Impact of family characteristics on the gender publication gap:

Evidence for physicists in France

Jacques Mairesse[a], Michele Pezzoni[b] and Fabiana Visentin*[c]

[a] CREST ENSAE (France); UNU-MERIT, Maastricht University (Netherlands);
EHESS (France); NBER (USA); email: *mairesse@ensae.fr*
[b] Université Côte d'Azur, CNRS, GREDEG (France); ICRIOS, Bocconi University (Italy);
email: *michele.pezzoni@unice.fr*;
[c] UNU-MERIT, Maastricht University (Netherlands); email: *visentin@merit.unu.edu*

*Corresponding author

Impact of family characteristics on the gender publication gap:

Evidence for physicists in France

Abstract

The publication gender gap in science has been largely studied. Despite women have been found to be less productive than men, little is known about the reasons behind gender differences. Unique longitudinal data collected by surveying a large sample of French physicists gave us the opportunity to investigate the role of family characteristics over time. Using panel data econometrics, we confirm the existence of an average gender gap of about two-third of a journal article per year, and of about one-third when taking into account several important control variables as age and career characteristics. We find that female scientists suffer an average productivity loss of about one article when they have a small child, while male scientists suffer a non-significant loss. We also find that female scientists benefit from having large families, with a productivity gain of 0.63 articles per year per child.

Keywords: Gender gap, Publication productivity, Family characteristics, Panel data, Physics.

Subject classification codes: I23, J13

Introduction

Since 1901, 904 individuals have been awarded a Nobel Prize and among those Nobel laureates, only 51 are women. Female scientists not only are underrepresented as recipients of the most prestigious award in science, the Nobel Prize, but have been found to have lower salaries than men (Ginter and Hayes, 2003), to encounter greater difficulties in obtaining career promotions (Long et al., 1993) and in accessing resources (Witterman et al., 2019). While women are increasing participating in science[1], they still have difficulties in succeeding in science and the scientific community is debating about the reasons behind the phenomenon (for an overview on the debate, see the recent Special issue appeared in the Lancet journal dedicated to the gender gap in science, Shannon et al. 2019).

One of the reasons that might explain gender inequality in scientific careers and access to resources is the women's lower publication productivity (Zuckerman 1987; Levin and Stephan 1998; Bentley and Adamson 2003; Duch et al. 2012). Scientists are mainly evaluated on their scientific outcome as measured by publications productivity: a strong publication record guarantees scientists' reward in terms of career progress and resource access (Stephan, 2012). If women have lower publication productivity than men, especially in the early stages of their career, they will have also more difficulties in obtaining career promotions and access to resources (Merton, 1968). Our paper contributes to the debate about the gender gap in science by looking at the factors that might explain differences in publication productivity between women and men.

Several studies quantify the gender gap in publication productivity. For instance, West et al. (2013) map differences in gender ratios across research disciplines finding that even if female

[1] According to the US National Science Foundation and the Eurostat data, the share of women doctorate recipients is increasing and has almost reached the parity across fields (Sources: US National Science Foundation Survey of Earned Doctorates (SED) and Eurostat (Science and Technology Database, Statistics on Research and Development).

authors are increasingly publishing, they are still far to reach equality with large differences across fields. According to their study women progressed from 15.1% of authorships in 1665-1989 to 27.2% of authorships in 1990-2012. Fields like mathematics remain male dominated (10.6% of women' authorships) while others like education are almost balanced (46.3% of women' authorships). Larivière et al. (2013) analyse gender disparity across geographical regions showing that gender disparity remains widespread worldwide. More recently Holman, Stuart-Fox, and Hauser (2018) moved a step further the gender gap mapping, adding a formal modelling to predict when the gap will close. They estimate that "many research specialties (e.g., surgery, computer science, physics, and math) will not reach gender parity this century" and claim that their results "support a need for further reforms to close the gender gap". Those reforms would need an understanding of the mechanisms behind gender disparities, however little is known about them. Our study aims to add new insights in disclosing these mechanisms by investigating the impact of family characteristics on publication productivity by gender.

Cole and Zuckerman (1984) coined the term "productivity puzzle" to identify the unsuccessful attempts to explain the documented gender gap in publication productivity and, after almost forty years, the puzzle still remains unsolved. Extant studies have considered a broad set of explanatory factors such as age, family characteristics, tenure, and academic ranking. Among the explanatory factors, family characteristics have been pointed out as the most relevant factors. However, extant studies that look at the childrearing responsibilities found divergent findings (Mary Frank Fox 1981; Mary Frank Fox and Faver 1985; Zuckerman, Cole, and Bruer 1991; Long, Allison, and McGinnis 1993). The heterogeneity of those findings can be attributed to the lack of appropriate controls in the econometric analysis. For instance, information on individual ability, commitment to work or family life, and social and cultural background are often neglected since they are difficult to measure. Most of the extant empirical analyses are based on samples of scientists observed at a given point in time, cross-section data (with some notable exceptions such as Long, 1990). Cross-sectional data allow comparing the productivity of individuals with different measurable characteristics. Differently, from cross-sectional studies, longitudinal studies allow to observe the same individual over time and to control for measurable and unmeasurable time-invariant characteristics (Wooldridge, 2006). New in our work is the use of fine-grained information on the scientists' family status to exploit the complexity of childbearing responsibilities. Additionally, we use a longitudinal dataset to proper overcome the issue of controlling for unmeasurable time-invariant characteristics.

We survey the entire population of French physicists active at the Institute of Physics (INP) in June 2017. Our survey has been addressed to 1,085 individuals and obtained positive replies from 57,23% of those individuals. For the 621 respondents, we were able to reconstruct their publication record, family characteristics, and career advancement history over the period 2001-2016. Physics, being one of the disciplines so-called 'male-dominated', appears a field where the understanding of differences in gender dynamics deserve exploration and the high-reputation of French physicists in the international scientific community makes France the ideal country to conduct our investigation. While the large majority of extant studies consider the US context (Aguinis, Ji, and Joo 2018), our focus on France allows us to add new insights about the productivity of European women in science.

We find that female scientists publish on average 0.68 articles less per year than male scientists. Controlling for the biographical, career, and family characteristics, female scientists the gap reduces to 0.32 papers per year. When we further investigate the role played by family characteristics, we discover that females benefit from having larger families while suffer from having children less than three-year-old. Differently, male scientists are not affected by the family size and by the presence of a child. Considering these findings, in the final part of our work we frame important policy suggestions on the possible interventions that can help females to reach equality with their male colleagues in science.

Family characteristics and the gender gap: A review of the extant studies

For a long time, scholars have debated on the existence of a differential in productivity between female and male scientists and on the reasons behind it. In the ‘80s Cole and Zuckerman (1984) coined the term “productivity puzzle” to point out the lack of definitive explanations for the observed lower productivity of females in science. Extant studies identify three main explanations for the puzzle: biological differences in abilities between females and males, gender discrimination, and differences in life choices. Up to date, empirical studies seem to exclude biological differences in abilities and are not convergent in detecting gender discrimination by journal reviewers, granting agencies or recruitment committees (Aguinis, Ji, and Joo 2018; Stephen J. Ceci and Williams 2011, 2010). The most credible explanation is that females and males seem oriented to make different life choices. According to this explanation, females are penalized by societal and cultural constraints that impose them to be the main childcare giver, to move to follow their partners' career needs, and to take care of old parents.

Childrearing responsibilities are an important part of work-home balance choices, and several studies have attempted to quantify the impact of having children both for females and males (Stephen J. Ceci and Williams 2011, 2010). Those studies have not yet reached convergent findings. For instance, Fox and Faver (1985) find a positive effect of having children on female scientists’ productivity, while no effect exists for male scientists. Stack (2004) and Mary Frank Fox (2005) find a limited positive effect of having children for both female and male scientists.

Other studies find negative effects of children on scientists' productivity (Long 1990; Kyvik 1990; Zuckerman, Cole, and Bruer 1991; Kyvik and Teigen 1996). Finally, no effect results in the studies of Toren (1991) and Sax et al. (2002).

The heterogeneity of results can be attributed to three main limitations: an unaccounted complexity of childbearing responsibilities, lack of appropriate controls in the econometric analysis, and scarcity of longitudinal data. Our study contributes to the gender literature by addressing those three limitations.

First, childbearing responsibilities have often been considered in a simplified way. Several studies do not account for children age, or for the interplay between children age and number of children within the family. Importantly, children do not impact equally over time on scientists' productivity, i.e. children in their pre-schooling age are more time demanding than older children and are associated with the loss in productivity (Mary Frank Fox and Faver 1985; Stack 2004). When considering the number of children, Leslie (2007) shows that when the number of children increases men and women adapt the amount of time devoted to their academic work. The hours worked per week decrease for females, while worked hours increase for males. Mary Frank Fox (2005) finds that a greater number of children has a positive but not significant effect on female scientists' productivity. Our study takes into account the children age distinguishing between children in the pre-maternity school period, i.e., the highest time demanding period, from older children. Moreover, it reconstructs the complexity of the family status both in term of children age and number of children.

Second, the studies listed above on the relationship between childbearing responsibilities and scientists' productivity consider different sets of control variables. While Mary Frank Fox and Faver (1985) and Stack (2004) add controls for scientists' personal characteristics, academic rank and co-authorships, other studies assess the publication productivity average without considering controls. Several factors might be correlated both with having children and with publication productivity and not controlling for these factors might bias the estimated effects of having children on productivity. For instance, having a child might interfere with the scientist's publication outcomes making less likely a career advancement and the access to additional resources granted by a higher academic rank. Therefore, females might experience a disadvantage with respect to male colleagues since the early phases of their career overlap with the age in which women are more likely to have a child (Stephen J. Ceci, Williams, and Barnett 2009). To overcome these econometric concerns, in our study we exploit a fine-grained set of controls including academic rank and individual characteristics. Additional controls on the work context are unnecessary since we consider a homogeneous sample of scientists belonging to the same institute, the INP.

The third limitation of the extant studies, i.e. the scarcity of longitudinal data, relates to the data gathering. The large majority of extant studies are based on cross-sectional data that gives a snapshot of family status and scientists' profile on a precise point in time (with some exceptions, e.g. Long 1992; Mairesse and Pezzoni 2015). In assessing the causal relationship between productivity and family status, those studies make a comparison across individuals at the same point in time. For instance, an individual A with three children is compared with and individual B with one child. However, A and B might also differ in other characteristics. Stack (2004) claims that '[female and male scientists] who have children [might have] higher organizational skills, energy, stamina, health, or other qualities which are also related to research productivity'. Similarly, personal factors like family-orientation attitude might affect the decision of having a child as well as productivity. The remarkable advantage of having longitudinal data is that we can access for the same individual the effect of changes in her family status. We compare the individual A at different points in time, when family characteristics changes while individual peculiar characteristics remain unchanged. With the addition of scientists' fixed effects, we control for all the unobservable time-invariant characteristics obtaining an unbiased estimation of the effect of family characteristics on scientists' productivity (Wooldridge, 2006).

Empirical setting: The Institute of Physics in France

The empirical context for investigating the role of family characteristics in explaining the gender gap in science is a large group of physicists affiliated to the Institute of Physics (INP) in France.

France has a reputation for producing cutting-edge research in the field. Looking back at the history of physics, important discoveries are attributed to French scientists. For instance, the international system of measurement units of electric current, i.e. the *ampere,* was introduced by André-Marie Ampère who was one of the founders of the science of electromagnetism. Several discoveries were awarded the highest worldwide reward, the Nobel Prize. For example, in 1903, Marie Curie was awarded for her pioneering research on radioactivity. French scientists experiment successfully in the fields of traditional as well as modern physics like quantum electronics. In 2012, Serge Haroche was awarded the Nobel Prize in Physics for his studies on the measurement and manipulation of quantum systems. In 2018, Gérard Albert Mourou was awarded for his discoveries on very high-intensity laser pulses.

The INP is the physics French national institute. It is part of a public organization under the responsibility of the French Ministry of Education and Research, the National Center for Science Research (CNRS). INP aggregates scientists with high academic profile and productivity.

Publication data retrieved from the Web of Science bibliometric dataset (Clarivate Analytics) shows that during the period 2001-2016, INP produced approximately 52,000 publications in Physics. Researchers affiliated to INP are French civil servants and follow a well-defined career progression that counts five career levels. Specifically, a researcher enters at INP as (1) Junior Researcher Second Class and, later in his career, she can be upgraded to (2) Junior Researcher First Class, (3) Senior Researcher Second Class, (4) Senior Researcher First Class and, if she is recognized for her outstanding academic achievements, she can reach the level of (5) Senior Researcher Exceptional Class. Researchers are appointed according to their expertise, and, in each career level, they have different responsibilities and salaries. A Junior Research conducts independent research under the supervision of a Senior Researcher who runs her own lab. Salaries vary from 2,000 euros for a Junior Researcher Second Class to 6,000 euros for a Senior Researcher Exceptional Class[2].

Being part of the French system, INP's researchers benefit from the national social security system. On this regard, France, as the majority of the EU countries, address a set of instruments to support parents when children are in their first years of life. Specifically, the French government considers as a critical turning point for childbearing the child's third birthday. For instance, parents can opt for reducing their working hours until their children are 3-year old[3]. Moreover, starting from this age the large majority of children attends the Nursery School ('Ecole Maternelle') and, to consolidate this habit of sending children to a pre-school, the French government intends to reduce the compulsory school age from 6 to 3 years starting from 2019.

[2] See the CNRS website for details on INP scientists' work conditions (http://www.dgdr.cnrs.fr/drhchercheurs/Travail/concoursch2010/chercheur/carriere-en.htm).
[3] For further details about the French Social Security System, see https://www.cleiss.fr/docs/regimes/regime_france/an_1.html

Data

Data collection

We surveyed all French physicists affiliated to INP and active in June 2017. We contacted 1,085 scientists working on the five main study areas in which the institute specializes: theoretical physics, condensed matter and optics, atoms, molecules, and quantum physics. The survey was developed and conducted in French avoiding challenges of translation. We surveyed scientists on their family characteristics. Specifically, we inquired if a scientist had children, and – if this was the case – we asked to detail the birth year of each child. We reconstructed from their Curricula their career advancements, identifying the year of promotion (if any) from Junior Researcher Second Level to Senior Research Exceptional Level. To complete the tracking of their career advancements, in the survey, we asked to specify the beginning and end year of promotion to head of a research unit, if any occurred. The survey was launched on line: the INP administration provided us the complete list of their scientists with their email contacts. Each scientist of the list received a link to access and fill a form with our set of questions[4]. In total, we received 621 usable answers. The overall response rate is 57,24%. We obtained 364 responses after the first round (June 2017), and 257 additional responses with the first and second reminders (177 and 80 additional responses, respectively in July 2017 and January 2018). We performed a series of statistical tests showing that respondents are similar to non-respondents (see Appendix A). Female responded in the 68.27% of the cases, while males responded in 54.62% of the cases.

We complemented the survey data with demographic data obtained from the French Ministry of Education and we reconstructed the full scientists' publication record by collecting her publications on the Web of Science dataset (Clarivate Analytics). In collecting the publications,

[4] Our questionnaire is available upon request.

we included only publications above a certain threshold of quality. Specifically, publications that appeared in journals with a 5-year impact factor greater than 0.5, meaning that the articles published in these journals received on average more than half a citation in the last five years. Journals with 5-year impact factor greater than 0.5 correspond to the 99% of entire poll of Physics journals covered by Web of Science. We end up with a sample of 273 journals.

Data description

The way in which we framed the questions allowed us to build a longitudinal dataset that is a remarkable advancement compared to similar studies. While extant studies used surveys to capture a snapshot of the situation in a given period of time, having asked the exact birth year of each child and the exact period when they had been the head of a research unit, we reconstructed the yearly scientist' family and academic rank. Observing each of our 621 scientists along the observation period 2001-2016, we end up with 9,021 observations. Not all our scientists are present in the observation period 2001-2016, due to their different entry and exit year at INP. In our sample, 67,3% of scientists are observed for the full period, whereas the minimum number of observations per scientist is 5 years (0.48% of the cases).

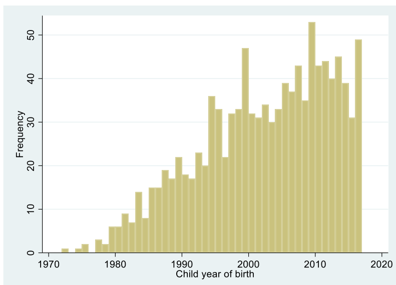
Table 1 and Figure 1 show a picture of our scientists' family status. According to Table 1, the most frequent case is having two children (36,88% of the total sample: 35.7% of the men has two children, 40.85% of women has the same family size). It is rare to have more than four children (1.45% of the total sample, 1.67% of men, 0.7% of women). The 22.55% of males have no children, 19.01% of females are in the same situation (21,74% of the total sample).

Figure 1 shows the distribution of the children birth years. 248 scientists had the first child during our observation period, i.e., from 2001 to 2016.

Table 1. Family size

Number of children	Female	Male	Total
0	27	108	135
1	28	70	98
2	58	171	229
3	19	99	118
4	9	23	32
5	1	5	6
6 and more	0	3	3
Total	142	479	621

Figure 1. Children year of birth



The 19.42% of males (93/479) have had the responsibility of a research unit (head of a research unit), 14.08% of female (20/142) had covered the same position.

Looking at the demographic and academic profile, scientists are middle age (the average age is about 41 for both males and females) with on average 2.66 publications per year (2.82 for males, 2.14 for females).

Scientists in our sample are almost equally distributed across the INP sub-disciplines: 18.68% belongs to Physical theories, 24.48% to Condensed Matter Physics with a focus on structures and electronic properties, 28.66% to Atoms and Molecules, Optics and Lasers and the remaining 28.18% to Condensed Matter Physics with a focus on organizations and dynamics.

Variable construction

Based on the survey data, the data released by the French Ministry of Education, and the bibliometric data collected on the Web of Science database, we calculate our variables of interest and controls. We measure scientist's publication productivity as the *Number of publications* authored by the scientist in a given year *t*. Among the variables related to the scientist's demographic characteristics, we consider the dummy variable *Female* that equals one if the scientist is a woman, zero otherwise, and *Age* that corresponds to the age of the scientist in year *t*[5]. To measure family characteristics, we look at the *Total number of children* that equals to the number of children born until year *t.* Considering that children have different needs at the various stages of their life, we distinguish the cases where in the family there is a child in her first years of life. The dummy variable *One child born in the last 3 years* equals one if the scientist has at least one child born in the last three years, i.e. in *t*, *t-1*, and *t-2*. The French government considers 3-year old as the maximum age until when parents can opt for reducing their working-time (see the section Empirical setting). We assume that 3-years old is the threshold since when parents devote more time to their child. Then, to take into account the effect of the career advancement on scientists' productivity, we construct a set of dummy variables reflecting the scientists' career level. The dummy variables *Junior researcher (Second class)*, *Junior researcher (First class)*, *Senior researcher (Second class)*, *Senior researcher (First class)*, and *Senior researcher (Exceptional)*, assume the value one if the scientist is at the corresponding career level, zero otherwise. We also consider the period of the scientist's career before being appointed at INP (*Pre-entry INP*). Moreover, we include the dummy variable *Head of a research unit in the last 3 years* that equals one if the scientist had the full responsibility of the research unit in *t*, *t-1*, and *t-2.* To

[5] The variable age is centred on the age of 40, meaning that we subtract 40 to the actual age at time *t*

account for the past publication productivity, the variable *Initial number of papers* counts the number of papers published by the scientist in the year before entering the panel. Finally, we construct a dummy variable for each physics sub-discipline (Physical theories, Condensed matter physics -structures and electronic properties-, Atoms and Molecules, Optics and Lasers, Hot Plasma Physics, and Condensed matter physics -organizations and dynamics-) and for each calendar year from 2001 to 2016.

Table 2 shows the corresponding descriptive statistics at scientist-year level.

Table 2. Descriptive statistics

	Female (2,074)		Male (6,947)		All (9,021)	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Number of publications	2.14	4.61	2.82	5.77	2.66	5.53
Publishing in year *y*	0.62	0.49	0.73	0.45	0.7	0.45
Female	1	0	0	0	0.23	0.42
Age	41.69	9.32	41.36	9.09	41.43	9.14
Total number of children	1.39	1.24	1.37	1.29	1.38	1.28
No children	0.34	0.47	0.36	0.48	0.35	0.48
One child born in the last 3 years	0.16	0.36	0.19	0.39	0.18	0.39
Director of research	0.26	0.44	0.32	0.47	0.31	0.46
Pre-entry INP	0.17	0.37	0.17	0.38	0.17	0.38
Junior researcher (Second class)	0.11	0.31	0.12	0.32	0.11	0.32
Junior researcher (First class)	0.46	0.50	0.39	0.49	0.40	0.49
Senior researcher (Second class)	0.20	0.40	0.23	0.42	0.23	0.42
Senior researcher (First class)	0.06	0.24	0.08	0.27	0.077	0.27
Senior researcher (Exceptional)	0.005	0.69	0.007	0.087	0.007	0.083
Head of a research unit	0.07	0.25	0.08	0.27	0.08	0.27
Initial number of papers	1.99	3.83	2.49	3.59	2.37	3.66
Section 2: Physical theories	0.1	0.3	0.21	0.41	0.19	0.39
Section 3: Condensed matter physics (structures and electronic properties)	0.22	0.41	0.26	0.44	0.25	0.43
Section 4: Atoms and Molecules, Optics and Lasers, Hot Plasma Physics	0.3	0.46	0.28	0.45	0.28	0.45
Section 5: Condensed matter physics (organizations and dynamics)	0.38	0.49	0.25	0.44	0.28	0.45

Method and results

The aim of this study is to investigate the impact of the family characteristics on the gender publication gap. To this purpose, we proceed in two steps. First, we conduct an exploratory analysis based on the descriptive statistics of our data and, second, we estimate the parameters (β) of Equation 1 with a formal multiple regression model applying Ordinary Least Squares (OLS).

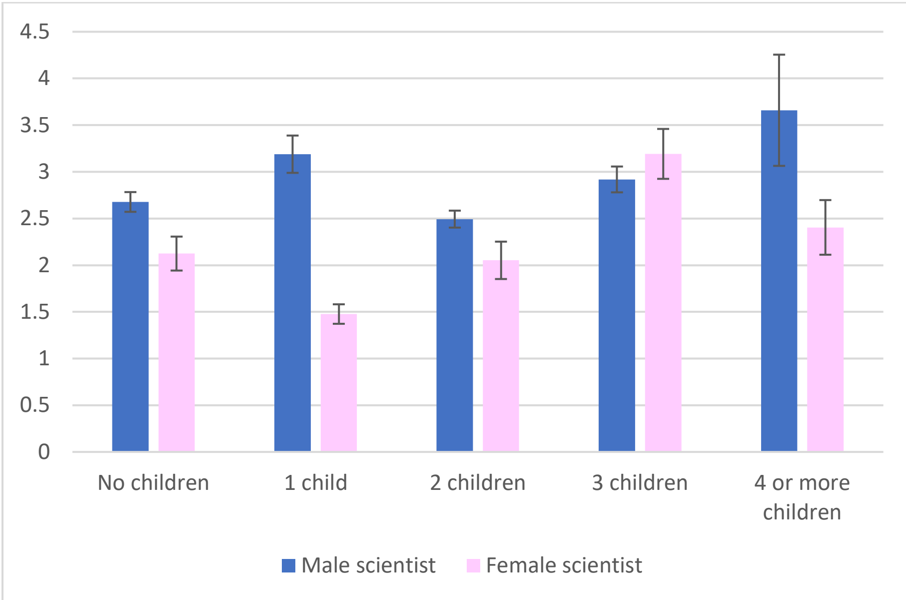
$$Publication\ productivity_{it} = \beta_0 + \beta_1\ demographic\ characteristics_{it} + \beta_2\ family\ characteristics_{it} + \beta_3\ Career\ advancement_{it} + \varepsilon_{it}$$

(Equation 1)

Where *i* and *t* refer to the scientist *i* observed at time *t*. Our 621 individuals are observed over the period 2001-2016, for a total of 9,021 observations. Each individual is observed on average for 14.52 years.

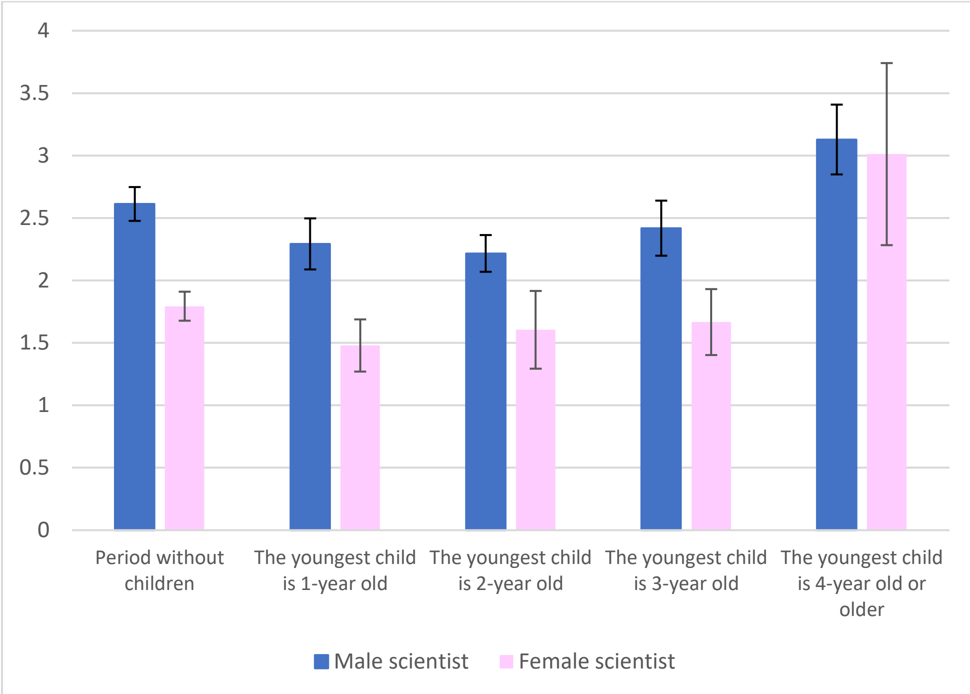
As explorative analysis, Figures 2 and 3 visualize how publication productivity varies by family size and children age. Comparing the productivity of scientists across a number of children classes (Figure 2), we observe males being more productive when they have one child or four children or more, while they show lower productivity when they have no children or a medium size family (two or three children). Females show lower productivity when they have one child, while they show higher productivity when they have three children. Male and female scientists show the closest productivity values in the case of 3-children families.

Figure 2. Publication productivity by family size (standard errors are reported on the top of the bars)



When adding the temporal dimension into the analysis, Figure 3 shows that when children grow up gender disparities attenuates. In our study sample, 38.33% of scientists (238 scientists) enter the observation period already having children while 21.74% (135 scientists) remain for the entire observation period without children. To quantify the impact of having children over time, we consider the subsample of 248 scientists (39.94%) for which we can observe the full family history, i.e. those scientists who had the first child during our observation period. Figure 3 shows that male and female scientists with children older than 4 years have no significant differences in publication productivity. When the youngest child is younger than 3-year old, we observe lower publication productivity both for female scientists (about one publication less) and for male scientists (about half a publication less).

Figure 3. Publication productivity by children age in the sub-sample of 248 scientists for which we can observe the full family history.



In Table 3, we present the results of the OLS estimation of Equation 1. In columns from 1 to 5, we progressively add to scientists' demographic characteristics, controls for family characteristics and for career advancement. Column 1 presents the baseline model including demographics' characteristics (being a *Female* scientist, and *Age*) and scientist's productivity at the beginning of the observation period (*Initial* number *of papers*). Column 2 adds controls for career advancement (from the entry level - *Junior researcher (Second class)* – to the highest one - *Senior researcher (Exceptional)* - and being the *Head of a research unit*). Columns 3 and 4 separately include controls for family characteristics: the total number of children (*Total number of children*), and a dummy if the youngest child is 3-year old or younger (*One child born in the last 3 years*) respectively. Column 5 includes both controls for family characteristics at the same time.

According to Column 1, female scientists publish 0.29 papers less than male scientists per year. The gender gap in favour of males increases up to 0.31 papers when we add career advancement and family characteristics. The impact of controls on publications productivity is in line with conventional expectations: young scientists are more productive, having a high-level managerial tasks as head of a research unit has a negative effect on publication productivity, high productivity tend to persist[6]. Interestingly, having a little child has a detrimental effect on scientists' productivity: -0.30 papers per year (Column 5).

The latter results on family characteristics might be biased by a series of omitted unobservable variables such as the commitment to work, the willingness to create a family, or other cultural and social background characteristics. To further investigate the family characteristics impact correcting for these omitted unobservable factors, we include scientists' fixed effects and we conduct separate analyses on the sub-samples of male (479 scientists observed yearly for a total of 6,947 observations) and female scientists (142 scientists observed yearly for a total of 2,074 observations). Table 4 reports the results. In columns from 1 to 4, we progressively add controls for family characteristics and career advancements. The model in column 1 includes controls for scientist's age (*Age*), academic rank (*Junior researcher* and *Senior researcher* dummies), being the *Head of a research unit* and *Scientists' fixed effects*. Columns 2 and 3 include separately the controls for family characteristics, i.e., the total number of children (*Total number of children*) and a dummy that equals one if the youngest child is 3-year old or younger (*One child born in the last 3 years*). The full model reported in Column 4 includes both the family characteristics at the same time. The values of the R-squared reported in Column 1, 2, 3 and 4 of

[6] There is a positive correlation between the number of papers published by a scientist when she enters the observation period and the later yearly productivity. One paper more in the initial stock guarantee 0.72 paper more by year.

Table 4 indicate the proportion of the variation of the publication productivity that our models explain. In a range that goes from 0 to 1, all the R-squared values are above 0.54 indicating that our models are explaining more than half of the publication productivity variation. The increase of the R-squared values between the models reported in Table 3 and those reported in Table 4 including scientist fixed effects, shows the importance of controlling for the time-invariant unobserved characterises of the scientists in explaining their publication productivity.

Interestingly, results of the full model reported in Column 4 show than when scientists are appointed *Head of a research unit* they decrease significantly their productivity (-0.52 papers per year for males, -0.78 papers per year for females). For male scientists, having a larger family or having at least one little child in the family does not affect publication productivity. Female scientists have productivity gain when the number of children increases (+0.63 papers per year for each child). However, for females having a young child has detrimental effects on publication productivity (-0.99 papers per year).

Table 3. Publication productivity gender gap, OLS estimations

	(1)	(2)	(3)	(4)	(5)
	All	All	All	All	All
Dependent variable:	N. pubs	N. pubs	N. pubs	N. pubs	N. pubs
Female	-0.29**	-0.31**	-0.31**	-0.32***	-0.32***
Age	-0.068***	-0.065***	-0.066***	-0.068***	-0.072***
Total number of children			0.031		0.065
One child born in the last 3 years				-0.23	-0.30**
Junior researcher (Second class)		-0.21	-0.21	-0.15	-0.15
Junior researcher (First class)		0.12	0.099	0.18	0.15
Senior researcher (Second class)		0.077	0.050	0.12	0.078
Senior researcher (First class)		-0.23	-0.26	-0.18	-0.23
Senior researcher (Exceptional)		-0.21	-0.24	-0.16	-0.20
Head of a research unit		-0.49**	-0.49**	-0.49**	-0.49**
Initial number of papers	0.72***	0.72***	0.72***	0.72***	0.72***
Constant	0.58**	0.61**	0.59**	0.61**	0.58**
Scientists' fixed effects	No	No	No	No	No
Observations	9,021	9,021	9,021	9,021	9,021
R-squared	0.227	0.228	0.228	0.228	0.228
Dummy calendar year	Yes	Yes	Yes	Yes	Yes
Dummy section	Yes	Yes	Yes	Yes	Yes
Number of scientists	621	621	621	621	621

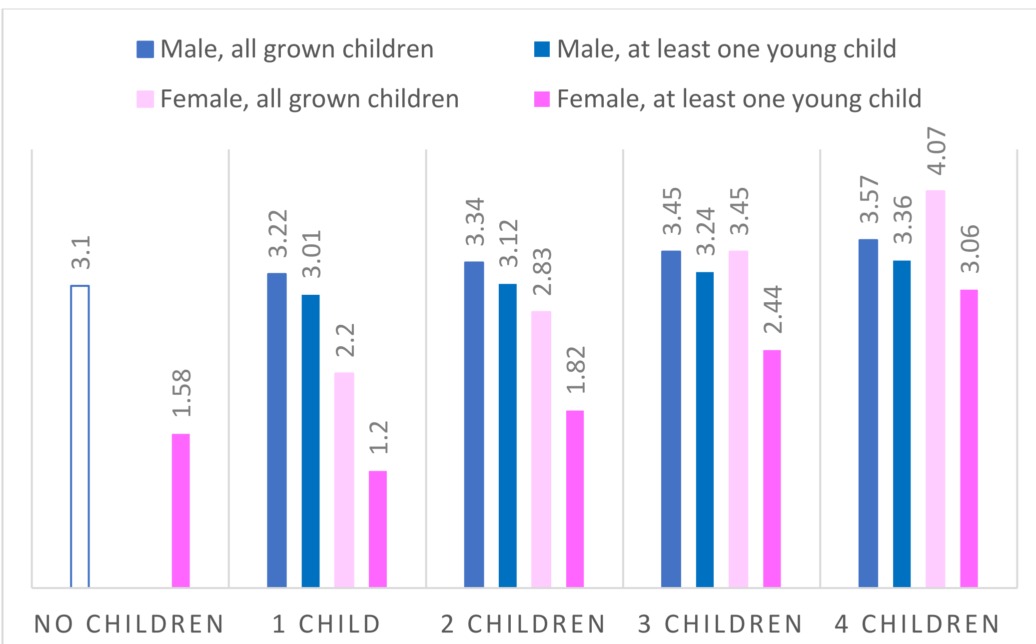
Table 4. Family characteristics, scientists' fixed effect estimations

	(1)		(2)		(3)		(4)	
	Male	Female	Male	Female	Male	Female	Male	Female
Dependent variable:	N.pubs	N.pubs	N.pubs	N.pubs	N.pubs	N.pubs	N.pubs	N.pubs
Age	0.0080	0.29	0.0045	0.29	0.0076	0.27	-0.00043	0.25
Total number of children			0.050	0.30*			0.11	0.63***
One child born in the last 3 years					-0.16	-0.69***	-0.22	-0.99***
Junior researcher (Second class)	0.32	0.31	0.31	0.25	0.36*	0.46	0.36*	0.39
Junior researcher (First class)	0.86***	-0.0054	0.83***	-0.21	0.91***	0.20	0.85***	-0.13
Senior researcher (Second class)	0.53	-0.28	0.50	-0.45	0.55	-0.12	0.50	-0.40
Senior researcher (First class)	-0.10	-0.75	-0.11	-0.86	-0.082	-0.57	-0.10	-0.72
Senior researcher (Exceptional)	-0.83	-0.64	-0.82	-0.69	-0.80	-0.42	-0.77	-0.41
Head of a research unit	-0.52*	-0.72*	-0.52*	-0.74*	-0.53*	-0.75*	-0.52*	-0.78*
Constant	2.09***	1.52*	2.04***	1.24	2.09***	1.51*	1.99***	0.94
Scientists' fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,947	2,074	6,947	2,074	6,947	2,074	6,947	2,074
R-squared	0.597	0.545	0.598	0.545	0.599	0.545	0.602	0.545
Dummy year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of individuals	479	142	479	142	479	142	479	142

Figure 4 provides a graphical representation of the regression results reported in Table 4, Column 4[7]. Specifically, it shows the predicted scientist's publication productivity for a representative scientist when family characteristics vary. For a representative scientist, we intend a junior researcher (first class), who is not the head of a lab, and 40-year-old. This scientist's profile is the most frequent in our dataset (see the average characteristics of the scientists in our sample reported in Table 2). For family characteristics, in Figure 4, we predict the publication productivity for female and male scientists, having or not a young child, and with different family sizes.

For the representative scientist, the figure shows the compensation effect between the number of children and children age. Small children penalize females, however female scientists catch up their male colleagues when they have a family of at least three grown up children.

Figure 4. Publication productivity predictions



[7]The predictions are based on the two models reported in Column 4 of Table 4 excluding calendar year dummies. We opted for this exclusion since calendar dummies are not statistically significant (nor for male neither for female). The exclusion also allows us to interpret predictions as the average all over the observation period.

Discussion and conclusions

This paper contributes to the gender literature by analysing the role played by family characteristics of male and female scientists' publication productivity. Unique to our study is the use of longitudinal data. We surveyed all the physicists affiliated to INP in June 2017 and reconstructed their full data record, including family status, academic rank, and publications over the period 2001-2016. We find that family characteristics play a different role for male and female.

The paper addresses a number of econometric challenges faced in the attempt to properly quantify the relationship between family related explanatory factors and publication productivity. Family characteristics are the observable characteristics most commonly suggested as an explanatory factor of the productivity gender gap (Stephen J. Ceci and Williams 2010, 2011). Few extant studies in the literature have been able to reconstruct the entire scientists' family history. Mary Frank Fox and Faver (1985) and Kyvik (1990) use cross-sectional data showing not convergent evidence on the effect of having children. By asking our surveyed scientists the exact date of birth of their children, we have been able to identify how female and male scientists' productivity vary when scientists' family characteristics change over the scientists' career. We find that while family characteristics do not impact on male scientists' productivity, while female scientists benefit in having a large family but are penalized in the early years of their children life. Concerning the number of children, we have two possible explanations for its positive impact on publication productivity. First, there might be an incentive to publish more and to progress in career to assure better life conditions to a larger family. Second, we cannot completely exclude a selection effect, i.e. there might be some unobserved time-variant characteristics that could be correlated both with the willingness to have a larger family and to higher publication productivity such as significant cultural value changes along the career of an individual.

The French government is investing a lot of efforts to support families. Parents are allowed to devote more time to family life when they have young children by reducing their working time. Later on, when parents are asked to go back to their normal worktime schedule, children enter the pre-school. Despite all these efforts, our results show that they are not enough, and greater investments are needed. The French government seems to perceive this need. For instance, the government intends to reduce the compulsory school age from 6 to 3 years starting from 2019 to consolidate the habit of sending children to pre-school.

References

Aguinis, Herman, Young Hun Ji, and Harry Joo. 2018. "Gender Productivity Gap among Star Performers in STEM and Other Scientific Fields." *Journal of Applied Psychology* 103 (12): 1283–1306. https://doi.org/10.1037/apl0000331.

Bentley, Jerome T., and Rebecca Adamson. 2003. "Gender Differences in the Careers of Academic Scientists and Engineers: A Literature Review. Special Report."

Ceci, Stephen J., and Wendy M. Williams. 2011. "Understanding Current Causes of Women's Underrepresentation in Science." *Proceedings of the National Academy of Sciences* 108 (8): 3157–62. https://doi.org/10.1073/pnas.1014871108.

———. 2010. "Sex Differences in Math-Intensive Fields." *Current Directions in Psychological Science* 19 (5): 275–79. https://doi.org/10.1177/0963721410383241.

Ceci, Stephen J., Wendy M. Williams, and Susan M. Barnett. 2009. "Women's Underrepresentation in Science: Sociocultural and Biological Considerations." *Psychological Bulletin* 135 (2): 218–61. https://doi.org/10.1037/a0014412.

Cole, Jonathan R., and Harriet Zuckerman. 1984. *The Productivity Puzzle: Persistence and Change in Patterns of Publication of Men and Women Scientists*. Vol. 2.

Duch, Jordi, Xiao Han T. Zeng, Marta Sales-Pardo, Filippo Radicchi, Shayna Otis, Teresa K. Woodruff, and Luís 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.

Fox, Mary Frank. 2005. "Gender, Family Characteristics, and Publication Productivity among Scientists." *Social Studies of Science* 35 (1): 131–50. https://doi.org/10.1177/0306312705046630.

———. 1981. "Sex, Salary, and Achievement: Reward-Dualism in Academia." *Sociology of Education* 54 (2): 71–84.

Fox, Mary Frank, and Catherine A. Faver. 1985. "Men, Women, and Publication Productivity: Patterns among Social Work Academics." *The Sociological Quarterly* 26 (4): 537–549.

Shannon, Geordan, Melanie Jansen, Kate Williams, Carlos Cáceres, Angelica Motta, Aloyce Odhiambo, Alie Eleveld, and Jenevieve Mannell. 2019. "Gender Equality in Science, Medicine, and Global Health: Where Are We at and Why Does It Matter?" *The Lancet* 393(10171):560–69.

Ginther DK, Hayes KJ. 2003. Gender Differences in Salary and Promotion for Faculty in the Humanities 1977–95. *Journal of Human Resources* 38(1): 34–73.

Holly O Witteman, Michael Hendricks, Sharon Straus, Cara Tannenbaum, Are gender gaps due to evaluations of the applicant or the science? A natural experiment at a national funding agency, *The Lancet*, Volume 393, Issue 10171, 2019, Pages 531-540, ISSN 0140-6736.

Holman, Luke, Devi Stuart-Fox, and Cindy E. Hauser. 2018. "The Gender Gap in Science: How Long until Women Are Equally Represented?" Edited by Cassidy Sugimoto. *PLOS Biology* 16 (4): e2004956. https://doi.org/10.1371/journal.pbio.2004956.

Kyvik, Svein. 1990. "Motherhood and Scientific Productivity." *Social Studies of Science* 20 (1): 149–160.

Kyvik, Svein, and Mari Teigen. 1996. "Child Care, Research Collaboration, and Gender Differences in Scientific Productivity." *Science, Technology & Human Values* 21 (1): 54–71.

Larivière, Vincent, Chaoqun Ni, Yves Gingras, Blaise Cronin, and Cassidy R. Sugimoto. 2013. "Bibliometrics: Global Gender Disparities in Science." *Nature News* 504 (7479): 211.

Leslie, David. 2007. "The Reshaping Of America's Academic Work Force." *Research Dialogue*, 23.

Levin, Sharon G., and Paula E. Stephan. 1998. "Gender Differences in the Rewards to Publishing in Academe: Science in the 1970s." *Sex Roles* 38 (11–12): 1049–1064.

Long, J. Scott. 1990. "The Origins of Sex Differences in Science." *Social Forces* 68 (4): 1297–1316.

———. 1992. "Measures of Sex Differences in Scientific Productivity." *Social Forces* 71 (1): 159–178.

Long, J. Scott, Paul D. Allison, and Robert McGinnis. 1993. "Rank Advancement in Academic Careers: Sex Differences and the Effects of Productivity." *American Sociological Review* 58 (5): 703–722.

Mairesse, Jacques and Pezzoni, Michele, 2015. "Does gender affect scientific productivity?". *Revue economique*, *66*(1), pp.65-113.

Sax, Linda J., Linda Serra Hagedorn, Marisol Arredondo, and Frank A. Dicrisi III. 2002. "Faculty Research Productivity: Exploring the Role of Gender and Family-Related Factors." *Research in Higher Education* 43 (4): 423–446.

Stack, Steven. 2004. "Gender, Children and Research Productivity." *Research in Higher Education* 45 (8): 891–920.

Stephan, Paula. 2012. How Economics Shapes Science. Harvard University Press.

Toren, Nina. 1991. "The Nexus between Family and Work Roles of Academic Women in Israel: Reality and Representation." *Sex Roles* 24 (11–12): 651–667.

West, Jevin D., Jennifer Jacquet, Molly M. King, Shelley J. Correll, and Carl T. Bergstrom. 2013. "The Role of Gender in Scholarly Authorship." *PloS One* 8 (7): e66212.

Wooldridge, Jeffrey M. 2006. Introductory Econometrics: A Modern Approach, Third Edition. South-Western, Thomson.

Xie, Yu, and Kimberlee A. Shauman. 1998. "Sex Differences in Research Productivity: New Evidence about an Old Puzzle." *American Sociological Review*, 847–870.

Zuckerman, Harriet. 1987. "Persistence and Change in the Careers of Men and Women Scientists and Engineers. Inn Linda S. Dix, Editor." *Women: Their Underrepresentation and Career Differentials in Science and Engineering*.

Zuckerman, Harriet, Jonathan R. Cole, and John T. Bruer. 1991. *The Outer Circle*. Norton. http://www.law.yale.edu/documents/pdf/uncertain_inheritance.pdf.

Appendix A

Table A.1 compares respondents and non-respondents across their principal characteristics, i.e. birth year, gender, research sub-field, and academic rank. Respondents and non-respondents show minimal evidence of bias on birth year and gender.

Table A.1. Respondents versus non-respondents' characteristics (comparison on the means).

	Respondents (621) Mean variable	Non-respondents (464) Mean variable	P-value
Year of birth	1967.47	1969.72	0.00
Female	0.14	0.23	0.00
Section 2: Physical theories	0.24	0.19	0.02
Section 3: Condensed matter physics (structures and electronic properties)	0.27	0.24	0.45
Section 4: Atoms and Molecules, Optics and Lasers, Hot Plasma Physics	0.25	0.29	0.18
Section 5: Condensed matter physics (organizations and dynamics)	0.24	0.28	0.14
Junior researcher (Second class)	0.06	0.07	0.62
Junior researcher (First class)	0.47	0.46	0.70
Senior researcher (Second class)	0.27	0.28	0.83
Senior researcher (First class)	0.17	0.17	0.87
Senior researcher (Exceptional)	0.03	0.02	0.57