated system that identifies and extracts an extensive range of over 200,000 scientific fields from the abstracts and titles of scholarly papers. This vast array of fields significantly adds to the complexity of categorizing papers into various research domains. We exploit a recent dataset provided by Marx and Fuegi (2020) to overcome this issue. The authors mapped papers into 6

[5]In contrast, 35 percent of articles report at least one author from an institution that has not joined Hinari as of 2018. Percentages do not add due to multiple authorships and affiliations. Similar to the approach in Cuntz et al. (2024), in case of multiple affiliations, the single author's contribution to the publication is defined as a fractional count (e.g., the institution-author contribution of a single author with two affiliations is 0.5 for each listed institution).

OECD fields and 39 sub-fields, establishing a bridge between the OECD classification and the 251 fields in the Web of Science (WoS). These WoS categorizations enable us to differentiate between papers related to biomedicine and health, and those unrelated to health.

In addition, Marx and Fuegi (2020) also provide data on the journal impact factor (JIF) of scientific articles, which we use as a proxy for the quality of articles.[6]

4.3 Gender assignment

Another important challenge of this paper is the gender disaggregation of authors to measure the representation and contribution of women to science before and after the program's adoption. While MAG provides a unique identifier for the 657,923 authors in our database over time, assigned at the time of first publication and repeated for all subsequent publications, it does not record the gender of the authors. Therefore, we use Stata Statistical Software's *genderit* command, which relies on the second version of the World Gender Name Dictionary (WGND) to determine the most likely gender of authors based on their first names and country (Raffo and Lax-Martínez, 2018; Lax-Martínez et al., 2021; Raffo, 2021).

The dictionary contains more than 26 million records linking first names and 195 countries and territories to a binary gender. Because authors' nationality or country of origin is not available in bibliometrics data, we use the country of institutional affiliation as a second-best option. We perform the gender assignment using a 0.8 confidence threshold, meaning that name-country pairs should be associated with a given gender at least 80 percent of the time across all sources within the dictionary. Unisex names for which the probability is around 0.5 and rare names not found in the dictionary are assigned an "unknown" gender and authors are not included in the analysis.[7] At the end of this process, we identified a binary gender for 75 percent of all authors, 36 percent of whom are women and 64 percent men (see Table 1).[8]

[6]Some journals have missing values for the JIF in specific years, affecting a substantial number of scientific articles. We use a linear interpolation to fill some of the gaps when the gap is no more than two years.

[7]When all authors of an article have an unknown gender, the article is removed from the database.

[8]Please note that the *genderit* command statistics reported in the appendix table indicate the

4.4 Educational attainment

Finally, to further assess the importance of the "leaky pipeline" and its potentially mediating role in the relationship between the program adoption and gender composition in science, we collect data on countries' gender balance at different stages of educational attainment. The World Economic Forum provides a Global Gender Gap Index allowing us to assess and quantify gender disparities across various countries and over time. It measures gender gaps in four key areas: economic participation and opportunity, educational attainment, health and survival, and political empowerment.

We retrieve values of the Gender Gap Educational Attainment (GGEA) index for 68 countries over the years 2006-2018, as the composite indicator was first published in 2006 and does not cover all 80 countries in our database. The GGEA index range from 0 to 1 wherein a score of 1 denotes the absence of disparity. This index intends to delineate the variance between women and men in access to education. It builds upon gender-based ratios in enrollment and participation rates across primary, secondary, and tertiary educational tiers.

We divide countries into two strata as follows: for each country, we compute the average value of the index on the time span, then set as threshold the median on this average (time-invariant) index across all countries.[9] To illustrate, Chad scores the lowest (.52) on the index, while Lesotho scores the highest (1). A total of 40 countries are classified as having a high level of gender equality in educational attainment, with an average index above .93, while a further 28 countries are classified as having a low level of gender equality in educational attainment.

4.5 Outcome variables

We define three main dependent variables to assess the representation of women in publishing faculty and their contribution to local science output and quality.

First, we focus on scientific output and articles published, by defining as dependent

non-unique number of entries, while the statistics reported here differs because we report them for unique authors.

[9]We perform a similar exercise setting the threshold at the mean instead of the median. The analysis' outcomes are discussed in Section 6.

variable the share of publications with at least one female author. Second, we construct a dependent variable that measures gender composition in the publishing workforce, namely, the share of female academics among all local authors. As higher publication output might come at the expense of higher research quality, we build a third outcome variable that captures and approximates the contribution of women authors to the average quality of local science output. The construction of these variables originated from the (enriched) MAG database at the paper-author-affiliation-publication date level.

Specifically, as first dependent variable, we focus on the share of publications authored by at least one woman (SPW):

$$SPW_{i,f,t} = \left(\frac{\sum_{p=1}^{P} Papers\ with\ (at\ least)\ one\ woman_{i,f,t}}{\sum_{p=1}^{P} Papers_{i,f,t}} \right) \tag{1}$$

This variable is constructed on the paper-author level. It distinguishes between papers co-authored and single-authored papers by women faculty vis-à-vis all publications by publishing faculty in a given field f, institution i, and year t. We note that shares likely also capture underlying changes in team composition and not only the changes in the contribution of women to local output.

Hence, the second variable is defined as the share of unique women authors (SWA) among all publishing authors in a given institution and science field as follow:

$$SWA_{i,f,t} = \left(\frac{\sum_{a=1}^{A} Women\ Authors_{i,f,t}}{\sum_{a=1}^{A} Authors_{i,f,t}} \right) \tag{2}$$

Where $Women\ Authors_{i,f,t}$ relates to the institution i in research field f at time t (for the academic body denoted as a, where a ranges from 1 to A). This variable approximates overall gender composition in academic staff, assuming that (most) researchers in the local faculty actively publish and are also credited as authors for their contribution to publications.

Moreover, we generate yet another outcome variable measuring the relative contribution of women to publication quality vis-à-vis the quality of all publications by local faculty. This exploits available information on the journal impact factor for each published paper and the associated outlet. As a first step, we define the average impact factor in local publications listing female authors as:

$$Average\,Women's\;JIF_{i,f,t} = \frac{\left(\sum_{p=1}^{P} Women\;JIF_{p,i,f,t,}\right)}{\left(\sum_{p=1}^{P} Women\;in\;Papers_{p,i,f,t}\right)} \quad (3)$$

where the nominator is given by the yearly sum of impact factors for publications (associated outlets) that women (co-)authored at the institution-field level. The denominator corrects and weights $Women\;JIF$ by the yearly sum of listed women authors as contributors to these publications. The average impact factor of publications with male authors is constructed analogously.

In the next step, we define the relative quality contribution by women to publications (RPQ) as:

$$RPQ_{i,f,t} = \frac{\left(Average\,Women's\;JIF_{i,f,t}\right)}{\left(Average\,Women's\;JIF_{i,f,t} + Average\,Men's\;JIF_{i,f,t}\right)} \quad (4)$$

The third dependent variable represents a proxy of the relative contribution of women scientists to the average publication quality in a given institution and scientific field, weighted by the average quality of publications authored by all faculty.

4.6 Descriptive statistics

Table 2 reports the statistics for the entire dataset, while Tables 3 to 8 split the data and separate out samples for (i) scientific fields (not) covered by the Hinari program, (ii) observations before and after the creation of the Hinari program in 2002, and (iii) treated and control institutions, i.e., institutions that eventually joined Hinari at some point in time as compared to those that never did in the observation period.

First, Table 2 highlights that, for the full sample, roughly 49% of the papers include at least one female author (SPW), with a minimum female participation equal to zero, and a maximum female participation in the scientific productivity in scientific publications with only female authors. Focusing on the average number of authors, the average publishing faculty has 34 associated researchers (i.e., authors listing the affiliation), out of which 28 percent are women, namely, the share of women among authors (SWA). The smallest institution in our dataset employs just one actively publishing researcher, while the biggest amounts to 2,511 researchers publishing in a given year. On average, women authors publish articles in journal outlets of slightly lower average quality than men, i.e., the average JIF of women authors sits at 1.53 as compared to 1.68 for male authors. Still, their research shows a higher JIF maximum value, i.e., their journal outlets are cited up to 154 citations against just 63 for male authors.

As more women graduate from health-related fields than any other field, we observe in our data sample that women represent 31 percent of authors in those fields against 25 percent in other non-health related fields (Tables 3 and 4). Similarly, on average, the share of papers with at least one woman is higher in health-related fields of study. In addition, articles published in health fields of study also expect a greater number of citations and have higher average JIFs associated with outlets (1.36 vs 1.08). Furthermore, the average impact factor of published research listing women authors is higher compared to the one in non-health fields.

Tables 5 and 6 provide evidence on the evolution over time and compares the two sub-periods of interest, i.e., 1990–2002 and 2002–2018. The number of authors increases over time, which is aligned with global trends. The share of papers with at least one woman rises from 33 to 53 percent. Similarly, the share of women among authors grows from about 22 percent to 29 percent over time. Still, despite the general increase in male and female contributions to average publication quality, the relative contribution of women authors (RPQ) stays on similar levels in the two periods.

Comparing institutions eventually joining the Hinari program (treated) to those never joining it (controls) in Tables 7 and 8, it appears that treated units are twice as big on average, with 41 actively publishing researchers in treated units against 26 researchers hosted in control units. The gender distribution and average number of women is different in treated and control groups, with 14 and 8 percent, respectively. Moreover, treated institutions have a slightly lower share of papers with at least one woman. Still, treated and untreated institution-fields do not show a different average value concerning the share of women among authors (SWA), and with regard to the relative journal impact factor (RPQ).

The descriptive statistics suggest that the average institution that joined the program is larger and typically has more active authors, both female and male academics. However, we note that treated and control units do not substantially differ in the gender-related dimensions we adopt in the empirical analysis.

5 Empirical Strategy

5.1 Identification strategy

In order to assess the impact of the Hinari program on gender participation and contribution of women to science in developing countries, we set as the unit of analysis an *institution-field* in a given year, where the field is broadly define as related to biomedical and health, or not. The rationale behind such a unit of analysis is that, although institutions may differ substantially in their propensity to join the program, only scholars in the relevant fields are affected by the program. One one hand, institutions are most likely not enrolling into the program at random, but rather learn about the initiative and, based on the relevance of the program with regards to their needs and resources, decide or not to undertake the enrollment procedure. Therefore, institutions that choose to participate in the program may differ significantly from other institutions, for example in terms of their size, research production, ties with other countries, or even their interest and efforts in addressing sustainable development goals, including gender equality. On the other hand, by dividing institutions into the group of scholars working in the health-related fields, which are targeted and affected by the Hinari program, and the group of scholars conducting research in non-health-related fields, we may be able control for the potential selection bias at the institutional level.

A triple difference model allows us to implement our empirical strategy and properly assess the impact of the Hinari program on women's representation and contribution to local science (Gruber, 1994; Olden and Møen, 2022). We compare the treatment institutions (i.e., those enrolled in the Hinari program) in the health field to a group of control institutions (never treated or not-yet-treated) in the same field and measure the change in the treatment outcomes relative to the non-health field. Formally, we estimate the following equation:

$$y_{ift} = \beta_0 + \beta_1 Hinari\ Field_{ift} \times Hinari\ Institution_{ift} \times Post\ Hinari_{ift} + \beta_2 X_{ift} + \gamma_c + \delta_t + \rho_{ct} + \epsilon_{ift} \quad (5)$$

Where if is an institution-field observed at time t, and $Hinari\ Field$, $Hinari\ Institution$, and $Post\ Hinari$ are a set of dummies indicating the health field, in the treated institutions, and years after treatment, respectively. y_{ift} is one of the three dependent variables defined in Section 4.5, namely SPW, or SWA or RPQ. X_{ift} is a set of time-varying control variables that further control for potential selection bias. γ_c is a city or country fixed effect (depending on the specification), δ_t is a time fixed effect, and ρ_{ct} is a city or country time trend included in the most sophisticated specifications.

The main coefficient of interest is β_1, which captures the impact of enrolling in the Hinari program on gender equality in health fields conditional on the validity of our model.

Several circumstances may compromise the validity of our estimations. If there are pre-existing differences in our main measures of gender representation and contribution to science, then our model cannot properly disentangle the program's impact from pre-existing trends. We use event studies to verify the parallel trend assumption for all of our dependent variables. Another confounding factor may be the collaborative nature of science. Since the contribution of individual authors is difficult to determine, we use equal weights for each co-author on papers in the main analysis.

A third challenge may arise from academics needing time to learn and digest research findings from the body of prior knowledge made available to them through the Hinari program. We use dynamic estimators to account for the institutions' staggered adoption of the program over time (Borusyak, Jaravel and Spiess, 2024). The economic intuition behind these alternative approaches is that there could be a temporal dimension of knowledge absorption and learning that models should account for. This potentially means that the program's impact on gender composition and the contribution of women authors could be delayed or might only gradually unfold over time.

As highlighted in Goodman-Bacon (2021), in a staggered difference-in-differences setup the coefficient of interest is computed as a weighted average of all possible (2x2) comparisons. Negative weights in the coefficient estimates could arise since the two-way fixed effect model identifies weighted sums of the average treatment effects in each group and period. In other words, the $\hat{\beta}$ coefficient is computed comparing the not-yet-treated groups, and the already-treated groups, and the estimates could lead to negative weights.[10] In this regards, a recent strand of the literature highlights the issue and proposes several ways to deal with the problem (De Chaisemartin and D'Haultfœuille, 2020; Callaway and Sant'Anna, 2021; Goodman-Bacon, 2021; Sun and Abraham, 2021; Athey and Imbens, 2022; Borusyak, Jaravel and Spiess, 2024).

Finally, our models includes a set of variables controlling for institutions' different propensity to join Research4Life. At the institution-field level, we control for: the average number of co-authorships with European scholars; the average number of co-authorships with US scholars; the average number of co-authorships with scholars in the same world region;[11] the average number of co-authorships with scholars in the same world region in an institution-field enrolled in the Hinari program; the average number of authors (in log terms); and a set of dummies indicating the subscription of the research institution in the other subprograms if the R4L initiative.

6 Results

6.1 Baseline results

Table 9 reports the coefficients from the triple difference model on the Hinari effect on the share of papers with at least one woman (SPW). From column (1) to column (8), we add the control variables, several fixed effects, and country or city time trends. Overall, we

[10]For instance, Sun and Abraham (2021) show that, in the case of variation in treatment across units, the regression coefficients are not robust to the heterogeneous or dynamic treatment effects across group and time.

[11]World regions are defined using the World Bank definition. Specifically: East Asia and Pacific; Europe and Central Asia; Latin America and Caribbean; Middle East and North Africa; North America; South Asia; and Sub-Saharan Africa.

find a positive and significant impact of the program on female academic production. The estimated impact ranges from roughly 20% in the less demanding model specification to 3% in more demanding ones. Table 10 shows the coefficient estimates once we adopt the share of women among authors (SWA) as dependent variable. The outcomes are positive and significant, ranging from about 2% to 9%. The coefficients are stable and slightly significant, also in the most demanding specifications. In Table 11, we investigate the impact of the Hinari initiative on the relative female contribution to overall publication quality (RPQ). Column (1) presents a significant coefficient, roughly 2%. Adding controls and fixed effects, the estimates remain stable, but barely significant with the inclusions of the full set of fixed effects, and not significant once we include the geographical time trends.

Regardless of the outcome variable, the results demonstrate a positive and significant impact of the Hinari program on the academic gender balance and greater participation of women researchers in local scientific production and research quality. Results are consistent with the idea that women researchers are ex ante more mobility constrained and exposed to human bias then men, and they might therefore benefit relatively more from anonymous and remote digital access to resources provided through the program.

In the heterogeneity section, we discuss characteristics that potentially existed before institutions enrolled in the program such as the gender education gap at country level. This could help to explain the differences in production and quality levels we observe in the data. The next section examines the robustness of main findings by verifying key assumptions of the triple difference model and by using alternative modelling approaches.

6.2 Robustness

Event study and dynamic estimators

As noted in the empirical strategy, the identification of effects rests on the assumption that, conditional on the set of controls and fixed effects, the adoption of the Hinari program is not related to pre-existing differences between treated and untreated institution-fields. Accordingly, we test for pre-existing differences in our gender metrics that could potentially invalidate our estimates.

Moreover, the staggered adoption of the Hinari program at the institution level represents one of the major concerns related to our baseline estimates. To validate the results discussed in Section 6, we allow for treatment heterogeneity by presenting results using recent estimators proposed in the literature by Borusyak, Jaravel and Spiess (2024). To the best of our knowledge, this estimator is the only one in the current literature which deals with the triple difference approach by adopting an imputation estimator approach.

Figure 2 presents the institution-field event study graph that plots the evolution of yearly changes in the share of papers with at least one woman (SPW). The model includes all institution-field level controls as well as country, year, institution, and research-field fixed effects. Plotted estimates do not show systematic differences in female productivity between treated and control at the institution-field level. On the other hand, the program adoption appears to have a barely positive and significant effect roughly ten years after the introduction (confidence interval at 90%).

In Figure 3, we adopt the same model, but present event study results for the share of women among authors (SWA) as the dependent variable. Again, we take note that point estimates do not exhibit a pre-trend. However, we observe a positive and strongly significant impact (at 95% confidence interval) of the program after seven years of adoption. This result is consistent with the idea that any gains in academic productivity associated with program adoption and improved access might require learning and knowledge absorption.

Focusing on the female quality contribution metric (RPQ), Figure 4 reveals no pre-trend issues. Furthermore, coefficient estimates show a significant and positive impact from three years after the initial Hinari program adoption. According to the graph, we observe a positive impact of the initiative at year three and five, hence again the coefficient are positive and significant only ten years after Hinari program adoption.

Taken together, the evidence presented in this section corroborates our main results and rules out the idea that institution-field level trends explain systematic differences on gender metrics.

6.3 Continuous treatment

In this section, we exploit the data regarding actual usage and online logins to the Hinari online platform instead of the timing of Hinari program adoption. In the model specification 5, we replace the binary treatment variable indicating the program adoption by the number of logins to the Hinari online platforms.[12]

Table 12 presents the outcome for the share of papers with at least one woman (SPW) as dependent variable. The results are positive, significant and stable in all model specifications. Hence, a 10-percentage-point increase in actual usage implies a positive and significant impact of the program resources on the share of papers with at least one woman of 4.6% (see Column (8)). In other words, given the adoption of the program, a 10-percentage-point increase in login access to the relevant study literature implies an increase in the share of papers with at least one woman from 48.8% to 51%.[13]

Table 13 presents the results for the share of women among authors (SWA). These results are less stable in significance compared to the ones discussed above. This may be due to the downward bias of the continuous treatment variable. According to the model specification (8), a 10-percentage-point increase of actual usage and logins recorded relates to a 3.4% increase of the share of women among authors.

Despite the limited data coverage for program resources usage, these findings corroborate the baseline estimates discussed in the previous Section 6. Moreover, we duly note that results derived from restricted usage data sample may suffer from downward bias. In turn, by observing zero logins prior to the year 2007 also in treated institution-fields, we may expect higher magnitude in the estimated coefficients.

More broadly, general outcomes in this section complement earlier findings extensively discussed in Cuntz et al. (2024), indicating that journals included in the Hinari program are, in effect, also more heavily cited in local publications after program adoption.

[12]We define this variable as the logarithm of one plus the number of logins to the Hinari platform.
[13]See the variable's mean in Table 2.

6.4 Heterogeneity

Time-varying, country-level characteristics could imply a heterogeneous impact beyond what country-fixed effects can absorb in previous models. In this subsection, we investigate whether pre-existing gender differences in educational attainment across countries could lead to different program impacts.

Table 14 presents the coefficient estimates for the share of papers with at least one woman (SPW) as the dependent variable, by dividing the sample using the median of the Gender Gap Educational Attainment (see Section 4.4). Interestingly, the Hinari impact appears statistically significant for the treated institutions-fields located in countries with lower female attainment and smaller gender gaps in education, respectively.

Similarly, Table 15 studies the impact of the program on the share of papers with at least one woman (SPW), by dividing the sample by the mean of the gender educational attainment gap instead of the median. The results are similar in magnitude and significance compared to the ones discussed in Table 15. That is, the Hinari initiative mainly affects countries with a minor gender gap in educational attainment. These results highlight how gender equality achievements in educational attainment by women can be further amplified through programs like Hinari.[14]

Seemingly, promoting gender equality at different levels of education, in parallel with providing local researchers with free or low-cost access to knowledge, can lead to a greater impact on gender imbalances in academia. In this way, e.g., by providing higher-quality teaching and better resources to students, improved knowledge access to local faculty seems to also help overcome gaps in educational attainment and activate the potential of less-educated women in academic labor markets, effectively reducing the loss of domestic women talent along the leaky pipeline.

[14] We also perform a similar analysis for the SWA and RPQ as dependent variables. The results do not display significant coefficients but are available upon request from the authors.

7 Discussion

Our main research findings are as follows. Causal effects demonstrate that free or low-cost access to online scientific resources through the UN-led Hinari program has yielded significant positive outcomes for researchers located in developing countries, particularly regarding gender equality in scientific publishing. The findings reveal a substantial improvement in local gender composition among publishing faculty, alongside moderate but meaningful increases in both women's contributions to scientific output and average publication quality. This is consistent with the idea that women researchers are ex ante more mobility constrained and exposed to human bias then men when accessing resources. Hence, the evidence shows that they benefit relatively more from anonymous and remote digital access to knowledge resources provided through the program. Effects are robust to alternative model specifications, dynamic effect estimators and continuous treatment based on actual usage of Research4Life resources. Still, notably, effects only gradually emerge and manifest over time, consistent with established theories of knowledge absorption and learning. Taken together, our main findings show how reduced online barriers to scientific information can contribute and help improve participation of researchers in more marginalized positions and less-well resourced places in global science as in the case of women scientists in developing countries.

Second, while our research design addresses potential bias arising from institutional selection into the Hinari program, results could still be challenged on the basis that selection might also happen on the publishers' side and via the resources they provide. Arguably, rather than the true effect of better access to prior knowledge, our estimates might reflect changes in the publishers involved, or the selection of journals and articles made available by them under the program. Companion research to this paper shows that this is unlikely to be a concern and more systematically studies this issue (Cuntz et al., 2024).[15]

[15] Hinari publishers account for approximately 65 percent of all possible titles and hence include most of the major publishers operating in the global scholarly publishing market. Moreover, tests comparing characteristics of journals (not) included in the program indicate that, at least for top journals in biomedical and health sciences, there is no selection bias resulting from the inclusion of higher-quality journals and larger outlets by publishers contributing to the Hinari program. Finally, based on interviews conducted with Hinari program owners and WHO officials, we can also rule out systematic selection bias

Third, we note that the data collected and analysis of program impact also lends itself to study effects on the level of the individual researcher and their career progression in academia. By changing the level of the analysis, one could explore how academic networks of women possibly change relative to men's, or whether or not there are changes in citation behavior among women and men after program adoption. This could provide further insights into the underlying mechanisms and potential career prospects, e.g., whether or not access to prior science effectively lowers search cost and helps the identification of potential research collaborators worldwide and might redirect some women to pursue other research areas. We leave these interesting dimensions for future research.

Finally, the bibliometric data used to explore gender composition is not without limits. We note that baseline measurement is constrained to women involved and credited in published research. On the one hand, this will exclude some women who are active researchers but might not be credited for their contribution to publications. On the other hand, women among local faculty might also be working in ancillary technical functions in research projects, but, given the data at hand, are also disregarded in the current approach. With new and more granular research project data becoming available, future research can also help overcome this measurement issue.

8 Conclusion

In this paper, we document the effect of free or low-cost online access to science through the UN-led Hinari program on follow-on science and academic gender balance in developing countries.

The evidence points to a positive and significant impact of the Hinari program on local gender composition in publishing faculty, and moderate increases in the contribution of women scientists to scientific output and average publication quality. Causal effects are shown to be robust using dynamic estimators, event study and lead time analysis,

at the article level with regard to state-of-the-art research and researchers' information access to the scientific frontier. As explained by WHO officials, "every journal included [in Hinari] does provide access to the present time, unless the journal is no longer available from the publisher or has ceased publication. [..] If a publisher includes a journal in Hinari, it will be all [author emphasis] the articles in that journal within the extent time period of year, volume, issues that they have included in their commercial offer."

but they also suggest that effects are slow to unfold over time, which is in line with knowledge absorption and learning theories. Consistent with the above argument, we find a substantial increase in the share of women among all authors as well as the share of papers with at least one woman only several years after programs have been adopted by local research institutions. The differential impact on women as compared to men supports the idea that women researchers are ex ante more mobility constrained and exposed to human bias then men when accessing knowledge resources.

More broadly, our research reveals that policies mandating free or low-cost and online access to codified knowledge can not only induce follow-on science. They also play a crucial role in reducing gender disparities in local science, thereby contributing to the United Nations' Sustainable Development Goals. More specifically, notable differences in program impact emerge for countries with low and high gender balance in educational attainment. Precisely, the Hinari initiative appears more effective in countries where the educational attainment gender gap is already relatively low. These findings provide important takeaways for policy-makers and stakeholders of the R4L initiative: Better access to prior knowledge holds the potential to improve and make science production more inclusive, by activating the potential of women along the entire leaky pipeline, from earlier educational to later science career stages.

References

Aghion, P, L Boustan, C Hoxby and J Vandenbussche. 2009. "The Causal Impact of Education on Economic Growth: Evidence from US.".

Agrawal, Ajay and Avi Goldfarb. 2008. "Restructuring Research: Communication Costs and the Democratization of University Innovation." *American Economic Review* 98(4):1578–90.

Alper, Joe. 1993. "The pipeline is leaking women all the way along." *Science* 260(5106):409–411.

Andrews, Michael. 2019. "Bar talk: Informal social interactions, alcohol prohibition, and invention." *Alcohol Prohibition, and Invention (November 18, 2019)* .

Athey, Susan and Guido W. Imbens. 2022. "Design-based analysis in Difference-In-Differences settings with staggered adoption." *Journal of Econometrics* 226(1):62–79.

Azoulay, Pierre, Jeffrey L Furman, Krieger and Fiona Murray. 2015. "Retractions." *Review of Economics and Statistics* 97(5):1118–1136.

Barley, Stephen R. 1986. "Technology as an occasion for structuring: Evidence from observations of CT scanners and the social order of radiology departments." *Administrative science quarterly* pp. 78–108.

Berkes, Enrico and Peter Nencka. 2019. 'Novel'Ideas: The Effects of Carnegie Libraries on Innovative Activities. In *2019 Meeting Papers.* Number 1315 Society for Economic Dynamics.

Berkes, Enrico and Peter Nencka. 2021. "Knowledge access: The effects of Carnegie libraries on innovation." *Available at SSRN 3629299* .

Biasi, Barbara and Petra Moser. 2021. "Effects of copyrights on science: Evidence from the WWII book republication program." *American Economic Journal: Microeconomics* 13(4):218–60.

Bikard, Michaël, Isabel Fernandez-Mateo and Ronak Mogra. 2025. "Standing on the Shoulders of (Male) Giants: Gender Inequality and the Technological Impact of Scientific Ideas." *Administrative Science Quarterly* 0(0).

Borusyak, Kirill, Xavier Jaravel and Jann Spiess. 2024. "Revisiting Event-Study Designs:

Robust and Efficient Estimation." *The Review of Economic Studies* 91(6):3253–3285. **URL:** *https://doi.org/10.1093/restud/rdae007*

Buckles, Kasey. 2019. "Fixing the Leaky Pipeline: Strategies for Making Economics Work for Women at Every Stage." *Journal of Economic Perspectives* 33(1):43–60.

Buffington, Catherine, Benjamin Cerf, Christina Jones and Bruce A Weinberg. 2016. "STEM training and early career outcomes of female and male graduate students: Evidence from UMETRICS data linked to the 2010 census." *American Economic Review* 106(5):333–338.

Butler, Daniel M, Richard J Butler and Jesse T Rich. 2008. "The equalizing effect of the internet on access to research expertise in political science and economics." *PS: Political Science & Politics* 41(3):579–584.

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

Carlana, Michela. 2019. "Implicit stereotypes: Evidence from teachers' gender bias." *The Quarterly Journal of Economics* 134(3):1163–1224.

Carpentier, Elodie and Julio Raffo. 2023. "The Global Gender Gap in Innovation and Creativity: An International Comparison of the Gender Gap in Global Patenting over Two Decades." *WIPO Development Studies Series* .

Ceci, Stephen J, Donna K Ginther, Shulamit Kahn and Wendy M Williams. 2014. "Women in academic science: A changing landscape." *Psychological science in the public interest* 15(3):75–141.

Ceci, Stephen J, Wendy M Williams and Susan M Barnett. 2009. "Women's underrepresentation in science: sociocultural and biological considerations." *Psychological bulletin* 135(2):218.

Chang, Wan-Ying, Wei Cheng, Julia Lane and Bruce Weinberg. 2019. "Federal funding of doctoral recipients: What can be learned from linked data." *Research Policy* 48(6):1487–1492.

Cuntz, Alexander, Frank Müeller-Langer, Alessio Muscarnera, Prince C. Oguguo and Marc Scheufen. 2024. "Access to science and innovation in the developing world." *WIPO - Economic Research Working Paper* (78).

Cuntz, Alexander and Matthias Sahli. 2024. “Intermediary liability and trade in follow-on innovation.” *Journal of Cultural Economics* 48(1):1–42.

Currie, Janet and Enrico Moretti. 2003. “Mother’s Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings.” *The Quarterly Journal of Economics* 118(4):1495–1532.

De Chaisemartin, Clément and Xavier D’Haultfœuille. 2020. “Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects.” *American Economic Review* 110(9):2964–96.

Ding, Waverly W, Sharon G Levin, Paula E Stephan and Anne E Winkler. 2010. “The impact of information technology on academic scientists’ productivity and collaboration patterns.” *Management Science* 56(9):1439–1461.

Duch, Jordi, Xiao Han T Zeng, Marta Sales-Pardo, Filippo Radicchi, Shayna Otis, Teresa K Woodruff and Luis A Nunes Amaral. 2012. “The possible role of resource requirements and academic career-choice risk on gender differences in publication rate and impact.” *PloS one* 7(12):e51332.

Furman, Jeffrey L, Markus Nagler and Martin Watzinger. 2021. “Disclosure and subsequent innovation: Evidence from the patent depository library program.” *American Economic Journal: Economic Policy* 13(4):239–70.

Furman, Jeffrey L. and Megan J. MacGarvie. 2007. “Academic science and the birth of industrial research laboratories in the U.S. pharmaceutical industry.” *Journal of Economic Behavior & Organization* 63(4):756–776.

Furman, Jeffrey L. and Scott Stern. 2011. “Climbing atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research.” *American Economic Review* 101(5):1933–63.

Ginther, Donna K and Shulamit Kahn. 2009. Does science promote women? Evidence from academia 1973-2001. In *Science and engineering careers in the United States: An analysis of markets and employment*. University of Chicago Press pp. 163–194.

Giorcelli, Michela and Petra Moser. 2020. “Copyrights and Creativity: Evidence from Italian Opera in the Napoleonic Age.” *Journal of Political Economy* 128(11):4163–4210.

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

Graddy-Reed, Alexandra, Lauren Lanahan and Jesse D'Agostino. 2021. "Training across the academy: The impact of R&D funding on graduate students." *Research Policy* 50(5):104224.

Gruber, Jonathan. 1994. "The incidence of mandated maternity benefits." *The American economic review* pp. 622–641.

Hofstra, Bas, Vivek V Kulkarni, Sebastian Munoz-Najar Galvez, Bryan He, Dan Jurafsky and Daniel A McFarland. 2020. "The diversity–innovation paradox in science." *Proceedings of the National Academy of Sciences* 117(17):9284–9291.

Huang, Junming, Alexander J Gates, Roberta Sinatra and Albert-László Barabási. 2020. "Historical comparison of gender inequality in scientific careers across countries and disciplines." *Proceedings of the National Academy of Sciences* 117(9):4609–4616.

Hussinger, Katrin and Lorenzo Palladini. 2024. "Information accessibility and knowledge creation: the impact of Google's withdrawal from China on scientific research." *Industry and Innovation* pp. 1–31.

Iaria, Alessandro, Carlo Schwarz and Fabian Waldinger. 2018. "Frontier Knowledge and Scientific Production: Evidence from the Collapse of International Science." *The Quarterly Journal of Economics* 133(2):927–991.

Kaiser, Franziska, Alexander Cuntz and Christian Peukert. 2023. "Batman forever? The role of trademarks for reuse in the US comics industry." *Research Policy* 52(8):104820.

Kim, E Han, Adair Morse and Luigi Zingales. 2009. "Are elite universities losing their competitive edge?" *Journal of Financial Economics* 93(3):353–381.

Koffi, Marlène and Matt Marx. 2023. Cassatts in the Attic. Technical report National Bureau of Economic Research.

Kong, Dongmin, Chen Lin, Lai Wei and Jian Zhang. 2022. "Information accessibility and corporate innovation." *Management Science* 68(11):7837–7860.

Koning, Rembrand, Sampsa Samila and John-Paul Ferguson. 2021. "Who do we invent for? Patents by women focus more on women's health, but few women get to invent." *Science* 372(6548):1345–1348.

Larivière, Vincent, Chaoqun Ni, Yves Gingras, Blaise Cronin and Cassidy R Sugimoto. 2013. "Bibliometrics: Global gender disparities in science." *Nature* 504(7479):211–213.

Lax-Martínez, Gema, Helena Saenz de Juano-i Ribes, Deyun Yin, Bruno Le Feuvre, Intan Hamdan-Livramento, Kaori Saito and Julio Raffo. 2021. "Expanding the World Gender-Name Dictionary: WGND 2.0." *WIPO working papers* .

Liu, Junwan, Yinglu Song and Sai Yang. 2020. "Gender disparities in the field of economics." *Scientometrics* 125:1477–1498.

Marwell, Gerald, Rachel Rosenfeld and Seymour Spilerman. 1979. "Geographic constraints on women's careers in academia." *Science* 205(4412):1225–1231.

Marx, Matt and Aaron Fuegi. 2020. "Reliance on science: Worldwide front-page patent citations to scientific articles." *Strategic Management Journal* 41(9):1572–1594.

Merton, Robert K. 1973. *The sociology of science: Theoretical and empirical investigations*. University of Chicago press.

Mokyr, Joel. 2011. *The gifts of Athena: Historical origins of the knowledge economy*. princeton university press.

Murray, Fiona and Scott Stern. 2007. "Do formal intellectual property rights hinder the free flow of scientific knowledge?: An empirical test of the anti-commons hypothesis." *Journal of Economic Behavior & Organization* 63(4):648–687.

Müeller-Langer, Frank, Marc Scheufen and Patrick Waelbroeck. 2020. "Does online access promote research in developing countries? Empirical evidence from article-level data." *Research Policy* 49(2).

Nagaraj, Abhishek. 2018. "Does Copyright Affect Reuse? Evidence from Google Books and Wikipedia." *Management Science* 64(7):3091–3107.

Olden, Andreas and Jarle Møen. 2022. "The triple difference estimator." *The Econometrics Journal* 25(3):531–553.

Orlikowski, Wanda J. 1992. "The duality of technology: Rethinking the concept of technology in organizations." *Organization science* 3(3):398–427.

Partha, Dasgupta and Paul A. David. 1994. "Toward a new economics of science." *Research Policy* 23(5):487–521.

Raffo, Julio. 2021. "WGND 2.0.".

URL: *https://doi.org/10.7910/DVN/MSEGSJ*

Raffo, Julio and Gema Lax-Martínez. 2018. "WGND 1.0.".

URL: *https://doi.org/10.7910/DVN/YPRQH8*

Rose, Michael E and Co-Pierre Georg. 2021. "What 5,000 acknowledgements tell us about informal collaboration in financial economics." *Research Policy* 50(6):104236.

Scotchmer, Suzanne. 1991. "Standing on the Shoulders of Giants: Cumulative Research and the Patent Law." *Journal of Economic Perspectives* 5(1):29–41.

Shauman, Kimberlee A and Yu Xie. 1996. "Geographic mobility of scientists: Sex differences and family constraints." *Demography* 33(4):455–468.

Sproull, Lee, Sara Kiesler and Sara B Kiesler. 1991. *Connections: New ways of working in the networked organization.* MIT press.

Stephan, Paula E. 2010. The economics of science. In *Handbook of the Economics of Innovation.* Vol. 1 Elsevier pp. 217–273.

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

Visser, Martijn, Nees Jan van Eck and Ludo Waltman. 2021. "Large-scale comparison of bibliographic data sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic." *Quantitative Science Studies* 2(1):20–41.

Winkler, Anne E, Sharon G Levin and Paula E Stephan. 2010. "The diffusion of IT in higher education: Publishing productivity of academic life scientists." *Economics of Innovation and New Technology* 19(5):481–503.

Zheng, Yanfeng and Qinyu Wang. 2020. "Shadow of the great firewall: The impact of Google blockade on innovation in China." *Strategic Management Journal* 41(12):2234–2260.

Appendix

Fig. 1: Number of institutions enrolled in Hinari each year

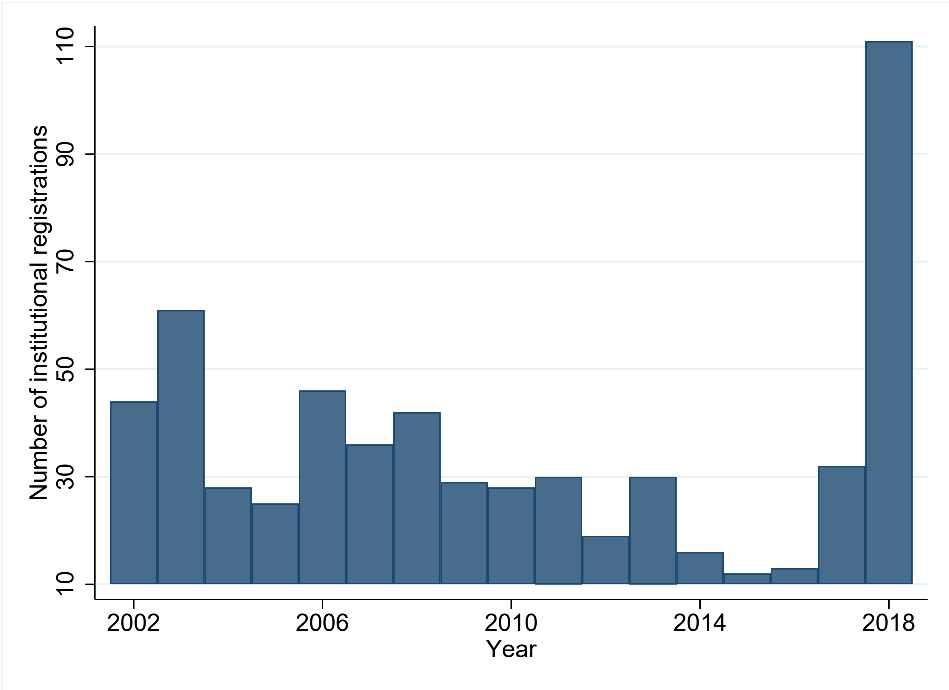


Table 1: *Genderit* command results

Most likely gender		
	Freq.	Percent
?	671,393	27.98
F	586,559	24.44
M	1,034,288	43.10
U	107,444	4.48
Total	2,399,684	100.00

Note: the table reports the Stata *genderit* command output.The probability threshold for gender assignment is ≥ 0.8.

Table 2: Summary statistics

	mean	sd	min	max	count
SPW	48.81	32.35	0.00	100.00	29,522
Nr. authors	34.68	103.88	1.00	2,511.00	29,522
Nr. women authors	11.51	44.60	0.00	1,354.00	29,522
Nr. men authors	23.17	62.88	0.00	1,269.00	29,522
SWA	27.83	26.85	0.00	100.00	29,522
Average Women's JIF	1.53	2.04	0.00	153.54	21,171
Average Men's JIF	1.68	1.86	0.00	62.59	27,830
RPQ	0.44	0.18	0.00	1.00	19,314
Triple DiD	0.14	0.35	0.00	1.00	29,522
Triple DiD (logins)	0.38	1.52	0.00	11.19	29,522
Hinari program adoption	0.29	0.45	0.00	1.00	29,522
Logins	118.34	1,303.93	0.00	72,309.00	29,522
Avg JIF	1.21	1.08	0.00	21.69	29,325
Avg nr. authors	13.43	103.86	1.00	3,056.50	29,522
Avg non-R4L countries co-authorship	0.33	0.32	0.00	1.00	29,522
Avg US coauthorship	0.11	0.21	0.00	1.00	29,522
Avg EU coauthorship	0.16	0.25	0.00	1.00	29,522
Agora program adoption	0.22	0.42	0.00	1.00	29,522
Ardi program adoption	0.11	0.31	0.00	1.00	29,522
Goali program adoption	0.04	0.18	0.00	1.00	29,522
Oare program adoption	0.19	0.40	0.00	1.00	29,522
Global Gender Gap Educational Attainment Subindex	0.90	0.09	0.47	1.00	18,035

Table 3: Summary statistics - Hinari field of study

	mean	sd	min	max	count
SPW	57.02	32.21	0.00	100.00	13,673
Nr. authors	33.92	104.01	1.00	2,510.00	13,673
Nr. women authors	12.62	49.13	0.00	1,295.00	13,673
Nr. men authors	21.30	58.34	0.00	1,269.00	13,673
SWA	30.52	27.75	0.00	100.00	13,673
Average Women's JIF	1.82	2.50	0.00	153.54	9,944
Average Men's JIF	1.95	2.11	0.00	62.59	12,797
RPQ	0.45	0.18	0.00	1.00	9,012
Triple DiD	0.30	0.46	0.00	1.00	13,673
Triple DiD (logins)	0.82	2.15	0.00	11.19	13,673
Hinari program adoption	0.30	0.46	0.00	1.00	13,673
Logins	255.50	1,906.87	0.00	72,309.00	13,673
Avg JIF	1.36	1.15	0.00	18.89	13,583
Avg nr. authors	8.79	55.65	1.00	2,697.33	13,673
Avg non-R4L countries co-authorship	0.35	0.34	0.00	1.00	13,673
Avg US coauthorship	0.13	0.23	0.00	1.00	13,673
Avg EU coauthorship	0.17	0.26	0.00	1.00	13,673
Agora program adoption	0.23	0.42	0.00	1.00	13,673
Ardi program adoption	0.11	0.31	0.00	1.00	13,673
Goali program adoption	0.04	0.19	0.00	1.00	13,673
Oare program adoption	0.20	0.40	0.00	1.00	13,673
Global Gender Gap Educational Attainment Subindex	0.90	0.09	0.47	1.00	8,358

Table 4: Summary statistics - non-Hinari field of study

	mean	sd	min	max	count
SPW	41.72	30.74	0.00	100.00	15,849
Nr. authors	35.34	103.77	1.00	2,511.00	15,849
Nr. women authors	10.55	40.25	0.00	1,354.00	15,849
Nr. men authors	24.79	66.50	0.00	1,157.00	15,849
SWA	25.52	25.83	0.00	100.00	15,849
Average Women's JIF	1.27	1.47	0.00	25.22	11,227
Average Men's JIF	1.44	1.59	0.00	54.43	15,033
RPQ	0.43	0.19	0.00	1.00	10,302
Triple DiD	0.00	0.00	0.00	0.00	15,849
Triple DiD (logins)	0.00	0.00	0.00	0.00	15,849
Hinari program adoption	0.27	0.45	0.00	1.00	15,849
Logins	0.00	0.00	0.00	0.00	15,849
Avg JIF	1.08	1.01	0.00	21.69	15,742
Avg nr. authors	17.44	131.86	1.00	3,056.50	15,849
Avg non-R4L countries co-authorship	0.32	0.31	0.00	1.00	15,849
Avg US coauthorship	0.09	0.20	0.00	1.00	15,849
Avg EU coauthorship	0.15	0.24	0.00	1.00	15,849
Agora program adoption	0.22	0.41	0.00	1.00	15,849
Ardi program adoption	0.10	0.30	0.00	1.00	15,849
Goali program adoption	0.03	0.18	0.00	1.00	15,849
Oare program adoption	0.19	0.39	0.00	1.00	15,849
Global Gender Gap Educational Attainment Subindex	0.91	0.09	0.47	1.00	9,677

Table 5: Summary statistics - Before 2002

	mean	sd	min	max	count
SPW	33.02	32.19	0.00	100.00	5,920
Nr. authors	13.85	30.73	1.00	812.00	5,920
Nr. women authors	3.31	8.44	0.00	202.00	5,920
Nr. men authors	10.53	22.98	0.00	610.00	5,920
SWA	21.40	26.93	0.00	100.00	5,920
Average Women's JIF	0.77	0.96	0.00	17.41	3,419
Average Men's JIF	0.82	1.13	0.00	29.96	5,582
RPQ	0.44	0.19	0.00	1.00	3,014
Triple DiD	0.00	0.00	0.00	0.00	5,920
Triple DiD (logins)	0.00	0.00	0.00	0.00	5,920
Hinari program adoption	0.00	0.00	0.00	0.00	5,920
Logins	0.00	0.00	0.00	0.00	5,920
Avg JIF	0.62	0.66	0.00	14.98	5,874
Avg nr. authors	5.08	23.73	1.00	1,240.31	5,920
Avg non-R4L countries co-authorship	0.26	0.32	0.00	1.00	5,920
Avg US coauthorship	0.09	0.21	0.00	1.00	5,920
Avg EU coauthorship	0.13	0.25	0.00	1.00	5,920
Agora program adoption	0.00	0.00	0.00	0.00	5,920
Ardi program adoption	0.00	0.00	0.00	0.00	5,920
Goali program adoption	0.00	0.00	0.00	0.00	5,920
Oare program adoption	0.00	0.00	0.00	0.00	5,920

Table 6: Summary statistics - After 2002

	mean	sd	min	max	count
SPW	52.77	31.16	0.00	100.00	23,602
Nr. authors	39.91	114.56	1.00	2,511.00	23,602
Nr. women authors	13.56	49.48	0.00	1,354.00	23,602
Nr. men authors	26.34	69.01	0.00	1,269.00	23,602
SWA	29.45	26.59	0.00	100.00	23,602
Average Women's JIF	1.67	2.16	0.00	153.54	17,752
Average Men's JIF	1.89	1.95	0.00	62.59	22,248
RPQ	0.44	0.18	0.00	1.00	16,300
Triple DiD	0.17	0.38	0.00	1.00	23,602
Triple DiD (logins)	0.48	1.69	0.00	11.19	23,602
Hinari program adoption	0.36	0.48	0.00	1.00	23,602
Logins	148.02	1,456.82	0.00	72,309.00	23,602
Avg JIF	1.36	1.12	0.00	21.69	23,451
Avg nr. authors	15.53	115.45	1.00	3,056.50	23,602
Avg non-R4L countries co-authorship	0.35	0.32	0.00	1.00	23,602
Avg US coauthorship	0.12	0.22	0.00	1.00	23,602
Avg EU coauthorship	0.17	0.25	0.00	1.00	23,602
Agora program adoption	0.28	0.45	0.00	1.00	23,602
Ardi program adoption	0.13	0.34	0.00	1.00	23,602
Goali program adoption	0.04	0.21	0.00	1.00	23,602
Oare program adoption	0.24	0.43	0.00	1.00	23,602
Global Gender Gap Educational Attainment Subindex	0.90	0.09	0.47	1.00	18,035

Table 7: Summary statistics - Treated

	mean	sd	min	max	count
SPW	47.72	31.34	0.00	100.00	16,802
Nr. authors	41.02	123.79	1.00	2,511.00	16,802
Nr. women authors	13.82	54.26	0.00	1,354.00	16,802
Nr. men authors	27.20	73.98	0.00	1,269.00	16,802
SWA	27.65	25.97	0.00	100.00	16,802
Average Women's JIF	1.46	2.15	0.00	153.54	12,266
Average Men's JIF	1.61	1.84	0.00	62.59	15,970
RPQ	0.44	0.18	0.00	1.00	11,347
Triple DiD	0.24	0.43	0.00	1.00	16,802
Triple DiD (logins)	0.67	1.97	0.00	11.19	16,802
Hinari program adoption	0.50	0.50	0.00	1.00	16,802
Logins	207.92	1,723.04	0.00	72,309.00	16,802
Avg JIF	1.17	1.03	0.00	18.89	16,677
Avg nr. authors	13.44	102.23	1.00	2,744.18	16,802
Avg non-R4L countries co-authorship	0.34	0.32	0.00	1.00	16,802
Avg US coauthorship	0.11	0.21	0.00	1.00	16,802
Avg EU coauthorship	0.16	0.24	0.00	1.00	16,802
Agora program adoption	0.39	0.49	0.00	1.00	16,802
Ardi program adoption	0.19	0.39	0.00	1.00	16,802
Goali program adoption	0.06	0.24	0.00	1.00	16,802
Oare program adoption	0.34	0.47	0.00	1.00	16,802
Global Gender Gap Educational Attainment Subindex	0.91	0.09	0.47	1.00	9,868

Table 8: Summary statistics - Untreated

	mean	sd	min	max	count
SPW	50.24	33.57	0.00	100.00	12,720
Nr. authors	26.31	68.41	1.00	1,603.00	12,720
Nr. women authors	8.46	26.65	0.00	624.00	12,720
Nr. men authors	17.86	43.55	0.00	979.00	12,720
SWA	28.08	27.99	0.00	100.00	12,720
Average Women's JIF	1.62	1.87	0.00	41.21	8,905
Average Men's JIF	1.77	1.89	0.00	54.43	11,860
RPQ	0.43	0.19	0.00	1.00	7,967
Triple DiD	0.00	0.00	0.00	0.00	12,720
Triple DiD (logins)	0.00	0.00	0.00	0.00	12,720
Hinari program adoption	0.00	0.00	0.00	0.00	12,720
Logins	0.00	0.00	0.00	0.00	12,720
Avg JIF	1.26	1.15	0.00	21.69	12,648
Avg nr. authors	13.43	105.98	1.00	3,056.50	12,720
Avg non-R4L countries co-authorship	0.33	0.33	0.00	1.00	12,720
Avg US coauthorship	0.11	0.22	0.00	1.00	12,720
Avg EU coauthorship	0.15	0.25	0.00	1.00	12,720
Agora program adoption	0.00	0.00	0.00	0.00	12,720
Ardi program adoption	0.00	0.00	0.00	0.00	12,720
Goali program adoption	0.00	0.00	0.00	0.00	12,720
Oare program adoption	0.00	0.00	0.00	0.00	12,720
Global Gender Gap Educational Attainment Subindex	0.89	0.10	0.65	1.00	8,167

Table 9: Effect of Hinari program on the share of publications (co)authored by women (SPW)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Field × Treated × Post	19.9798***	8.7300***	8.7300***	3.3008***	3.3008***	3.3008***	3.3764***	3.3461***
	(0.99)	(1.09)	(1.10)	(1.02)	(1.02)	(1.03)	(1.01)	(1.02)
Controls		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes			Yes		Yes	
City FE			Yes			Yes		Yes
Country-specific time trend							Yes	
City-specific time trend								Yes
N	29389	29389	29389	29389	29389	29389	29389	29389
R^2	0.36	0.44	0.44	0.46	0.46	0.46	0.47	0.49

Notes : Significance: *** p<0.01, ** p<0.05, * p<0.10. OLS model specification. Standard errors are clustered at the institution-field level. Dependent variable: share of papers with at least one woman. Control variables include: mean of US co-authorship, mean of EU co-authorship, mean of co-authorship in the same world region, mean of co-authorship in the same world region of institution in R4L program, log of the average number of authors per scientific publication, dummies indicating other R4L subprograms adoption years.

Table 10: Effect of Hinari program on the share of women authors (SWA)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Field × Treated × Post	9.1208***	4.9131***	4.9131***	1.8766**	1.8766**	1.8766**	1.8095*	1.6134*
	(0.90)	(0.96)	(0.97)	(0.94)	(0.94)	(0.95)	(0.93)	(0.95)
Controls		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes			Yes		Yes	
City FE			Yes			Yes		Yes
Country-specific time trend							Yes	
City-specific time trend								Yes
N	29389	29389	29389	29389	29389	29389	29389	29389
R^2	0.32	0.33	0.33	0.34	0.34	0.34	0.35	0.37

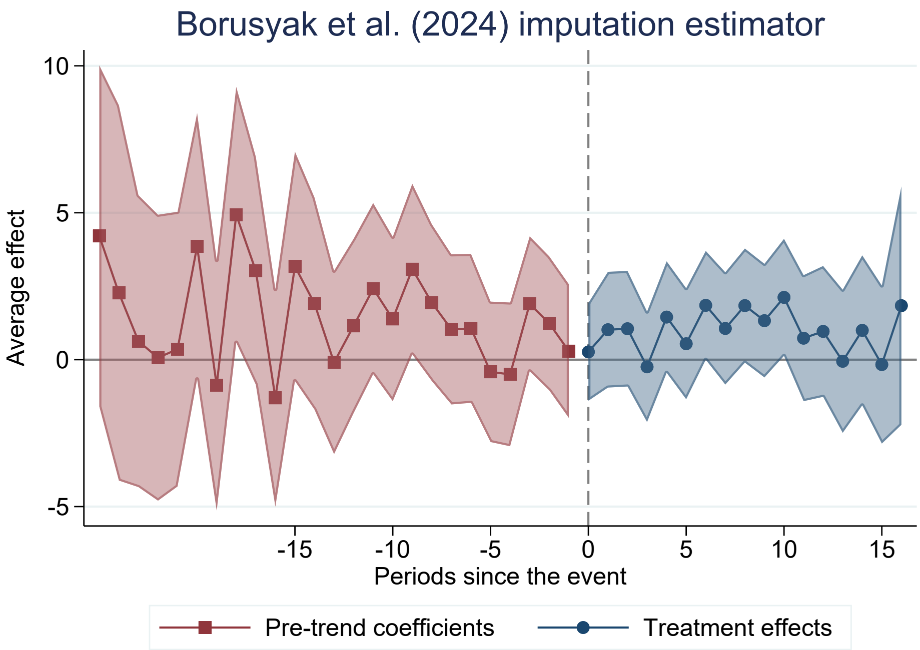
Notes : Significance: *** p<0.01, ** p<0.05, * p<0.10. OLS model specification. Standard errors are clustered at the institution-field level. Dependent variable: share of women among authors. Control variables include: mean of US co-authorship, mean of EU co-authorship, mean of co-authorship in the same world region, mean of co-authorship in the same world region of institution in R4L program, log of the average number of authors per scientific publication, dummies indicating other R4L subprograms adoption years.

Table 11: Effect of Hinari program on relative contribution of women to publication quality (RPQ)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Field × Treated × Post	0.0173**	0.0172**	0.0172**	0.0141*	0.0141*	0.0141*	0.0123	0.0125
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Controls		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes			Yes		Yes	
City FE			Yes			Yes		Yes
Country-specific time trend							Yes	
City-specific time trend								Yes
N	19123	19123	19123	19123	19123	19123	19123	19123
R^2	0.21	0.21	0.21	0.21	0.21	0.21	0.22	0.24

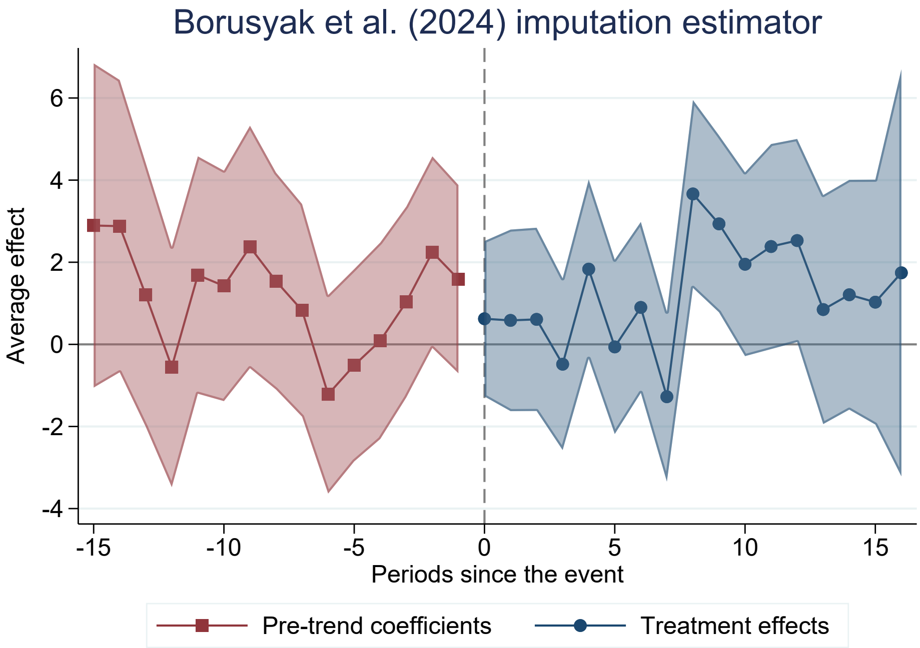
Notes : Significance: *** p<0.01, ** p<0.05, * p<0.10. OLS specification.Standard errors are clustered at the institution-field level. Dependent variable: relative JIF. Control variables include: share of US co-authorship, mean of EU co-authorship, mean of co-authorship in the mean world region, mean of co-authorship in the same world region of institution in R4L program, log of the average number of authors per scientific publication, dummies indicating other R4L subprograms adoption years.

Fig. 2: Event Study on share of papers with at least one woman



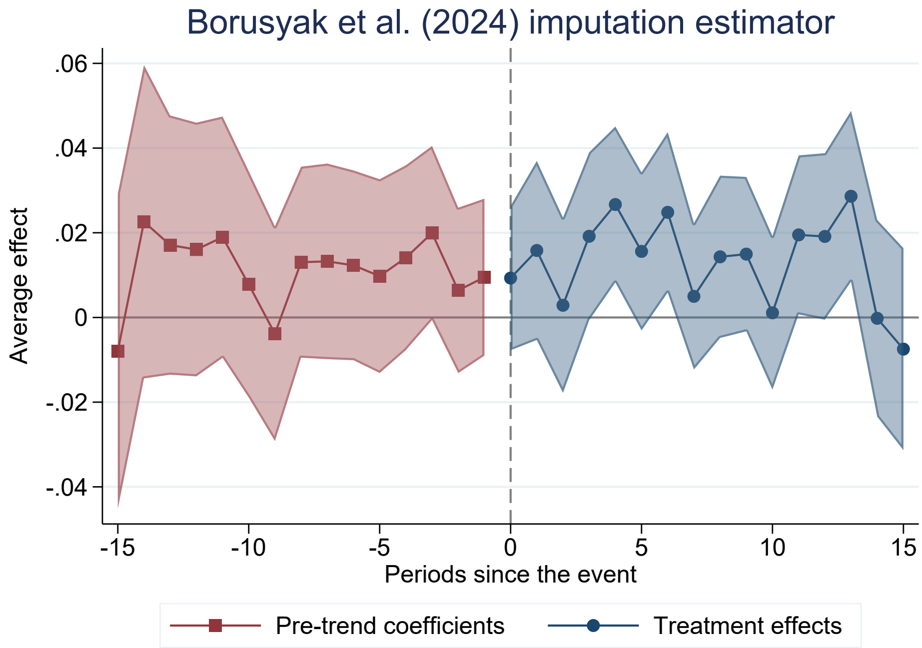
Notes : The graph reports the 90% confidence intervals adopting the Borusyak, Jaravel and Spiess (2024) estimator. Dependent variable: share of papers with at least one woman. Control variables include: mean of US co-authorship, mean of EU co-authorship, mean of co-authorship in the same world region, mean of co-authorship in the same world region of institution in R4L program, log of the average number of authors per scientific publication, dummies indicating other R4L subprograms adoption years. Standard errors clustered at the institutional-field level.

Fig. 3: Event Study on share of women among authors



Notes : The graph reports the 95% confidence intervals Pseudo-Poisson coefficient estimates. Dependent variable: SWA. The institution-discipline-quarter triplets constitute the unit of observation. Control variables include: mean of US co-authorship, mean of EU co-authorship, mean of co-authorship in the same world region, mean of co-authorship in the same world region of institution in R4L program, log of the average number of authors per scientific publication, dummies indicating other R4L subprograms adoption years. Standard errors clustered at the institutional-field level.

Fig. 4: Event Study on RPQ



Notes : The graph reports the 95% confidence intervals Pseudo-Poisson coefficient estimates. Dependent variable: RPQ. The institution-discipline-quarter triplets constitute the unit of observation. Control variables include: mean of US co-authorship, mean of EU co-authorship, mean of co-authorship in the same world region, mean of co-authorship in the same world region of institution in R4L program, log of the average number of authors per scientific publication, dummies indicating other R4L subprograms adoption years. Standard errors clustered at the institutional-field level.

Table 12: Effect of Hinari program on the share of publications with at least one woman author (SPW), with usage (login) data as continuous treatment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Triple DiD (logins)	2.4335***	0.9398***	0.9398***	0.3738***	0.3738***	0.3738***	0.4391***	0.4648***
	(0.15)	(0.14)	(0.14)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)
Controls		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes			Yes		Yes	
City FE			Yes			Yes		Yes
Country-specific time trend							Yes	
City-specific time trend								Yes
N	29389	29389	29389	29389	29389	29389	29389	29389
R^2	0.35	0.43	0.43	0.46	0.46	0.46	0.47	0.49

Notes : Significance: *** p<0.01, ** p<0.05, * p<0.10. OLS model specification. Standard errors are clustered at the institution-field level. Dependent variable: SPW. Control variables include: mean of US co-authorship, mean of EU co-authorship, mean of co-authorship in the same world region, mean of co-authorship in the same world region of institution in R4L program, log of the average number of authors per scientific publication, dummies indicating other R4L subprograms adoption years. Sample split by low and high educational attainment using the median.

Table 13: Effect of Hinari program on the share of women authors (SWA), with usage (login) data as continuous treatment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Triple DiD (logins)	1.1247***	0.5435***	0.5435***	0.1990	0.1990	0.1990	0.3314**	0.3417***
	(0.13)	(0.14)	(0.14)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)
Controls		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes			Yes		Yes	
City FE			Yes			Yes		Yes
Country-specific time trend							Yes	
City-specific time trend								Yes
N	29389	29389	29389	29389	29389	29389	29389	29389
R^2	0.32	0.33	0.33	0.34	0.34	0.34	0.35	0.37

Notes : Significance: *** p<0.001, ** p<0.01, * p<0.05. OLS model specification. Standard errors are clustered at the institution-field level. Dependent variable: SWA. Control variables include: mean of US co-authorship, mean of EU co-authorship, mean of co-authorship in the same world region, mean of co-authorship in the same world region of institution in R4L program, log of the average number of authors per scientific publication, dummies indicating other R4L subprograms adoption years. Sample split by low and high educational attainment using the mean.

Table 14: Heterogenous analysis - Effect of Hinari program on the share of publications with at least one woman author (SPW) by educational attainment

	Low Educational Attainment				High Educational Attainment			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Field × Treated × Post	3.7201***	3.7201***	3.9574***	3.9875***	0.5844	0.5844	1.1651	1.3168
	(1.32)	(1.33)	(1.32)	(1.31)	(1.59)	(1.60)	(1.55)	(1.58)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes		Yes		Yes		Yes	
City FE		Yes		Yes		Yes		Yes
Country-specific time trend			Yes				Yes	
City-specific time trend				Yes				Yes
N	16592	16592	16592	16592	11726	11726	11726	11726
R^2	0.47	0.47	0.47	0.49	0.46	0.46	0.47	0.48

Notes : Significance: *** p<0.01, ** p<0.05, * p<0.10. OLS model specification. Standard errors are clustered at the institution-field level. Dependent variable: SPW. Control variables include: mean of US co-authorship, mean of EU co-authorship, mean of co-authorship in the same world region, mean of co-authorship in the same world region of institution in R4L program, log of the average number of authors per scientific publication, dummies indicating other R4L subprograms adoption years. Sample split by low and high educational attainment using the median.

Table 15: Heterogenous analysis - Effect of Hinari program on the share of publications with at least one woman author (SPW) by educational attainment

	Low Educational Attainment				High Educational Attainment			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Field × Treated × Post	3.4887**	3.4887**	3.9958**	3.8584**	1.7046	1.7046	2.1361*	2.3137*
	(1.73)	(1.74)	(1.69)	(1.65)	(1.27)	(1.28)	(1.26)	(1.28)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes		Yes		Yes		Yes	
City FE		Yes		Yes		Yes		Yes
Country-specific time trend			Yes				Yes	
City-specific time trend				Yes				Yes
N	10914	10914	10914	10914	17404	17404	17404	17404
R^2	0.48	0.48	0.48	0.50	0.46	0.46	0.47	0.48

Notes : Significance: *** p<0.001, ** p<0.01, * p<0.05. OLS model specification. Standard errors are clustered at the institution-field level. Dependent variable: SPW. Control variables include: mean of US co-authorship, mean of EU co-authorship, mean of co-authorship in the same world region, mean of co-authorship in the same world region of institution in R4L program, log of the average number of authors per scientific publication, dummies indicating other R4L subprograms adoption years. Sample split by low and high educational attainment using the mean.