Working Paper Series

#2022-040

Gender bias in team formation: The case of the European Science Foundation's grants

Michele Pezzoni and Fabiana Visentin

Published 28 November 2022

Maastricht Economic and social Research institute on Innovation and Technology (UNU-MERIT)
email: info@merit.unu.edu | website: http://www.merit.unu.edu

Boschstraat 24, 6211 AX Maastricht, The Netherlands
Tel: (31) (43) 388 44 00

UNU-MERIT Working Papers
ISSN 1871-9872

Maastricht Economic and social Research Institute on Innovation and Technology
UNU-MERIT

UNU-MERIT Working Papers intend to disseminate preliminary results of research carried out at UNU-MERIT to stimulate discussion on the issues raised.

Gender bias in team formation:

The case of the European Science Foundation’s grants

Michele Pezzoni

Université Côte d’Azur, CNRS, GREDEG, France;
Observatoire des Sciences et Techniques, HCERES, Paris, France;
OFCE, SciencesPo, Paris, France;
ICRIOS, Bocconi University, Milan, Italy;
michele.pezzoni@univ-cotedazur.fr

Fabiana Visentin

UNU-MERIT, Maastricht University, the Netherlands
f.visentin@maastrichtuniversity.nl

Abstract

This paper investigates gender bias (if any) when teams are formed. We use data from the European Science Foundation to estimate if female scientists have the same opportunities as their male colleagues to join a team when applying for funds. To assess gender bias, we construct a control group of scientists with the competencies for being invited to join the team but do not join. By comparing the proportion of female scientists in the control group with the one in the observed teams, we find a gender bias against female scientists only when a project leader is a male scientist. At the same time, we do not observe gender bias when the project leader is a female scientist.

Keywords Gender bias, Team science, Team selection, Research grants

Acknowledgments: The authors acknowledge the financial support of GIGA (ANR-19-CE26-0014).

1. Introduction

Despite the call for gender parity coming from private actors, public actors, policymakers, and also advocated among the Sustainable Development Goals (SDGs) by the United Nations General Assembly, a remarkably significant gender gap against women persists across different sectors of our society. Academia is not exempted from gender inequality: female scientists suffer from lower productivity (Arensbergen et al., 2012; Kelchtermans and Veugelers, 2013) and fewer promotion opportunities than their male colleagues (Ceci and Williams, 2011; McDowell et al., 1999). Among the factors that help women to close their productivity gap, previous studies have shown the benefit of working in teams (Badar et al., 2013; Fox, 2020; Fox and Mohapatra, 2007; Kyvik and Teigen, 1996).

If scientists benefit from working in teams, it becomes crucial to understand if female and male scientists have equal opportunities to enter a team. Extant studies have considered the propensity of individual scientists to collaborate (Abramo et al., 2017, 2009; Boschini and Sjögren, 2007). However, no studies have considered whether female scientists have the same probability as their male colleagues of being selected as team members when a team is formed. Indeed, team members do not randomly join a team. Instead, scientists team up according to their characteristics, background, specialization, and skills. Therefore, this paper aims to answer the question: does gender affect team formation?

We rely on a unique dataset of 329 teams of scientists applying to the European Science Foundation (ESF) granting program between 2006 and 2010. The program was launched to promote an "*independent, bottom-up approach to collaborative research in Europe that allowed new scientific ideas to be developed*" (ESF - Evaluation of EUROCORES final report, 2015) and required one project leader (PL) to team up with other principal investigators (PIs) to respond to an open call. To assess the gender bias in team formation (if any), we needed to assess the presence of female scientists in a poll of potential PI candidates that the PL might have invited to join the team, but she did not invite. Being limited by observing only the final composition of the 329 teams analyzed, we implemented a quasi-natural experiment constructing a control group of potential teammates. Specifically, we compared the probability of having a female scientist in a control group of scientists with the potential to be selected with the probability of having a female scientist in the observed teams. Potential PIs are selected assuming three main mechanisms used by PLs to screen PIs. First, we assume that a PL invites PIs to join the team by searching among the scientists with whom she has already collaborated. Then, a PL looks for PIs among colleagues she knows for their research work. Finally, a PL looks for PIs in university departments well-known in the research areas of interest for the project application. Once we identified all the eligible controls in these three pools of scientists, we paired each observed PI with a corresponding potential PI with the most similar characteristics. Finally, we compared the probability of observing a female PI in the control group and the observed teams.

We find a gender bias in the team formation phase. The probability of having a female PI in an observed team equals only 17.5%, compared to the 22.6% probability of having a female scientist in the control group of potential PIs. By digging into this result, gender bias appears only when the PL is a male scientist. In this latter case, the probability of observing a female PI equals 15.8%, against the 22.9% probability of observing a female scientist in a control group of potential PIs. On the contrary, we do not observe any gender selection bias within teams led by female PLs: female PLs select 27.5% of female PIs, a value that is not statistically different from the 20.5% of female individuals observed in the control group of potential PIs.

Our work makes three main contributions. First, we add to the literature on team composition. The extant contributions focus on the role played by women in research teams (Bandiera et al., 2013; Dasgupta et al., 2015; Gaulé and Piacentini, 2018; Ivanova-Stenzel and Kübler, 2005; Pezzoni et al., 2016; Sosik and Godshalk, 2000; West et al., 2013). Although those studies have contributed to assessing how gender composition makes teams more productive, they have neglected that the observed teams' composition results from a selection process of individuals. We add to this literature by investigating the selection phase. Investigating selection is crucial because it might influence the team's performance. For instance, a gender bias against women might lead to not selecting women with competencies and skills that would have fitted the project's aims, causing a decrease in team performance. Our second contribution is that our empirical setting allows consider successful and unsuccessful teams. When reconstructing teams of scientists, extant studies look at co-authorships in published papers (Abramo et al., 2017, 2017, 2009; Boschini and Sjögren, 2007). By doing so, they observe teams that successfully deliver publication outcomes. In our work, we consider teams formed to compete in fundraising, regardless of the result of the competition. Our choice of the empirical setting mitigates a potential bias resulting from observing only successful teams of superior quality. Finally, we add to the gender parity debate. Recent statistics report that women are unrepresented in research teams (Abramo et al., 2009). However, a lower number of women in a team does not necessarily denote a gender bias against women because it might result from a scarcity of women among the potential teammates to select (Card et al., 2022). By comparing the proportion of potentially qualified women to enter a team with those who entered teams, our findings confirm gender bias against women in selection. In other words, female scientists suffer forced isolation that might have detrimental effects on their productivity and career progress.

The paper is structured as follows. Section 2 reviews the literature on teamwork in science. Section 3 describes our empirical strategy to assess the existence of a possible gender bias in team formation. Section 4 presents our data and describes how we construct the control sample of potential PL's teammates. Section 5 shows the results, and Section 6 concludes.

2. Teamwork in science

Teamwork is a growing phenomenon. Within the scientific community, Wuchty and colleagues (2007) analyzed paper co-authorships and patent co-inventorships between the late 50'ies and 2000 over a sample of about 20 million papers and 2 million patents, finding that teamwork has become the dominant working mode to produce knowledge in all fields. Not only individuals prefer teamwork to solo work, but when joining their efforts, their outcomes have a higher impact (Uzzi et al., 2013; Weitzman, 1998). For example, multiple authors' papers receive more citations than solo-author ones (Freeman and Huang, 2014; Wuchty et al., 2007). Teamwork also favors creativity (Jones, 2021). Recently, Wu et al. (2019) examined more than 65 million papers, patents, and software products realized over sixty years, from 1954 to 2014. Considering the teamwork outcomes over such a sample, the authors find that teams play a key role in the science and technology landscape, and, in particular, small teams produce disruptive outcomes with high impact.

Individuals team up to complement their knowledge and skills in response to the so-called 'burden of knowledge' challenges (Jones, 2009). Knowledge has accumulated with time and, for single individuals, it is impossible to keep pace, so they specialize in narrow areas of expertise. At the same time, the complexity of the problems often requires interdisciplinary contributors (Falk-Krzesinski et al., 2011; Milojević, 2014). As a consequence, individuals team up to complete complex tasks. By surveying almost 4,000 corresponding authors involved in collaborations with at least one US coauthor in the fields of Particle and Field Physics, Nanoscience and Nanotechnology, and Biology and Applied Microbiology, Freeman et al. (2014) find that searching for unique human capital is the main driver for collaborations: scientists opt for the division of labor and look for "unique knowledge, expertise, and capabilities" (page 14). Finally, another reason why individuals team up is to access special equipment or infrastructures that would not be accessible to single individuals due to their high cost (Katz and Martin, 1997; Stephan, 1996).

If teamwork allows individuals to access complementary skills and unique resources to boost productivity, creativity, and impact, it becomes crucial to understand how individuals access a team. In recent years collaboration costs have decreased, and teamwork faces lower entry barriers. The broad-scale adoption of internet has facilitated long-distance communication (Cairncross, 1997), and the substantial reduction of traveling costs made possible through solutions like access to low-cost flight connections boosted collaborations. For instance, Catalini et al. (2020) analyzed the effect of the entry of the low-cost Southwest in the airlines market, finding that the number of collaborations increased between 0.3 and 1.1 times and stimulated more novel projects of higher quality. However, those opportunities do not equally benefit individuals. Specifically, female scientists do not respond to lower travel cost incentives because women 'may have more constrained travel schedules' than men (Catalini et al., 2020).

Jones (2021) discusses more broadly the problem of having access to teams and claims that biases might limit the selection of specific categories of persons in teams, mainly when the team members' screening happens among people who do not know each other personally. One potentially important bias is the one against women. In the growing debate about gender equality in science (Holman et al., 2018; Larivière et al., 2013), understanding gender bias in team formation becomes crucial because the existence of such a gender bias would be detrimental to women's careers and the entire society when potential talents are limited or excluded. Extant studies have analyzed the role played by gender in team formation, looking at the coauthorship in published papers. Boschini and Sjögren (2007) analyze coauthorships in papers published in three top economics journals from 1991 to 2002 and find that team formation is not gender-neutral. "Women are more than twice as likely as men to have a female coauthor (p. 338)." Women are also more likely to publish solo-author papers than their male colleagues, and "as the fraction of female researchers [in the field] increases, women increasingly tend to write with other women" (p. 339). This latter finding shows that even if there is no lack of women in the field, men tend not to team up with women. As Boschini and Sjögren (2007), we choose academia as an empirical setting since, in this context, team formation is voluntary. Scientists choose the colleagues they want to work with based on individual preferences and expectations. Unlike Boschini and Sjögren (2007), we consider scientists who team up to apply for a grant, regardless of the result of the grant competition. By doing that, we overcome the limitation of analyzing only successful teams where the bias observed might result from a selection process discriminating against certain types of teams. We cover all the disciplines and implement a quasi-natural experiment where we compare the observed proportion of women in the observed teams with the proportion of women in a control group of potential candidates. As pointed out by Card et al. (2022), fewer women observed in teams might result from a scarcity of women among qualified candidates or from a bias in selecting between equally qualified women and men. The quasi-natural experiment design described in Section 3 aims to disentangle those two cases by identifying a bias in the selection of female teammates.

3. Empirical strategy

Using data on the observed teams, we calculate the probability of observing a female PI (*F*) conditional on having joined an observed team (*join*), *P(F|join)*. However, the pool of potential candidates from which the observed ones are selected can have a different gender composition depending on the discipline and abilities required. Having no female scientists in the team might not necessarily reflect a gender selection bias, but a lack of female scientists in the discipline or a lack of women having the competencies needed for the research project. For instance, a low share of women in historically male-dominated disciplines such as physics and mathematics (Ginther and Kahn, 2004; Stephan, 1996) might not denote a gender bias but simply a lack of women in the eligible pool of candidates. In other fields, like psychology, where women are highly represented, a low share of women in teams might be associated with a selection bias (Card et al., 2022).

To account for the difference in the share of potentially qualified women and to correctly estimate the existence of a gender bias (if any) in teammate selection, we compare the probability of having a female researcher in the observed teams, *P(F|join)*, with the probability of observing a female researcher in a matched control group of individuals similar to those selected according to their observable characteristics. Specifically, in the control group, we calculate the unconditional probability of observing a female researcher *P(F)* in the pool of potential qualified PIs. By calculating the ratio *P(F|join)/P(F),* we estimate if female researchers are more (or less) frequent among individuals in the observed teams than among individuals in the control group[1]. In other words, if the ratio *P(F|join)/P(F)* is lower than 1, a gender bias against selecting teammate female scientists exists, while if the ratio is higher than 1, a gender bias in favor of selecting female scientists exists.

Since assessing the existence of a gender bias in team formation crucially depends on the reliability of the control group, defining the pool of potential qualified PIs is the main challenge of our empirical analysis. Section 4 describes our data and details how we identify the control group of potential PIs.

4. Data

The European Science Foundation (ESF) is a non-governmental, internationally oriented, non-profit foundation that promotes high-quality research projects in Europe. In the ESF's funding opportunities portfolio, the European Collaborative Research initiative (EUROCORES) was launched to promote an "independent, bottom-up approach to collaborative research in Europe that allowed new scientific ideas to be developed." (Bianchini et al., 2022; *ESF - Evaluation of EUROCORES final report.pdf*, 2015). One of the EUROCORES objectives was to sponsor collaborations that would not have taken place without such funds. A team led by a project leader (PL) and a minimum of two principal investigators (PIs) originating from two different countries was eligible to apply with a project to the EUROCORES' open call on a broad range of disciplines, including Life, Earth, and Environmental Sciences, Biomedical Science, Physical and Engineering Science, Humanities and Social Science.

4.1 Observed teams

In our analysis, we consider 329 teams applying with a project to the ESF between 2006 and 2010. Figure 1 reports the team size distribution. The average team counts 3.83 PIs. Most teams count two or three PIs, 26.7% and 28.0% of the total number of teams, respectively. Only 17.6% of the teams have more than five PIs. The average share of female scientists does not vary substantially by team size (see Figure 2) or application year (see Figure 3). On average, each team has 17.5% of female PIs.

[1] An easily interpretable equivalent of the calculated ratio results from applying the Bayes Rule. Indeed, we can interpret *r* as the ratio between the probability of joining a team conditional on the PI gender and the unconditional probability of joining a team, i.e., *P(join|F)/P(join)*.

Figure 1. Team size distribution

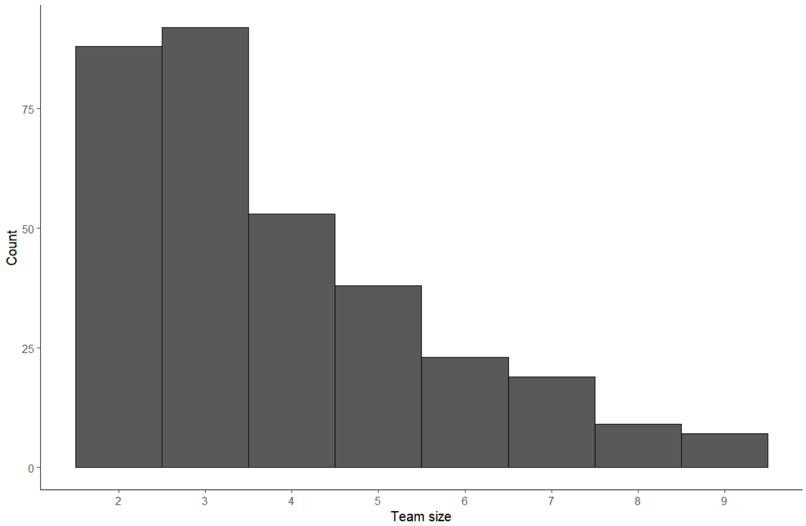


Figure 2. Average share of female scientists by team size

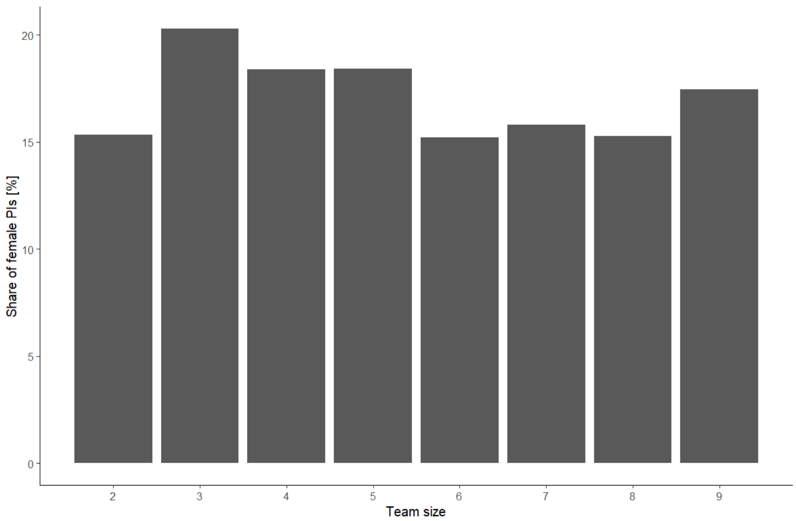
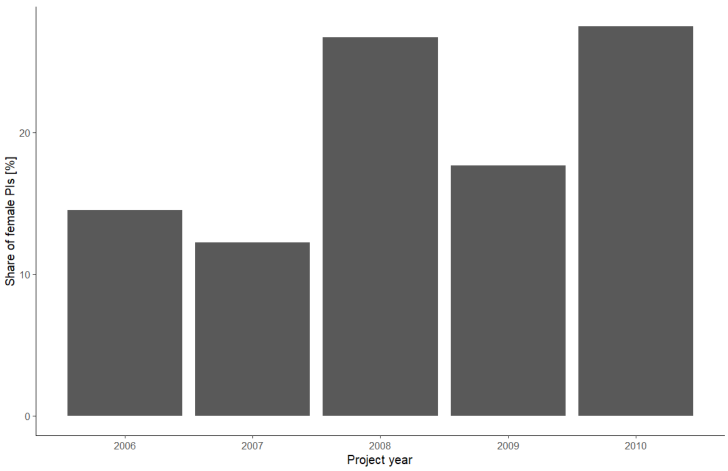


Figure 3. Average share of female scientists by application year



For each application, we obtained from the ESF information on the name of the applicants, gender, and affiliations. We complemented that information with publication records retrieved from Scopus Elsevier's database. Table 1 reports the main descriptive statistics for our sample of 1,260 PIs. On average, PIs are senior scholars with a solid academic curriculum. The average PI has 15 years of experience from the first publication, reports a record of 40 publications, and leverages an extensive network of co-authors (87.19). When considering the gender distribution of the PI-PL gender pairs, female PIs tend to be led more often by female PLs (21%) than male PIs (12%).

Table 1. Profiles of the PIs observed in our sample, by gender

	Observed PIs (1,260)			
	Female PIs (221)		Male PIs (1,039)	
	Average	SD	Average	SD
Number of publications	40.48	41.19	74.98	96.37
Career length	15.70	8.33	19.23	9.67
Number of coauthors	87.19	95.66	121.86	152.41
Female PL	0.21	0.41	0.12	0.32

4.2 Defining the control sample

To define our control sample of qualified PIs, we follow a 3-step procedure. First, we begin by constructing a large pool of researchers with the potential to be selected by the PLs (step 1). Then, among those potential PIs, we extract those who have competencies coherent with the project application team (step 2). Finally, we match observed PIs with potential PIs identified in steps 1 and 2, relying on observable characteristics such as publication profile, career length, and scientific network (step 3).

4.2.1 Identifying the pool of potential controls

As Step 1, we construct the group of potential PIs assuming that PLs choose their teammate PIs within three different pools of researchers: (i) those researchers with whom they have collaborated in the past, (ii) reputed colleagues of whom they know the work, and (iii) researchers working in departments well-known for conducting research on subjects functional to the project they intend to propose to ESF.

To reconstruct those three pools, we collect (i) the list of coauthors with whom the PLs have worked before applying for an ESF call (*PL's co-authors*), (ii) the list of authors cited by the PLs in the works published before applying for an ESF call (*PL's cited authors*), and (iii) the list of departmental colleagues of the PIs selected who are active before the application date of the ESF call (*Colleagues of the observed PIs*).

PL's co-authors - Looking at the coauthors' list, the 329 PLs of our sample collaborate, on average, with 62.86 coauthors each.

PL's cited authors - In identifying the authors cited in PLs' articles, we limit our search to papers with less than 20 references listed to exclude literature reviews. We also excluded the authors of papers older than ten years at the project application date to ensure that the authors considered are likely to be active researchers. On average, the PLs in our sample have 81.12 distinct authors in their reference lists.

Colleagues of the observed PIs - Since PLs are supposed to target researchers in well-known departments to conduct research in the field of interest of their projects, we assume that the departments of the selected PIs are potential sources of talents. Then, we identify as potentially talented PIs those scientists who are departmental colleagues of the selected PIs, i.e., those working at the same university and discipline as the selected PIs in the year of the project application. Using the Scopus database, we create the list of department colleagues, extracting the names of authors with the same affiliation as the observed PIs and the same discipline defined using the observed PIs' Scopus subject area in which they published the most. On average, we identify for each PL in our sample 80.62 distinct authors who are department colleagues of the selected PIs.

After polling together these three lists of scientists, i.e., coauthors, cited authors, and departmental colleagues, we identified 69,710 potential PIs for the 329 teams, with on average 211.89 potential PIs for each team. Among the PIs part of the control group, 32% are previous coauthors of the PL, 36% are scientists listed in the references of the PL publications, and 37% are departmental colleagues of the selected PIs. As the last step, we assigned to each potential PI the gender using the information gathered from the Gender Name Dictionary of the World Intellectual Property Organization (Martínez et al., 2016). In the pool of the 69,710 potential PIs, 27% are women.

4.2.2 Matching the controls

In the pool of potential PIs, each PL selects her teammates by looking closely at their scientific profiles. We assume that the observed PIs have the ideal characteristics the PL is looking for. To measure gender bias in the team selection, we extract from the large pool of potential PIs those scientists whose profiles match the ones of the selected PIs along a set of relevant characteristics excluding gender. To do that, we implement the remaining two steps of our procedure. In step 2, we identify the 20 potential PIs with the most similar scientific competencies to the selected PIs. By assuming that scientific competencies disclose in publications, we select all the articles published by the observed PIs before the project application date and extract the list of words[2] appearing in titles and abstracts. Similarly, we isolate the words appearing in the title and abstract of the potential PIs' articles. Once each observed PI is associated with a bag of words, by project, we extracted the 20 potential PIs sharing the highest number of words with the PIs selected. In doing so, we restrict our control group from 211.89 potential PIs per project to 20 potential PIs with similar competencies to the selected ones. Looking at the origin of the researchers in our control group, 39% of the researchers have coauthored with the PL, 32% have been cited by the PL, and 36% are department colleagues of the observed PIs[3].

In step 3, among the 20 potential PIs selected by project, we conducted a pairing exercise with the observed PIs. Specifically, for each observed PI, we identify the nearest potential PI in terms of productivity, career length, and the number of coauthors until the year of the project application. To do so, we construct three categorical variables. The first variable *High productivity* is a dummy variable that equals 1 if the PI has a number of publications higher than the median value in our sample, i.e., 19 publications at the time of the project application. The second dummy variable, *Long career*, considers the career's length. We define *Long career* as a dummy variable that equals 1 if the time elapsed from the date of the first PI's publication to the year of the application is greater than 13 years, where 13 is the median career length in our sample. Finally, we define the dummy variable *Large network* as a variable that equals 1 if the PI has a large coauthorship network, i.e., more than the median number of 39 coauthors in our sample.

Table 2 reports the matching equation estimated with a logistic regression where the explained variable is the dummy variable *Observed PI*, which equals one if the PI belongs to the list of the 1,260 observed PIs and zero if the PI belongs to the list of 6,580 potential PIs[4]. The estimates show that highly productive PIs with long careers and large networks are more likely to become part of the observed teams.

[2] We excluded stop words.

[3] The percentages do not sum up to 100% since one potential PI might be selected according to more than one criterion.

[4] Corresponding to 20 potential PIs for each of the 329 project applications, i.e., 20*329=6,580.

Table 2. Propensity Score Matching regression. Logit estimates and marginal effects reported

	(1) Observed PI
High productivity (D)	0.15***
	(0.014)
Long career (D)	0.028**
	(0.011)
Large network (D)	0.034***
	(0.012)
Missing career information	0.13***
	(0.041)
Observations	7,840 (1,260+6,580)
Pseudo R2	0.0728

Note: We do not have the information concerning the PIs' careers for 1% of our sample. In these cases, we set the dummy variable *Long career (D)* equal to zero, and we calculate a dummy variable *Missing career information* that equals one when the career information is missing, zero otherwise.

As a result of step 3, we match 1,260 observed PIs with 1,260 potential PIs. Table 3 compares the average values of the three matching variables for the observed and potential PIs, showing the quality of the matching exercise. Specifically, 78% of the observed PIs has high productivity, compared to 76% of the potential PIs. Seventy-two percent of the observed PIs has a long career, compared to the 70% of potential PIs. Finally, 72% of the observed PIs has a large network, compared to the 71% of potential PIs. All the differences between the two groups of PIs, i.e., observed and potential, are not statistically significant.

Table 3. Comparison between observed and potential PIs' characteristics

	Observed PIs	Potential PIs	H0: $\mu_{\text{Observed}} = \mu_{\text{Potential}}$ P-value
High productivity (D)	0.78	0.76	0.14
Long career (D)	0.72	0.70	0.29
Large network (D)	0.72	0.71	0.62

Table 4 summarizes the average number of potential PIs for each step of the 3-step procedure implemented to construct our control sample and the share of potential PIs drawn from each pool.

Table 4. Average number of controls per project application along the 3-step procedure implemented to identify the control sample of potential PIs and the share of researchers drawn from each pool

	Pool 1 PL's co-authors	Pool 2 PL's cited authors	Pool 3 Colleagues of the observed PIs	Average number of Potential PIs per project
Step 1: Controls in Pools 1, 2 and 3	32%	36%	37%	211.89
Step 2: Controls with the competencies needed for the project	39%	32%	36%	20
Step 3: Controls after Propensity Score Matching	39%	34%	35%	3.83 (one for each observed PI)

Note: The average number of PI controls per project reported for Step 3 (3.83) equals the average number of observed PIs. These values are expected because we retrieved one potential PI for each observed PI.

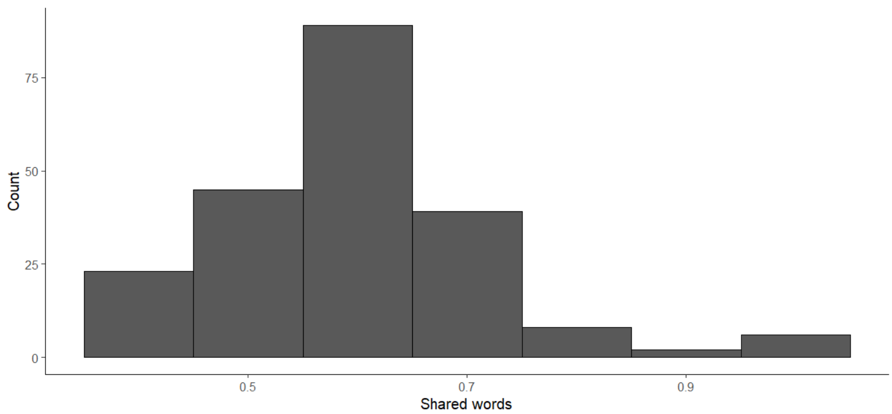
4.2.3 An illustrative example

To illustrate our 3-step procedure, we present in this Section an anonymized example. Project application X was crafted in 2009 and listed one male PL, PL1, affiliated with the university of Turku in Finland. The same application included two male PIs, PI1, and PI2, affiliated with the university of Jyväskylä in Finland and the Konrad Lorenz Institute in Austria, respectively. The project application was in the field of Ethology, the study of animal behavior. PL1 had 88 publications before 2009, listing 81 distinct coauthors and 184 referenced authors in their bibliographies[5]. Moreover, PI1 and PI2 have 41 colleagues publishing in the same field at Konrad Lorenz Institute (13 colleagues) and the University of Jyväskylä in 2009 (28 colleagues). Overall, we identified 283 potential PIs for this project application X. The number of potential PIs is not the algebraic sum of the 81 coauthors, 184 referenced authors, and 41 PIs' departmental colleagues because roles might overlap. Indeed, a scientist might be a coauthor, referenced author, or PIs' departmental colleague simultaneously. Specifically, 28.6% of the 283 potential PIs are PL's coauthors, 65.0% are referenced authors, and 14.5% are PIs' departmental colleagues.

[5] These figures are the results of applying the constraints defined in section 4.2.1. For instance, in searching for referenced authors, we consider only papers with less than 20 references listed to exclude literature reviews. We also excluded the authors of papers older than 10 years at the project application date to ensure that the authors considered are likely to be active researchers.

Once we identified the 283 potential PIs, we retrieved the articles' titles and abstracts in which they were listed as authors and created the 283 corresponding vectors of words gathered from those titles and abstracts. Then, we calculated a similar vector, including all the words in the articles' titles and abstracts published by PI1 and PI2 before the application date. This latter vector represents the competencies shown by the PIs in contributing to the project. Figure 3 shows the distribution of the share of words in common between the vector representing the competencies required for the project and each of the 283 vectors of words representing the content of the articles where the potential PIs were listed. Finally, we selected the potential PIs with the 20 highest shares, retrieving 12 female and 8 male researchers. Among those researchers, we selected the most similar to PI1 according to three observable characteristics: productivity, career, and network size. We did the same for PI2. For PI1, we found a female researcher with similar characteristics (control C1), while for PI2, we found a male researcher (control C2). In this example, a male researcher PI1 has been invited by PL1 to join the application team, but potentially a female researcher with the same characteristics (C1) could have been chosen.

Figure 3. Distribution of the share of words in common between the vector of the competencies needed for project X and the vectors representing the articles where we found the potential PIs



5. Results

To identify gender bias in team selection, we compare the probability of having a female scientist in the teams of observed PIs, *P(F|join)*, with the probability of having a female scientist in the control group of potential PIs identified with the 3-step procedure illustrated above, *P(F)*. A ratio *P(F|join)/P(F)* equal to 1 means no gender bias in team selection since the probability of having a female scientist in the observed team equals the probability of having a female scientist in the control group.

Results reported in Table 5 show that there is a gender bias against female PIs when PLs form their teams. Indeed, the ratio *P(F|join)/P(F)* equals 0.77 being the proportion of female PIs in the observed teams 17.5%, *P(F|join)*, and the one in the sample of potential PIs 22.6%, *P(F)*. When we test the hypothesis that *P(F|join) = P(F)*, we obtain a P-value of 0.001, below the standard significance level of 0.01. This result allows us to reject the null hypothesis that the probability of having a female PI in the observed teams equals the probability of having a female PI in the control sample of potential PIs. In other words, the choice of PIs is not gender-neutral.

Interestingly, when we split the sample by the gender of the PL, we find that the gender bias against selecting female PIs is driven by the PL's gender (Table 6). For the 1,089 PIs led by a male PL, the ratio *P(F|join)/P(F)* equals 0.69 depicting a gender bias against female PIs. When we test the hypothesis that *P(F|join) = P(F),* we obtain a P-value close to zero, rejecting the hypothesis for the standard significance level of 0.01. This latter result allows us to conclude that there is statistical evidence of a gender bias against female PIs when a male PL leads the team. On the contrary, for the 171 PIs in the teams led by a female PL, the ratio *P(F|join)/P(F)* equals 1.34 depicting a higher probability of having a female PI in the observed teams than in the control sample of potential PIs. When we test the hypothesis that *P(F|join) = P(F)*, we obtain a P-value of 0.129, which does not allow us to reject the null hypothesis for the standard significance level of 0.1. Therefore, based on our empirical evidence, we cannot conclude that teams led by female PLs are gender biased, favoring the selection of female PIs. Nonetheless, the high P-value might be due to the not negligible reduction in the observations used in the statistical test when considering only teams led by female PLs. Indeed, the test is conducted by comparing a sample of 171 observed PIs and 171 potential PIs, a sample size of almost one-tenth of the original sample of 1,260 observed PIs and 1,260 potential PIs.

Table 5. Ratio P(F|join)/P(F) in the entire study sample

	Probabilities, P-values, and ratios
P(F|join)	17.5% (1,260 Observed PIs)
P(F)	22.6% (1,260 Potential PIs)
H0: P(F|join) = P(F)	0.001 (Pvalue)
P(F|join)/P(F)	**0.77**

Table 6. Ratio P(F|join)/P(F) in the subsamples of teams led by male and female project leaders

	Probabilities, P-values, and ratios
Conditional on having a female PL	
P(F|join)	27.5% (171 Observed PIs)
P(F)	20.5% (171 Potential PIs)
H0: P(F|join) = P(F)	0.129 (Pvalue)
P(F|join)/P(F)	**1.34**
Conditional on having a male PL	
P(F|join)	15.8% (1,089 Observed PIs)
P(F)	22.9% (1,089 Potential PIs)
H0: P(F|join) = P(F)	0.000 (Pvalue)
P(F|join)/P(F)	**0.69**

When looking at the results presented in Table 5, a possible concern is that we are grouping projects of different quality when considering awarded and non-awarded project applications. It might be that awarded projects have a diverse team composition than non-awarded projects. To address this concern, we split the applications by awarded status. Table 7 shows a similar gender bias against female PIs in awarded and non-awarded projects. The ratio *P(F|join)/P(F)* equals 0.78 for the awarded projects and 0.77 for non-awarded projects. When we test the hypothesis that *P(F|join) = P(F),* we obtain a P-value below the standard significance level of 0.05, allowing us to reject the null hypothesis of equal opportunity in accessing the team for female and male PIs.

Table 7. Ratio P(F|join)/P(F) in the sub-samples of awarded and non-awarded projects

	Probabilities, P-values, and ratios
Conditional on the application being awarded	
P(F|join)	19.11% (492 Observed PIs)
P(F)	24.6% (492 Potential PIs)
H0: P(F|join) = P(F)	0.04 (Pvalue)
P(F|join)/P(F)	**0.78**
Conditional on the application not being awarded	
P(F|join)	16.5% (768 Observed PIs)
P(F)	21.4% (768 Potential PIs)
H0: P(F|join) = P(F)	0.02 (Pvalue)
P(F|join)/P(F)	**0.77**

6. Conclusions

Teamwork has become the dominant model for producing science (Wuchty et al., 2007). Scientists team up to produce new knowledge by joining their competencies and efforts. Moreover, teamwork stimulates learning and creativity, benefiting individual teammates (Ayoubi et al., 2017). Funding agencies have also recognized the advantage of teamwork and included dedicated funding schemes reserved for team applications in their funding portfolios. Several studies have analyzed the impact of team composition on performance once teams are formed but have neglected the team formation phase (Dasgupta et al., 2015).

Given the importance of teamwork, understanding if individuals have equal opportunities to join a team is crucial to guarantee equality in career progress (Jones, 2021). Women are one of the categories that might encounter barriers in accessing teams. Unique to our study, we look at the team formation phase assessing if this phase is affected by gender bias. By doing so, we contribute to the broad debate on gender equality. We analyze gender bias in team formation, considering academia, an ideal context in which individuals team up voluntarily, selecting teammates based on their preferences and expectations. Differently from previous studies, we observe teams in which the composition results from a deliberate choice of individuals without any filtering or external assessment. Indeed, we consider a sample of observed teams of scientists who, between 2006 and 2010, applied for the European Collaborative Research initiative promoted by the European Science Foundation. Our study includes teams awarded and non-awarded with the grant. During the team formation phase, a Project Leader (PL) is asked to form her own team by asking Principal Investigators (PIs) to join. To assess gender bias, we looked at the measurable characteristics of the observed PIs – with the exclusion of gender – and identified for each of those PI a potential PI with similar characteristics that might have joined the team but did not.

To obtain a reliable control sample of potential PIs, we implemented a 3-step procedure assuming that a PL identifies her teammates among past colleagues with whom she has worked, colleagues of whom she knows the work, and scientists working in departments well-known for their field or research. In the first step, we retrieved a pool of potential PIs, including the coauthors of the PL, scientists cited in the PL's work, and scientists affiliated with the departments from which the PL selected the PIs in our observed teams. In the second step, we restricted the pool to those scientists with the competencies to complete the PL's project. Finally, we extracted the potential PIs with the closest profile to the observed PIs. Specifically, we found for each observed PI a 'twin' similar in terms of publications, career length, and the dimension of the coauthorship networks. By comparing the proportion of female scientists in the observed PIs' group with the same proportion in the group of potential 'twin' PIs, we found a gender bias against females. Moreover, we found that this result is driven by those teams led by a male PL.

Our findings add to the current call for gender parity coming from multiple actors. For example, the most recent G7 dashboard on gender claims, "Gender equality is a fundamental prerequisite for resilient, inclusive democratic societies. Nevertheless, gender inequalities continue to persist in all areas of social and economic life. Closing these gaps is an important goal and priority for all G7 members." ("Gender - OECD," 2022). Our results confirm this gender inequalities persistence and invite us to reflect on mechanisms and consequences.

Further studies should explore the mechanisms and answer questions like "Why shall a PL prefer male colleagues to females?". A possible explanation might be the belief that females are less likely to be awarded grants (Boyle et al., 2015). In this case, the PLs' bias against the inclusion of female PIs might be due to the females' stigma of having less probability of being funded. However, since the capacity to attract external funds is crucial in running labs and promotions (Stephan, 1996), it is crucial to take action to support female scientists. In this regard, a possible policy intervention to help women could be the introduction of a mandatory quota of females in teams as a requirement to be eligible when applying for grants. Along this line, the European Union is already introducing rules to tackle gender equality, and, as an example, they introduced gender balance among the ranking criterion to evaluate projects that obtained the same evaluation score ("Tackling gender equality in Research and Innovation," 2022).

Bibliography

Abramo, G., D'Angelo, A.C., Murgia, G., 2017. The relationship among research productivity, research collaboration, and their determinants. Journal of Informetrics 11, 1016–1030. https://doi.org/10.1016/j.joi.2017.09.007

Abramo, G., D'Angelo, C.A., Di Costa, F., 2009. Research collaboration and productivity: is there correlation? High Educ 57, 155–171. https://doi.org/10.1007/s10734-008-9139-z

Arensbergen, P., van der Weijden, I., Besselaar, P., 2012. Gender differences in scientific productivity: a persisting phenomenon? Scientometrics 93, 857–868. https://doi.org/10.1007/s11192-012-0712-y

Ayoubi, C., Pezzoni, M., Visentin, F., 2017. At the origins of learning: Absorbing knowledge flows from within the team. Journal of Economic Behavior & Organization 134.

Badar, K., Hite, J.M., Badir, Y.F., 2013. Examining the relationship of co-authorship network centrality and gender on academic research performance: the case of chemistry researchers in Pakistan. Scientometrics 94, 755–775. https://doi.org/10.1007/s11192-012-0764-z

Bandiera, O., Barankay, I., Rasul, I., 2013. Team Incentives: Evidence from a Firm Level Experiment. Journal of the European Economic Association 11, 1079–1114. https://doi.org/10.1111/jeea.12028

Bianchini, S., Llerena, P., Öcalan-Özel, S., Özel, E., 2022. Gender diversity of research consortia contributes to funding decisions in a multi-stage grant peer-review process. Humanit Soc Sci Commun 9, 1–10. https://doi.org/10.1057/s41599-022-01204-6

Boschini, A., Sjögren, A., 2007. Is Team Formation Gender Neutral? Evidence from Coauthorship Patterns. Journal of Labor Economics 25, 325–365. https://doi.org/10.1086/510764

Boyle, P.J., Smith, L.K., Cooper, N.J., Williams, K.S., O'Connor, H., 2015. Gender balance: Women are funded more fairly in social science. Nature 525, 181–183. https://doi.org/10.1038/525181a

Cairncross, F., 1997. The death of distance: The trendspotter's guide to new communications. Harvard Business School Press, Boston.

Card, D., DellaVigna, S., Funk, P., Iriberri, N., 2022. Gender Gaps at the Academies 21.

Catalini, C., Fons-Rosen, C., Gaulé, P., 2020. How Do Travel Costs Shape Collaboration? Management Science 66, 3340–3360. https://doi.org/10.1287/mnsc.2019.3381

Ceci, S.J., Williams, W.M., 2011. Understanding current causes of women's underrepresentation in science. Proceedings of the National Academy of Sciences 108, 3157–3162. https://doi.org/10.1073/pnas.1014871108

Dasgupta, N., Scircle, M.M., Hunsinger, M., 2015. Female peers in small work groups enhance women's motivation, verbal participation, and career aspirations in engineering. Proceedings of the National Academy of Sciences 201422822. https://doi.org/10.1073/pnas.1422822112

ESF - Evaluation of EUROCORES final report.pdf, 2015.

Falk-Krzesinski, H.J., Contractor, N., Fiore, S.M., Hall, K.L., Kane, C., Keyton, J., Klein, J.T., Spring, B., Stokols, D., Trochim, W., 2011. Mapping a research agenda for the science of team science. Res. Eval. 20, 143–156. https://doi.org/10.3152/095820211X12941371876580

Fox, M.F., 2020. Gender, science, and academic rank: Key issues and approaches. Quantitative Science Studies 1001–1006. https://doi.org/10.1162/qss_a_00057

Fox, M.F., Mohapatra, S., 2007. Social-organizational characteristics of work and publication productivity among academic scientists in doctoral-granting departments. The Journal of Higher Education 78, 542–571.

Freeman, R.B., Ganguli, I., Murciano-Goroff, R., 2014. Why and Wherefore of Increased Scientific Collaboration 42.
Freeman, R.B., Huang, W., 2014. Collaboration: Strength in diversity. Nature 513, 305–305. https://doi.org/10.1038/513305a
Gaulé, P., Piacentini, M., 2018. An advisor like me? Advisor gender and post-graduate careers in science. Research Policy 47, 805–813. https://doi.org/10.1016/j.respol.2018.02.011
Gender - OECD [WWW Document], 2022. URL https://www.oecd.org/gender/ (accessed 8.10.22).
Ginther, D.K., Kahn, S., 2004. Women in Economics: Moving Up or Falling Off the Academic Career Ladder? Journal of Economic Perspectives 18, 193–214. https://doi.org/10.1257/0895330042162386
Holman, L., Stuart-Fox, D., Hauser, C.E., 2018. The gender gap in science: How long until women are equally represented? PLOS Biology 16, e2004956. https://doi.org/10.1371/journal.pbio.2004956
Ivanova-Stenzel, R., Kübler, D., 2005. Courtesy and Idleness: Gender Differences in Team Work and Team Competition (SSRN Scholarly Paper No. ID 825686). Social Science Research Network, Rochester, NY.
Jones, B.F., 2021. The Rise of Research Teams: Benefits and Costs in Economics. Journal of Economic Perspectives 35, 191–216. https://doi.org/10.1257/jep.35.2.191
Jones, B.F., 2009. The Burden of Knowledge and the "Death of the Renaissance Man": Is Innovation Getting Harder? REVIEW OF ECONOMIC STUDIES 35.
Katz, J.S., Martin, B.R., 1997. What is research collaboration? Research policy 26, 1–18.
Kelchtermans, S., Veugelers, R., 2013. Top research productivity and its persistence: Gender as a double-edged sword. Review of Economics and Statistics 95, 273–285.
Kyvik, S., Teigen, M., 1996. Child care, research collaboration, and gender differences in scientific productivity. Science, Technology & Human Values 21, 54–71.
Larivière, V., Ni, C., Gingras, Y., Cronin, B., Sugimoto, C.R., 2013. Bibliometrics: Global gender disparities in science. Nature News 504, 211.
Martínez, G.L., Raffo, J., Saito, K., 2016. Identifying the gender of PCT inventors [WWW Document]. https://doi.org/10.34667/tind.28963
McDowell, J.M., Singell, L.D., Ziliak, J.P., 1999. Cracks in the Glass Ceiling: Gender and Promotion in the Economics Profession. American Economic Review 89, 392–396. https://doi.org/10.1257/aer.89.2.392
Milojević, S., 2014. Principles of scientific research team formation and evolution. Proc. Natl. Acad. Sci. U.S.A. 111, 3984–3989. https://doi.org/10.1073/pnas.1309723111
Pezzoni, M., Mairesse, J., Stephan, P., Lane, J., 2016. Gender and the Publication Output of Graduate Students: A Case Study. PLoS ONE 11, e0145146. https://doi.org/10.1371/journal.pone.0145146
Sosik, J.J., Godshalk, V.M., 2000. The Role of Gender in Mentoring: Implications for Diversified and Homogenous Mentoring Relationships. Journal of Vocational Behavior 57, 102–122. https://doi.org/10.1006/jvbe.1999.1734
Stephan, P.E., 1996. The economics of science. Journal of Economic literature 34, 1199–1235.
Tackling gender equality in Research and Innovation [WWW Document], 2022. URL https://rea.ec.europa.eu/news/tackling-gender-equality-research-and-innovation-2022-03-07_en (accessed 8.10.22).
Uzzi, B., Mukherjee, S., Stringer, M., Jones, B., 2013. Atypical Combinations and Scientific Impact. Science 342, 468–472.

Weitzman, M.L., 1998. Recombinant growth. The Quarterly Journal of Economics 113, 331–360.
West, J.D., Jacquet, J., King, M.M., Correll, S.J., Bergstrom, C.T., 2013. The role of gender in scholarly authorship. PloS one 8, e66212.
Wu, L., Wang, D., Evans, J.A., 2019. Large teams develop and small teams disrupt science and technology. Nature 566, 378–382. https://doi.org/10.1038/s41586-019-0941-9
Wuchty, S., Jones, B.F., Uzzi, B., 2007. The Increasing Dominance of Teams in Production of Knowledge. Science 316, 1036–1039. https://doi.org/10.1126/science.1136099

Appendix A: Potential PIs' selection validity test

Crucial to our analysis is the selection of the potential PIs. To identify the potential PIs, we assumed that the PL draws from 3 pools of researchers: (i) researchers with whom the PL collaborated in the past, (ii) reputed colleagues she knows the work, and (iii) researchers working in departments well-known for having competencies functional to complete the project proposed to ESF. Although these seem reasonable ways for a PL to select PIs, there might be others. In this Appendix, to validate our assumption on the potential sources of talents, we assess how many of the observed PIs have been drawn by the PL from the three pools. Table A.1 shows the results of this validation exercise. Specifically, 53% of the PIs selected in the application teams are drawn from one of the three pools. This high value shows that PL co-authors, referenced authors, and PIs' departmental colleagues are pools of researchers relevant in the PL's team selection process.

Table A.1 Average number of observed PIs per project application drawn from the three pools and share of researchers drawn from each pool.

	Pool 1 PL's co-authors	Pool 2 PL's cited authors	Pool 3 Colleagues of the observed PIs	Average observed PIs per project
Observed PIs	33%	20%	78%	3.83 overall, 2.04 (53%) in the three pools

The UNU-MERIT WORKING Paper Series

2022-01 *Structural transformations and cumulative causation towards an evolutionary micro-foundation of the Kaldorian growth model* by André Lorentz, Tommaso Ciarli, Maria Savona and Marco Valente
2022-02 *Estimation of a production function with domestic and foreign capital stock* by Thomas Ziesemer
2022-03 *Automation and related technologies: A mapping of the new knowledge* base by Enrico Santarelli, Jacopo Staccioli and Marco Vivarelli
2022-04 *The old-age pension household replacement rate in Belgium* by Alessio J.G. Brown and Anne-Lore Fraikin
2022-05 *Globalisation increased trust in northern and western Europe between 2002 and 2018* by Loesje Verhoeven and Jo Ritzen
2022-06 *Globalisation and financialisation in the Netherlands, 1995 – 2020* by Joan Muysken and Huub Meijers
2022-07 *Import penetration and manufacturing employment: Evidence from Africa* by Solomon Owusu, Gideon Ndubuisi and Emmanuel B. Mensah
2022-08 *Advanced digital technologies and industrial resilience during the COVID-19 pandemic: A firm-level perspective* by Elisa Calza Alejandro Lavopa and Ligia Zagato
2022-09 *The reckoning of sexual violence and corruption: A gendered study of sextortion in migration to South Africa* by Ashleigh Bicker Caarten, Loes van Heugten and Ortrun Merkle
2022-10 *The productive role of social policy* by Omar Rodríguez Torres
2022-11 *Some new views on product space and related diversification* by Önder Nomaler and Bart Verspagen
2022-12 *The multidimensional impacts of the Conditional Cash Transfer program Juntos in Peru* by Ricardo Morel and Liz Girón
2022-13 *Semi-endogenous growth in a non-Walrasian DSEM for Brazil: Estimation and simulation of changes in foreign income, human capital, R&D, and terms of trade* by Thomas H.W.Ziesemer
2022-14 *Routine-biased technological change and employee outcomes after mass layoffs: Evidence from Brazil* by Antonio Martins-Neto, Xavier Cirera and Alex Coad
2022-15 *The canonical correlation complexity method* by Önder Nomaler & Bart Verspagen
2022-16 *Canonical correlation complexity of European regions* by Önder Nomaler & Bart Verspagen
2022-17 *Quantile return and volatility connectedness among Non-Fungible Tokens (NFTs) and (un)conventional assets* by Christian Urom, Gideon Ndubuisi and Khaled Guesmi
2022-18 *How do Firms Innovate in Latin America? Identification of Innovation Strategies and Their Main Adoption Determinants* by Fernando Vargas
2022-19 *Remittance dependence, support for taxation and quality of public services in Africa* by Maty Konte and Gideon Ndubuisi
2022-20 *Harmonized Latin American innovation Surveys Database (LAIS): Firm-level microdata for the study of innovation* by Fernando Vargas, Charlotte Guillard, Mónica Salazar and Gustavo A. Crespi
2022-21 *Automation exposure and implications in advanced and developing countries across gender, age, and skills* by Hubert Nii-Aponsah

2022-22 *Sextortion in access to WASH services in selected regions of Bangladesh* by Ortrun Merkle, Umrbek Allakulov and Debora Gonzalez

2022-23 Complexity research in economics: past, present and future by Önder Nomaler & Bart Verspagen

2022-24 *Technology adoption, innovation policy and catching-up* by Juan R. Perilla and Thomas H. W. Ziesemer

2022-25 *Exogenous shocks and proactive resilience in the EU: The case of the Recovery and Resilience Facility* by Anthony Bartzokas, Renato Giacon and Corrado Macchiarelli

2022-26 *A chance for optimism: Engineering the break away from the downward spiral in trust and social cohesion or keeping the fish from disappearing* by Jo Ritzen and Eleonora Nillesen

2022-27 *Dynamic dependence between clean investments and economic policy uncertainty* by Christian Urom, Hela Mzoughi, Gideon Ndubuisi and Khaled Guesmi

2022-28 *Peer networks and malleability of educational aspirations* by Michelle Gonzalez Amador, Robin Cowan and Eleonora Nillesen

2022-29 *Linking the BOPC growth model with foreign debt dynamics to the goods and labour markets: A BOP-IXSM-Okun model* by Thomas H.W. Ziesemer

2022-30 *Countries' research priorities in relation to the Sustainable Development Goals* by Hugo Confraria, Tommaso Ciarli and Ed Noyons

2022-31 *Multitasking* by Anzelika Zaiceva-Razzolini

2022-32 *Informal institution meets child development: Clan culture and child labor in China* by Can Tang and Zhong Zhao

2022-33 *Income per-capita across-countries: Stories of catching-up, stagnation, and laggardness* by Juan Ricardo Perilla Jimenez

2022-34 *Mission-oriented R&D and growth of Japan 1988-2016: A comparison with private and public R&D* by Thomas H.W. Ziesemer

2022-35 *The macroeconomic implications of financialisation on the wealth distribution - a stock-flow consistent approach* by Huub Meijers and Joan Muysken

2022-36 Community Mobilization as a tool against sexual and gender-based violence in SADC region by Choolwe Mphanza Muzyamba

2022-37 *The empirics of technology, employment and occupations: Lessons learned and challenges ahead* by Fabio Montobbio, Jacopo Staccioli, Maria Enrica Virgillito & Marco Vivarelli

2022-38 *Money is not enough: Unintended negative effects of cash transfer design* by Juan Carlos Palacios Mora, Denis de Crombrugghe, Franziska Gassmann

2022-39 *Climate change and migration: Reviewing the role of access to agricultural adaptation measures* by Manisha Mukherjee

2022-40 *Gender bias in team formation: The case of the European Science Foundation's grants* by Michele Pezzoni and Fabiana Visentin