Impact of Subsidies for Researchers on the Gender Scientific Productivity Gap

Evidence from Paraguay

Prepared for the Competitiveness, Technology, and Innovation Division by:

Diego Aboal
Maren Vairo

Institutions for Development Sector

Competitiveness, Technology, and Innovation Division

TECHNICAL NOTE Nº IDB-TN-1303

August 2017

Impact of Subsidies for Researchers on the Gender Scientific Productivity Gap

Evidence from Paraguay

Prepared for the Competitiveness, Technology, and Innovation Division by:

Diego Aboal
Maren Vairo

August 2017

Inter-American Development Bank

Cataloging-in-Publication data provided by the
Inter-American Development Bank

Aboal, Diego.
Impact of subsidies for researchers on the gender scientific productivity gap: evidence from Paraguay / Diego Aboal, Maren Vairo.
p. cm. — (IDB Technical Note ; 1303)
Includes bibliographic references.
1. Gender mainstreaming-Paraguay. 2. Women in science-Paraguay. 3. Research grants-Paraguay. 4. Science and state-Paraguay. I. Vairo, Maren. II. Inter-American Development Bank. Competitiveness, Technology and Innovation Division. III. Title. IV. Series.
IDB-TN-1303

http://www.iadb.org

Abstract*

This paper evaluates the impact of Paraguay's National Researcher Incentive Program (Programa Nacional de Incentivo a los Investigadores, or PRONII) on the gender production gap in academic science. Data has been obtained from the electronic versions of resumes provided by applicants to the program and from bibliographic databases. This paper first quantifies the impact of PRONII's gender science production gap. This is followed by an analysis of whether or not the program's selection process may be gender-biased. Finally, an evaluation is made of the impact of the gender differential of the program. The results indicate that there is a preexistent gender gap in productivity among PRONII researchers. There is no evidence, however, of intended gender discrimination in favor of male researchers at the selection stage of the program. The outcome also demonstrates that the impact of the program is heterogeneous across genders.

JEL classifications: C21, H43, J16, O30, O38

Keywords: economics of gender, economics of science, policy impact evaluation, subsidies

* The authors thank Federico Veneri for the empirical work that led to the estimates reported in Table 2 of this paper. Also appreciated are the comments and suggestions of Jose Miguel Benavente, Matteo Grazzi, Jacques Mairesse, Jocelyn Olivari, Janet Stotsky, and the participants at the Gender Summit 8 North & Latin America 2016 and Jornadas Académicas 2016 of the Faculty of Economic Science and Administration of Uruguay's Universidad de la República. This paper is part of the research project, Science, Technology and Innovation Gender Gaps and their Economic Costs in Latin America and the Caribbean, led by the Competitiveness, Technology, and Innovation Division of the Inter-American Development Bank (IDB) and financed by the IDB Gender and Diversity Fund. The information and opinions presented here are entirely those of the authors, and not the endorsement of the IDB, its Board of Executive Directors, or the countries they represent.

1. Introduction

There is extensive evidence that indicates that there are gender gaps in labor markets. Not only do women receive lower wages than men, they also are underrepresented in various occupations; work fewer hours, in general, than their male counterparts; and have less access to productive inputs (Cuberes and Teignier, 2016).[1] This scenario may play negatively on economic variables such as aggregate income, productivity and economic growth due to the underutilization of female human capital (Klasen and Lamanna, 2009; Cuberes and Teignier, 2016).

Science, technology, and innovation (STI) activities are no different in displaying such gender gaps. Entry into knowledge production activities, historically, has been conditioned on gender (Halbert, 2006). In Latin America, despite the recent narrowing of the education gap reflecting greater female access to higher education, there still remain significant gaps in academic careers. In fact, including the Caribbean, only 6.5 percent of patents filed in 2006−11 were registered by women compared to 69.6 percent by men. Jointly filed patents, on the other hand, rose to 23.9 percent (Morales and Sifontes, 2014). The patenting gap most likely relates to low female participation in the science, technology, engineering, and mathematics disciplines (Castillo, Grazzi, and Tacsir, 2014).

The challenges for women that prevent them from achieving their full STI potential limit the positive externalities that could emerge from greater knowledge production; they also prevent society from reaching its welfare economic optimum. The argument gives ground for policy action aimed at reducing the STI gap between men and women. Unfortunately, few national STI policies reflect this objective (UNCTAD, 2011).

The rationale for the gender gap in STI careers is complex and multidimensional and involves, among others, a wide range of cultural factors[2] and biological considerations (Ceci and Williams, 2009). Reports on the labor market suggest that women, when applying for particular academic positions, are faced with selection (Steinpreis, Anders, and Ritzke, 1999) and wage inequality (Petersen, Saporta and Seidel, 2000, Moss-Racusin et al., 2012).

Evidence suggests that gender-biased recruitment and hiring procedures remain an issue for women (Nielsen, 2015). This is confirmed by a lesser propensity by faculty members to provide positive references for women, challenging the progression of their academic careers (Trix and Pzenka, 2003). Furthermore, women are frequently excluded from networking events and programs (i.e., social gatherings or research and teaching collaboration) impeding their access to information on funding opportunities (UNESCO, 2007).

[1] Also see Klasen and Lamanna (2009), Olivetti and Petrongolo (2008, 2014) and Blau and Kahn (2007, 2013).
[2] See Castillo, Grazzi, and Tacsiret (2014) for a broader analysis of this issue.

Gender gap explanations focus on the factors that discourage the supply of women in STI occupations. Firstly, the prevailing male-dominated culture in STI careers may contribute to an unpleasant environment for women, thus preventing their ingress (Fox, 2005). This is reinforced by discrimination and a cultural stereotype found in higher education, dissuading women from opting for a career in science (Blickenstaff, 2005).

Furthermore, the demands of motherhood, resulting in absence from work during maternity leave, as well as the significant burden of childcare, limits the time required to devote to careers and professional development. In fact, as Muller et al. (2011) indicate, the critical stage of academic careers, such as earning a PhD and Postdoctoral, usually overlaps with the greatest fertility period of women. Nevertheless, Goldin (2014) suggests that STI occupations are relatively motherhood-friendly because they have certain characteristics (i.e., greater work flexibility and independence) whereby hours worked and wages are less directly related, leading to more equality in terms of remuneration. This contrasts with business occupations, where the high salaries are usually over-proportionally linked to long hours of work.

As Bagues et al. (2017) have stated, however, gender discrimination in academia as a cause of women's underrepresentation in science remains a highly controversial issue. For example, a recent meta-analysis of the literature (Ceci and Williams, 2011) does not find evidence of discrimination in journal reviews, grant funding, and recruitment. Moreover, some studies implicitly question the rationale for some gender policies as mechanisms to reduce the gender gap by requiring gender quotas. For instance, with data from Italy and Spain, Bagues et al. (2017) find no evidence that more gender-balanced evaluation committees improve the chances of success for female candidates.

The focus of this research rests on Paraguay's National Researcher Incentive Program (Programa Nacional de Incentivo a los Investigadores, or PRONII). The National Council for Science and Technology (Consejo Nacional de Ciencia y Tecnología, or CONACYT) initiated PRONII in 2011 to boost careers in research by providing researchers a fixed monthly subsidy according to their scientific productivity.[3] CONACYT is responsible for the design and implementation of STI policy in Paraguay. PRONII evaluates researchers with academic criteria. Since gender is not considered in the selection process, in theory it is therefore considered a gender-neutral program.[4]

This paper addresses three questions that relate to PRONII, namely, whether or not (i) there was a pre-existing gender gap in academic productivity among researchers who applied to the program; (ii) the program, at the selection stage, implicitly discriminated

[3] Similar programs exist in other countries in Latin America, (e.g. Argentina, Mexico, and Uruguay). The lessons learned from the study on Paraguay could be useful to other countries in the region.
[4] Part of this empirical approach involves analyzing whether this principle holds in practice.

against women; and (iii) there is evidence of differential impacts of the program on academic productivity (e.g. written research publications, technical outputs, number of theses supervised, level of education of researchers) of men vis-à-vis women.

In its contribution to existing literature, and in its response to the first two questions above, this study provides evidence that there is a gender gap in science in Latin America—until now a relatively unexplored issue; and that there is open debate on gender discrimination in academic careers. In response to the third question, this examination will provide (i) an analysis of the impacts on the gender gap of a particular type of subsidy—seldom examined by the literature—and its pervasiveness in Latin America, based on direct subsidies rather than on researchers; (ii) an evaluation of the program's impact on the productivity of researchers (e.g., technical production, background education, training of researchers, and the extent between bibliography and published articles), based on resume data; and (iii) evidence that relates to a developing country, since most of the available corroboration on research subsidy impact, so far, has been carried out for developed countries.

Section 2 describes science policy in Paraguay and in terms of PRONII; Section 3 elaborates on the data that has been used; Section 4 discusses the methods applied; and Section 5 presents the primary results obtained in terms of quantifying the STI gender gap, assessing the gender-neutrality of the program at the selection stage, and evaluating the gender specific impacts of PRONII. Finally, Section 6 discusses the main findings, with a presentation of some of the conclusions arrived at and the policy implications.

2. Science Policy in Paraguay and Its National Researcher Incentive Program

The Government of Paraguay has undertaken considerable effort to support research and development (R&D) activities in recent years. It has done so by tripling its R&D investment from US$6.5 million in 2005 to US$21.7 million in 2012 (CONACYT, 2012). Given that economic growth has been strong in Paraguay in recent years, however, the increase in spending has generated a very minor rise in gross domestic product per capita, from 0.080 percent to 0.085 percent. In parallel, there has been considerable growth in the number of researchers—from 543 to 1,521 in the same period (CONACYT, 2012). Furthermore, the production of local knowledge has increased significantly, so that publications included in the Science Citation Index and in Elsevier's Scopus (more commonly referred to in Mexico as CONACYT) have risen from 41 and 45 in 2005 to 101 and 135 in 2012, respectively (CONACYT, 2012).

This has eventuated due to the strengthening of CONACYT, in particular, its launch of PRONII in 2011, which was inspired by the respective National Councils of Science and Technology of Mexico and Uruguay. Its objective is to strengthen and expand Paraguay's scientific community, while promoting research and productivity. PRONII classifies researchers according to their scientific and technological production, and provides them with a monetary subsidy according to their ranking in the system.

The program evaluates the applications of researchers and defines the classification to which they are assigned. It requests the candidate to complete a standard resume on its electronic CVPY platform, the latter of which is publicly accessible through the CONACYT website. The following evaluation criteria are applied:

- Production of basic research, applied research and technological outputs of proven quality
- Level of education
- Participation in the development of other researchers' capabilities (mainly through the supervision of undergraduate and graduate theses)
- Participation in the creation and strengthening of institutional capacities for research and experimental development
- Quality of research, judged by taking into account the number of publications accepted in reputable international journals, which are indexed, as well as regional and national journals; patents and original technological outputs; and leadership in the field

The evaluation and categorization of researchers is a three-phase process. Firstly, the applicant is evaluated by a Technical Commission panel (based on scientific discipline) composed of up to five member researchers, either Paraguayan or foreign. The composition of the commission changes after each panel session. Secondly, the commission issues a nonbinding recommendation to the Scientific Committee, which is composed of up to five members in the same discipline, with a two-year tenure. The Scientific Committee then issues its own recommendations. Finally, an Honorary Scientific Committee decides whether or not to admit the applicant into the program, grouping those who are accepted into four tiers.

The tiers into which the researchers are classified are Candidate, Level 1, Level 2 and Level 3. In 2012, researchers at Level 1, Level 2, and Level 3 were offered a monthly subsidy of approximately US$700, US$1,400, and $2,100 dollars, respectively. Researchers at the Candidate level did not receive a monetary incentive in the program's 2011 call for applicants, apart from the prestige of joining the program with the expectation of progress within its objectives.[5]

Given that the annual gross domestic product per capita of Paraguay in 2012 was US$3,860 dollars (i.e. US$322 dollars a month). These subsidies were of some value in the context of Paraguay. Each subsidy is offered for two (Level 1), three (Level 2), and five (Level 3) years, after which time the researchers are re-evaluated. Should their performance have dropped during these years, they are expelled from the program. Such expulsion, therefore, incentivizes the researcher to improve her/his performance, based on the criteria listed previously.

The data that is used in this study derives from PRONII's first call for applications in 2011, relating to 236 researchers. The impact evaluation will consider effects of the policy on various outcomes following two years of program implementation.

Figure 1 shows PRONII's theory of change, whereby the program (subsidies, classification, and periodic evaluation of researchers) should, in the short term, incentivize researchers to dedicate more time to research and development activities in lieu of other activities such as teaching and consultancy; to invest in their own education; and to train other researchers (e.g. supervision of theses). It is important to highlight that in developing countries such as Paraguay, most university professors depend only on a teacher's wage, requiring them to complement their income by way of consultancies or other assignments outside university. Few hold paid research positions. Consequently, incentives such as those provided by PRONII are likely to generate a shift away from research to non-research activities. This displacement effect is doubtless of more significance than the intensity effect that incentives could have on the few who are full-time professors involved in research activities.[6]

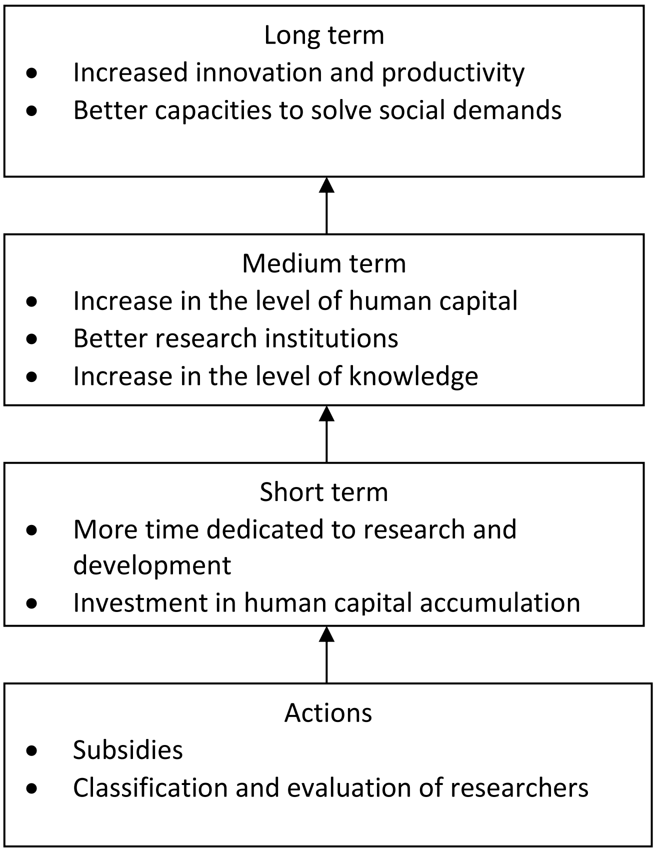
In the medium term, this shift in Paraguay will impact on the number of scientists available and the level of human capital; the quality of Paraguay's research institutions; and the production of knowledge. In the long term, the increased quantity and quality of human capital and knowledge should lead to further innovation and productivity at the firm level and to better capacities to resolve social issues.

On the hypothesis that female researchers are, indeed, discriminated against in universities and research centers, a factor behind the productivity gap, this alternative source of funding (if unbiased) could unleash the potential of female researchers. This will contribute to the closing of the gender science production gap.

[5] Since its second call for applications in 2013, PRONII also offers a subsidy for Candidates of approximately US$400 a month..

[6] From the PRONII database, there are only 70 researchers who claim to have a university job of at least 40 hours a week. This does not imply that they are paid to do research, since some may only be teaching.

Figure 1. The Theory of Change: Paraguay's National Researcher Incentive Program (PRONII)



3. Data and Auxiliary Statistics

The data based on PRONII's electronic applicant resumes relate to those used previously by Aboal and Tacsir (2016) to assess the program's impact. Gender issues, however, were not included in the study. Since the data include information on the program's classified researchers, this study relates to the entire population, rather than a sample of the population.

The performance of researchers is determined by four criteria. These include research production, technical production, level of education (i.e. completion of a Master's (MA) or Doctor of Philosophy (PhD) degree) and contribution to the teaching of new researchers through thesis supervision.

Total research production includes working and conference papers, published and to-be-published research, and books/book chapters. A separate analysis is carried out on scientific journal publications and those indexed with Scopus, as well as their quality in the context of the publishing researcher, by applying the SCImago Journal Rank (SJR) indicator for journals.[7] In terms of technical production, the items include technical work (e.g., advisory activities, consulting, development of regulations and ordinances); technological products (e.g., new varieties of plants, prototypes, software); and the introduction of new processes or techniques (e.g., management processes or analytical techniques).

Table 1 indicates the number of researchers by gender, science discipline, and classification by PRONII during its 2011 call for applications (excluding the rejection category that relates to the 2013 round). Overall, 236 researchers were classified, 109 of which were admitted as Candidates, and 89, 26, and 12 were admitted under Level 1, Level 2, and Level 3, respectively. The gender composition was found to be similar in the rejected group, while the share of women in the Candidate, Level 1, and Level 2 categories was larger than that of men. In Level 3, only one woman was represented compared to 11 men.

Table 1. Paraguay's National Researcher Incentive Program: Number of Researchers by Discipline and Classification

Classification/Discipline	1	2	3	4	Total
Rejected	25	22	12	52	**111**
% Female	40%	18%	58%	60%	**47%**
Candidate	32	7	46	24	**109**
% Female	56%	14%	85%	58%	**66%**
Level 1	36	9	23	21	**89**
% Female	42%	33%	91%	57%	**57%**
Level 2	7	3	11	5	**26**
% Female	71%	33%	73%	20%	**58%**
Level 3	3	3	4	2	**12**
% Female	0%	0%	25%	0%	**8%**
Total	103	44	96	104	**347**
% Female	47%	20%	79%	56%	**55%**

Notes: Science disciplines: (1) Agricultural and Natural, (2) Engineering and Technology, (3) Health, and (4) Social and the Humanities. The Rejected category includes those who were rejected in the 2013 round and those who did not enter the program in 2011, either because they were rejected or because they did not apply.

Table 2 shows the aggregate production of researchers in the PRONII program between the time they obtained their Bachelor degree and 2012. The simple comparison of the mean by gender indicates that women in the sample had lower levels of output than men prior to entering the program, overlooking the fact that other determinants relating to output are not held constant between women and men. In the regression analysis in Section 5, an in-depth test will be made of the differences in outputs across genders.

[7]See Guerrero-Bote and Moya-Anegón (2012) for details about how the SJR indicator is computed for each journal. For each researcher, the mean was computed of the SJR indicator relating to where the journals were published for every year, ranked by the SJR. Note that this indicator relates to the Scopus database to compute the SJR indicator and, therefore, journals have to be indexed in Scopus to appear in the SJR ranking.

Table 2. Paraguay's National Researcher Incentive Program: Accumulated Production prior to Program by Gender

Gender	Statistics	Written research production	Technical production	Papers in scientific journals	Scopus papers	Thesis supervised
Female	mean	33.5	3.6	13.4	2.0	7.8
	sd	34.5	7.9	19.1	3.7	13.9
	min	0	0	0	0	0
	max	160	56	130	23	83
Male	mean	36.6	4.2	13.9	2.6	8.8
	sd	54.7	8.1	30.0	5.1	12.0
	min	0	0	0	0	0
	max	447	49	235	34	53
Total	mean	34.7	3.8	13.6	2.2	8.2
	sd	43.8	8.0	24.1	4.4	13.2
	min	0	0	0	0	0
	max	447	56	235	34	83

Note: mean = average number of items per researcher; sd = standard deviation; min = minimum; max = maximum.

Table 3 reflects the amount of production during two years before and after program implementation (see the definition of variable in the Appendix). When considering the entire sample, the number of researchers whose maximum education level is an MA degree increases over time in the Rejected and Candidate categories, but remains stable in the higher levels, possibly due to the fact that having a PhD degree is a requirement to enter the latter. Conversely, an increase in the number of researchers with a PhD exists is shown in all levels. With regard to other production variables, outputs generally increase over time, observed in terms of contribution to others' human capital through the supervision of theses and in technical and written research production. There is no clear trend in the quality of publications in Level 1 to Level 3, given that the mean SJR indicator decreased, although the number of Scopus publications (another indicator of quality) only moderately increased.

By analyzing production before PRONII (i.e. 2010−11), pretreatment production increases depending on the applicant category in 2011. It is worth pointing out here the exception of Level 3 researchers, who demonstrate a very similar performance to their Level 2 colleagues.

Table 3. Paraguay's National Researcher Incentive Program: Researchers' Performance before and after, by Category and Gender

	Total Sample					*Female Researchers*					*Male Researchers*				
Variables\Category	Rej,	Cand.	Level 1	Level 2	Level 3	Rej.	Cand.	Level 1	Level 2	Level 3	Rej.	Cand.	Level 1	Level 2	Level 3
MA degree (number of people who attained up to a MA degree)															
2011	45	46	29	3	1	25	28	14	2	0	20	18	15	1	1
2013	59	53	31	3	1	30	36	16	2	0	29	17	15	1	1
PhD degree (number of people who attained up to a PhD)															
2011	16	20	27	20	11	9	10	12	10	1	7	10	15	10	10
2013	20	24	30	22	11	12	11	14	12	1	8	13	16	10	10
Theses directed–concluded (mean per annum)															
2010–2011	0.653	0.771	2.219	2.385	1.792	1.038	0.757	2.108	2.500	1.500	0.314	0.797	2.368	2.227	1.818
2012–2013	0.824	0.972	2.865	3.538	1.667	0.904	0.868	2.814	3.867	0.500	0.754	1.176	2.934	3.091	1.773
Theses directed–in process (mean per annum)															
2010–2011	0.054	0.170	0.258	0.365	0.708	0.106	0.132	0.216	0.367	0.000	0.008	0.243	0.316	0.364	0.773
2012–2013	0.509	0.408	1.652	0.942	0.750	0.365	0.306	1.461	0.933	0.500	0.636	0.608	1.908	0.955	0.773
Technical production (mean per annum)															
2010–2011	0.297	0.335	0.438	0.596	1.667	0.250	0.382	0.382	0.700	2.500	0.339	0.243	0.513	0.455	1.591
2012–2013	0.347	0.454	0.798	0.654	3.125	0.433	0.424	0.804	0.667	1.000	0.271	0.514	0.789	0.636	3.318

Written research production (mean per annum)															
2010–2011	0.914	2.413	4.612	5.462	6.167	0.731	2.569	4.941	5.767	5.500	1.076	2.108	4.171	5.045	6.227
2012–2013	1.536	2.193	5.567	7.442	7.792	1.269	1.951	5.039	8.500	8.500	1.771	2.662	6.276	6.000	7.727
Papers in scientific journals (mean per annum)															
2010–2011	0.095	0.784	1.635	2.192	2.125	0.087	0.854	1.931	2.367	5.000	0.102	0.649	1.237	1.955	1.864
2012–2013	0.239	0.881	1.697	2.923	2.000	0.240	0.868	1.941	3.967	7.500	0.237	0.905	1.368	1.500	1.500
Scopus Papers (mean per annum)															
2010–2011	0.041	0.101	0.438	0.865	0.917	0.038	0.104	0.441	1.033	1.000	0.042	0.095	0.434	0.636	0.909
2012–2013	0.032	0.161	0.483	1.212	0.875	0.029	0.132	0.392	1.733	2.000	0.034	0.216	0.605	0.500	0.773
Quality of papers– mean SJR (mean per annum)															
2010–2011	0.022	0.078	0.217	0.482	0.446	0.020	0.098	0.208	0.625	1.024	0.024	0.038	0.229	0.287	0.393
2012–2013	0.029	0.082	0.128	0.373	0.147	0.029	0.048	0.124	0.509	0.558	0.029	0.148	0.134	0.188	0.110

Source: Authors' elaboration, based on CVPY.
Notes: "Rej." and "Cand." stand for Rejected and Candidates, respectively. The Rejected category includes those who were rejected in the 2013 round and did not enter the system in 2011, either because they were rejected or they did not apply that year.

When dividing the sample by gender, the overall pattern remains unchanged in terms of pre-PRONII characteristics, with the exception that female researchers appear to have been more active in scientific publishing in 2010-11 than their male counterparts. Also, the behavior of both groups appears to have evolved similarly over time, in that female and male researchers alike show a general improvement in scientific performance. In the case of candidates, however, the exception is that male ones reflect a larger proportional improvement in most production indicators than women.

4. Methodology

4.1 Gender Gap in Scientific Production

Firstly, an assessment is proposed on whether or not there is a gender gap in scientific production prior to program implementation, requiring a model of the scientific production function, whereby relevant outputs result from certain inputs. Given the characteristics of the output variables that account for the scientific outputs of the researcher (i.e., number of written research and technical papers; papers in scientific journals, including Scopus; and supervised theses), it is appropriate to select a model from the count-data family in order to achieve consistent and efficient estimates.

These dependent variables exhibit another key characteristic, which is that of right-skewed distribution (Figure 2). While most researchers have an accumulated number of publications below 50, there are some rare instances where the number of total publications is over 100,[8] typical for scientific production (Lotka, 1926; Baccini et al., 2014; Kelchtermans et al., 2011).

Therefore, in order to account for the right-skewed distribution of the response variable, a negative binomial regression model is specified (Hausman, Hall, and Griliches, 1984; Long and Freese, 2014).[9] Some variants of this model have been previously used in the literature to assess the production of researchers and the number of patents by universities. For example, Gonzalez-Brambila and Veloso (2007) use a fixed-effect negative binomial model to address determinants of the research output of Mexican researchers. More formally, the negative binomial model is written as:

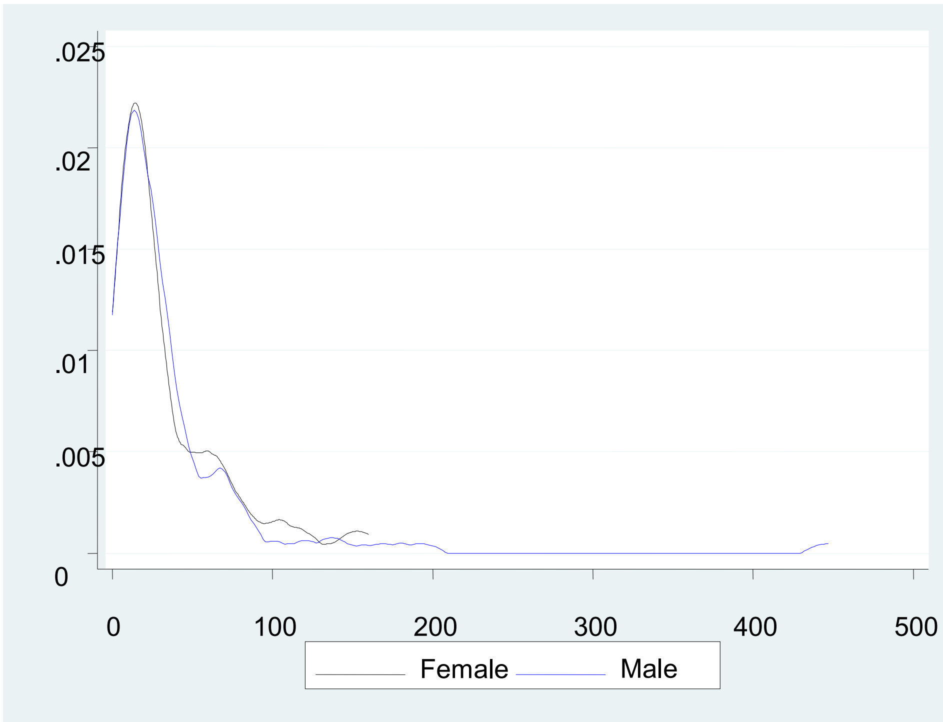
$$\mu_i = E(y_i|x_i) = \exp(x_i\beta + \alpha_i)$$

where μ_i is the expected number of outputs (where the output is y_i) for individual *i* (it can relate to the number of publications, number of theses supervised, etc.), x_i is a vector of explanatory

[8] The figures for the other dependent variables can be found in the Appendix.
[9] This contrasts with the alternative of using a Poisson count model, which assumes a constant variance.

variables, β is a vector of parameters and α_i, with the assumption that $E(e^{\alpha i})=1$, is a fixed effect (unobserved heterogeneity among individuals) that, in practice, allows for right-skewed distribution in the data. The exponential function in this count data model is a key to avoiding the prediction of negative number of outputs.[10]

Figure 2. Paraguay's National Researcher Incentive Program: Density of Written Research Production until 2011, by Gender



This model is applied to five of the selected researcher output variables. The total accumulated production is used, based on the researcher having attained her/his undergraduate degree by 2011 (the pre-program year). As for the selection of the relevant factors that explain scientific production, the work of Gonzalez-Brambila and Veloso (2007) is followed, despite the fact that some of the variables suggested are not included here due to the lack of data. As a result, the explanatory variables in the model relate to gender, age, education (i.e., if the researcher has an MA and/or PhD degree) and the field of research.

Based on the counting of outcomes, it is important to normalize the count by the number of potential years of production. The number of outcomes will not be the same between young

---

[10]The distribution of observations, given the value of x_i and e^{α_i}, $\Pr(y_i|x_i, e^{\alpha_i})$, follows a Poisson. Under the assumption that e^{α_i} is a draw from a Gamma distribution, $\Pr(y|x)$ follows a negative binomial distribution, on which the model is named.

researchers and older researchers (the latter of which has the potential for more production based on the longer opportunity to produce). This is done in the following way,

$$\mu_i * t_i = \exp(x_i\beta + \alpha_i + \ln(t_i)),$$

where t_i is the number of year after the researcher has attained an undergraduate degree (t_i is also known as the exposure time in the literature).

The main purpose of this exercise is to establish whether gender-determined scientific output existed before the program, which could provide evidence that supports the presence of a gender gap prior to PRONII.

4.2 Impact Evaluation and Probability of Participation in the Program

After evaluating the gender gap in the pre-program period, an evaluation is made of the impact of PRONII on productivity and the possibility of heterogeneous treatment effects by gender. In order to identify the causal effect of subsidies on productivity, difference in differences (DiD) will be combined (Angrist and Krueger, 1999) with propensity score matching (PSM) (Rosenbaum and Rubin 1983; Abadie and Imbens, 2006). The combination of DiD and PSM produces credible estimates of causal impact. Heckman et al. (1997), Heckman, Ichimura, and Todd (1998), Smith and Todd (2005) and Chabé-Ferret (2015) find DiD matching performs better than matching for reproducing randomized control trial results. Moreover, as Chabé-Ferret (2014) points out "[c]ombining Difference in Difference with Matching is a non-experimental method of causal inference that reproduces the results of Randomized Controlled Trials (RCTs) very well." For an application of this approach, see Cattaneo et al. (2009).[11]

The impact of a program from the DiD model is,

$$\tau_{DD} = E(Y_2 - Y_1|D = 1) - E(Y_2 - Y_1|D = 0)$$

where Y_2 is the variable of interest for the post program period, Y_1 is the value of the variable for the baseline period, and D is the treatment indicator. Note that DiD estimation controls for time-invariant unobservables, such as ability.

To obtain an unbiased estimation using only a DiD approach, however, assignment to the program must be exogenous (i.e., uncorrelated to unobservable characteristics that may affect the treatment condition and results, and which are not invariant in time). This might not be the case in terms of PRONII, since researchers are placed into each level due to previous performance, so that treated and control units are plausibly different. In this case, an alternative

[11] The combination of approaches is also recommended in the literature on impact evaluation (e.g. Gertler et al., 2011, (Ch. 8) and Bernal and Peña, 2011 (Ch. 6)).

strategy that allows control for time-invariant observable differences between both groups and also for time-invariant unobservables is to combine DiD with PSM.

To implement a PSM strategy, the probability of being treated is estimated given the characteristics of the individual measured before treatment (i.e., age, gender, level of education, field of research, scientific production) with a probit model. Using this estimated probability (or propensity score), each treated individual is matched to its closest (i.e., in terms of the propensity score) in the control group. This allows for a more precise identification of the average treatment effect to be obtained by controlling and matching for observable characteristics. As Table A.3, Table A.4, Table A5, and Table A.6 in the Appendix show, this matching procedure allows for proper control for observable differences between treated and control researchers, thus achieving a balanced sample. The average treatment effect can be expressed as follows,

$$\tau_{\mathrm{ATT}}^{\mathrm{DiD-PSM}} = E[Y_2 - Y_1 | \mathrm{D} = 1, \mathrm{P(X)}] - E[Y_2 - Y_1 | \mathrm{D} = 0, \mathrm{P(X)}].$$

The DiD with PSM estimator $\tau_{\mathrm{ATT}}^{\mathrm{DD-PSM}}$ is the difference in the relevant variable before and after treatment, among the treated and the control group compared on the common support (P(X), using PSM).

The control group for each level of PRONII researchers will be those that were classified in the previous level in the 2011 selection round and, for the Candidates level, controls will be the applicants who were rejected in the 2013 round and did not enter the program in 2011.[12] As an identification strategy, the fact that the subsidy increases with the level in which researchers are categorized will be exploited, as explained in Section 2.

It is worth noting that, given the definition of the control group, PRONII impacts will be estimated on the margin; that is, the impact of belonging to a certain category compared to the possibility of being part of the previous category. For example, a comparison is made on how the higher subsidy received by Level 2 researchers has led to greater productivity by colleagues in Level 1, and similarly between Level 1 and the Candidate level. Thus, the counterfactual scenario is not representative in the absence of the program (except for Candidates compared with nonparticipant researchers), so that failure to identify significant impacts for Level 2 and Level 1 researchers does not imply that PRONII had no effect. It only establishes the incremental effect with respect to the previous level as zero.

This strategy is applied based on an improved control group of individuals in the previous level of the program compared to the pool of individuals that did not enter the program.

[12] Rejected applicants from the 2011 round were not used as a control group for Candidates, given the small size of the group.

Note that matching techniques were used—more specifically, PSM—so as to pinpoint similar individuals in the control group (i.e., computation of a balancing test to establish the reliability of the match).

Additionally, gender-specific effects are obtained by estimating the previous equations separately for male and female researchers. Any gender-differentiated impacts of the program would imply that it has some effect of narrowing (or expanding) the gender gap in scientific production. In summary, DiD-PSM will be used to test the existence of heterogeneous treatment effects by gender and, thus, the program's impact on the gender gap.

As discussed, researchers who were rejected serve as control units to evaluate the impact of being classified a Candidate, ultimately becoming a control group to assess the effect of being a part of Level 1, with Level 1 becoming the control group for Level 2 researchers. Table 2 above provides data on the maximum sample size in each of these exercises. To that extent, while the number of observations for evaluating impacts at the lower PRONII categories is acceptable, the sample size becomes a concern in terms of Level 2 and Level 3. This becomes more relevant when attempting to divide the sample by gender. As a result, an evaluation of the impacts at Level 3 is omitted, where there would be only 38 control and treated units to work with, with only one female Level 3 researcher. Furthermore, while separate results for male and female Level 2 researchers are presented, the reader is urged to interpret the results with caution, since it may be that the sample size is too low to adequately identify the impacts.[13]

To carry out the DiD strategy, information on researcher performance (and other relevant characteristics) is necessary before and after program implementation. Work is then carried out on the periods corresponding to the previous two years (2010 and 2011) and the two years following the program (2012 and 2013), whereby an estimate of the increase in different outputs is made on both, based on the program incentives.

Finally, it is worth noting that by estimating the propensity score (i.e., the probit model for probability of participation in the program) that includes a gender variable, it will establish whether or not gender considerations were used at the program's applicant selection stage—the response to the second question posed in the Introduction.

[13] In particular, the lower the sample size, the lower the power to identify small impacts (i.e., the lower the minimum detectable effect). This therefore becomes a concern whenever expected impacts are small.

5. Results

5.1 Gender and Researcher Outputs

This section estimates negative binomial models as a means to investigate the presence of gender bias in the aggregate output of researchers since they received a BA degree until 2011 (pre-program). Apart from the gender dummy (male=1) that is the main variable of interest in this section, the following control variables are included in the regressions: age dummies, dummies for level of education, dummies for research field and implicitly, years of experience (i.e., number of years after obtaining a BA degree).

Table 4. Paraguay's National Researcher Incentive Program: Estimate of Scientific Production Functions (marginal effects)

Variables	Written research outputs	Technical outputs	Papers in scientific journals	Scopus papers	Theses supervised
Gender (Male=1)	0.418***	0.0670	0.582***	0.358	0.233
	(0.121)	(0.294)	(0.164)	(0.285)	(0.237)
Age 31–40	-0.119	0.497	0.205	0.281	0.368
	(0.221)	(0.550)	(0.311)	(0.554)	(0.441)
Age 41–50	-0.208	0.713	0.245	-0.093	1.029**
	(0.214)	(0.530)	(0.303)	(0.541)	(0.431)
Age 51–60	0.0714	0.856	0.492	-0.215	1.169**
	(0.227)	(0.560)	(0.318)	(0.567)	(0.458)
Age >=61	-0.413	0.842	-0.341	-0.982	1.039**
	(0.272)	(0.673)	(0.376)	(0.676)	(0.507)
MA degree	-0.189	0.384	-0.301	-0.365	-0.274
	(0.143)	(0.338)	(0.189)	(0.347)	(0.277)
PhD degree	0.162	0.132	0.257	0.250	-0.102
	(0.144)	(0.334)	(0.185)	(0.315)	(0.273)
Discipline 2 (Engineering, etc.)	-0.118	0.0108	-1.058***	0.575	-0.145
	(0.199)	(0.457)	(0.273)	(0.435)	(0.354)
Discipline 3 (Health Sciences, etc.)	0.510***	-0.507	0.654***	0.550*	-0.743***
	(0.138)	(0.325)	(0.180)	(0.311)	(0.280)
Discipline 4 (Social Sciences, etc.)	-0.297**	-0.531	-0.830***	-1.398***	-0.674**
	(0.146)	(0.331)	(0.194)	(0.380)	(0.269)
Constant	0.308	-2.348***	-1.048***	-2.417***	-1.416***
	(0.225)	(0.588)	(0.312)	(0.584)	(0.483)
lnalpha	-0.569***	1.011***	-0.110	0.870***	0.566***
	(0.0955)	(0.125)	(0.104)	(0.155)	(0.110)
Observations	**230**	**230**	**230**	**230**	**230**

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Note that the model imposes the restriction that the coefficient of ln (years after Bachelor degree) must be equal to 1. In the regression, the exclusion age is 21−30 and the excluded discipline is Agricultural and Natural Sciences; therefore, the coefficients of age and discipline variables must be interpreted as differences with respect to these excluded categories.

The results in Table 4 demonstrate that women produce fewer written research outputs and papers published in scientific journals than men. Women issue 0.42 and 0.58 publications a year fewer written research outputs and papers in scientific journals, respectively. This is evidence of a gender scientific productivity gap in terms of publications. No gender gap was observed in the number of technical outputs, papers published in Scopus journals or theses supervised (concluded and in process). Moreover, the fact that age is an explanatory factor that is relevant to the supervision of theses (controlling for the number of years since the researcher obtained her/his degree) was established, as well as that of the science discipline appearing relevant when expounding on most of the output variables. To this extent, Agricultural and Natural Sciences (Discipline 1) and Health Sciences (Discipline 3) appear to be the most productive of disciplines. Surprisingly, the researcher's level of education does not appear as relevant input in the production function.

In conclusion, this first analysis points to the presence of a gender scientific production gap prior to program implementation. Going forward, an examination will be made on whether PRONII had any impact on this pre-existing gap.

- **Probability of program participation**

This section presents the results from the first stage of the impact evaluation, whereby a determination is made of a propensity score that models the probability of categorization at various PRONII levels, according to the characteristics of researchers. This will help establish whether or not there is gender-based discrimination at the selection stage of the program, by analyzing whether gender is a significant factor that explains the propensity score (or the probability to enter the program). Table 5 presents the results from the probit probability of entry into the various categories of the program. The data is then applied to match treatment units to their "nearest neighbors" in the control group.[14]

The exercise is carried out separately by taking into account the probability of Candidate, Level 1 and Level 2 researcher categories. For Level 2, an evaluation is made of two and three years following PRONII implementation. The first exercise resembles that of the remaining categories, since it compares Level 2 researchers from the 2011 intake with their Level 1 counterparts. In the second case, however, the control group comprises those researchers who entered Level 1 in the 2011 intake and remained at that level after the 2013 intake. A double exercise was done for this category since Level 2 researchers are requested to

[14]The probit estimate was carried out, dividing the sample by gender. Results are qualitatively similar to those reported in Table 5.

re-apply to the program every three years to preserve their classification. Therefore, it is reasonable to assume that their electronic resumes were updated by 2014 (contrary to those of other categories with no requirement for updating them post-2013). As stated previously, the results relating to Level 3 researchers are unreported, given the minimal observations.

The first 4 columns in Table 5 show the determinants from including every dimension that, *a priori,* could affect program participation: gender, age, education, various research and technical production indicators, and science discipline. All variables included are measured prior to program implementation (year 2011). This first specification allows for an assessment to determine which factors relate to the researcher categories. One worthy result from this exercise is that gender does not significantly determine PRONII entry potential at any level. This is not surprising, since the program does not have gender-specific criteria for selection, confirming the neutrality of the program in treatment allocation.

In addition, the determinants for entry vary according to PRONII's various classifications. In this regard, the probability of being classified as a Candidate is positively affected by age, earning a PhD, and previously written research production and scientific journal publication, and it negatively relates to being a researcher in the disciplines of Engineering and Technology or Social Sciences and Humanities d(i.e., compared to the discipline of Agricultural and Natural Sciences, which is omitted). This is consistent with entry criteria for this level, such as to have participated in research activities through publishing and having completed or taken part in a graduate program. Selection bias on particular activities or age, however, is not a feature in the program's screening process.

Entry to Level 1, however, is positively related to the number of supervised theses (concluded), written research production, and Scopus journal publications, while it is not for Health Sciences.[15] PRONII requires Level 1 researchers to hold an MA or PhD degree or equivalent scientific production, demonstrating ability to carry out original and independent research. Scientific publishing rather than graduate education appears to significantly influence the chance of entry into Level 1. Based on the second specification in Column 6 of the table (selection is based on the estimation of propensity scores), having an MA degree negatively affects Level 1 entry potential. As Aboal and Tacsir (2016) point out, this result—implying that an MA degree increases the probability of Candidate in lieu of Level 1 entry, may demonstrate that PRONII evaluators considered the applicant lacks sufficient education to be accepted into Level 1.

[15] In a second specification (Column 6 of Table 3), where other nonsignificant variables are removed, age appears as another factor that positively affects the probability of being classified as Level 1, while researchers whose maximum education level is an MA degree appear to be less likely to enter this level.

Additionally, acceptance into PRONII's Level 2 is positively determined by an MA or PhD degree, by quality of publications, and by being in the discipline of Health Sciences. This is consistent with Level 2 researchers being required to have a PhD degree or equivalent scientific production. Additional requirements for acceptance into this level are a "strong track record of work, particularly in the five years prior to each round, having developed its own line of research with sustained production of original knowledge and activities aimed at capacity building for research" (Aboal and Tacsir, 2016). Since the quality of publications may equate to a compliance indicator in terms of requirements, it appears reasonable to apply the probit model. Contribution to capacity building for research through thesis supervision, however, does not appear to affect the chance of entry into this level. Lastly, there is bias with regard to science discipline compared to the selection criteria.

To avoid over-specification of the probit model, only those variables that were significant at the 10-percent level are included in the estimation of propensity score. The results of the final specification are presented in Column 5, Column 6, Column 7, and Column 8.[16]

[16] To achieve a balanced matching procedure, a PhD degree was removed as a control in the Candidate estimation (this variable became nonsignificant when removing other nonsignificant variables in Column 1), and age and MA degree were added to the Level 1 estimation (both variables are nonsignificant in Column 2, but become significant in the new specification in Column 6).

Table 5. Paraguay's National Researcher Incentive Program: Probability of Program Participation by Category (marginal effects)

Variables\Level	(1) Candidate	(2) Level 1	(3) Level 2	(4) Level 2b	(5) Candidate	(6) Level 1	(7) Level 2	(8) Level 2b
Female	0.162	-0.043	0.022	0.009				
	(0.100)	(0.104)	(0.075)	(0.093)				
Age	0.116***	0.004	0.047	0.052	0.127***	0.022***		
	(0.045)	(0.043)	(0.034)	(0.041)	(0.039)	(0.005)		
Age^2	-0.0014**	0.0003	-0.0004	-0.0004	-0.0015***			
	(0.0005)	(0.0005)	(0.0004)	(0.0004)	(0.0005)			
MA degree	0.175*	-0.160	0.378*	0.490**		-0.170*	0.273*	0.441**
	(0.104)	(0.119)	(0.206)	(0.231)		(0.092)	(0.182)	(0.216)
PhD degree	0.295***	-0.0148	0.603***	0.699***			0.491***	0.646***
	(0.114)	(0.138)	(0.123)	(0.131)			(0.118)	(0.129)
Theses supervised (concluded)	-0.028	0.074***	0.004	0.014		0.072***		
	(0.029)	(0.027)	(0.013)	(0.015)		(0.026)		
Theses supervised (in process)	0.087	0.002	0.034	-0.041				
	(0.101)	(0.107)	(0.066)	(0.080)				
Technical production	0.027	-0.053	0.017	0.007				
	(0.057)	(0.051)	(0.043)	(0.048)				
Written research production	0.051**	0.072***	-0.007	0.011	0.045**	0.064***		
	(0.024)	(0.027)	(0.011)	(0.017)	(0.021)	(0.020)		
Papers in scientific journals	0.648***	-0.014	0.002	-0.029	0.619***			
	(0.129)	(0.058)	(0.022)	(0.029)	(0.116)			
Papers in Scopus	0.239	0.281*	-0.005	0.116		0.419***		
	(0.291)	(0.162)	(0.052)	(0.075)		(0.125)		
Mean SJR	-0.080	0.231	0.223**	0.214**			0.160**	0.224**
	(0.388)	(0.196)	(0.093)	(0.104)			(0.081)	(0.092)
Engineering and Technology	-0.289**	0.081	0.091	0.078	-0.315***			
	(0.138)	(0.177)	(0.168)	(0.181)	(0.115)			
Health Sciences	0.055	-0.476***	0.262	0.345*		-0.417***	0.232*	0.364**

	(0.139)	(0.110)	(0.166)	(0.201)		(0.094)	(0.130)	(0.166)
Social Sciences and Humanities	-0.375***	-0.182	-0.016	0.098	-0.350***			
	(0.116)	(0.116)	(0.101)	(0.167)	(0.089)			
Observations	205	186	113	100	210	186	114	101

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The 0 category relates to rejected applicants in 2013 in Column 1 and Column 5; Candidates in 2011 in Column 2 and Column 6; Level 1 researchers in 2011 in Column 3 and Column 7; Level 1 researchers in 2011 and 2013 in Columns 4 and Column 8. The discipline of Agricultural and Natural Sciences is omitted. Mean of variables in 2010−11, except in the case of MA and PhD degree dummies where degree attainment by 2011 is applied.

- **Gender-specific impact of Paraguay's National Researcher Incentive Program**

This section reviews the results from the PRONII impact evaluation. Table 6 indicates the average treatment effects of PRONII on the various researcher performance indicators, distinguishing them by classification and gender. As explained above, a five-neighbor PSM matching approach is applied, combined with DiD techniques.[17] This implies the adoption of a traditional matching approach, while defining the dependent variable as the change in the result of interest, pre- and post-implementation.

At the Candidate level, only a positive and significant program impact exists in the case of scientific journal publication. Candidate researchers reflect 0.45 more publications per annum compared to researchers rejected by the program. This effect also appears to be justified in that the program has no impact on women at this level compared to male counterparts. Thus, the program's contribution to the gender scientific production gap (except in the case of scientific journal publication, where the gender gap broadens) has no significance on the level of Candidate, at which there is no monetary incentive.

The results vary somewhat at Level 1, where there is a positive impact of entry compared to that of Candidate in terms of theses supervision in process (0.79 more) and technical and written production (0.75 additional technical output and 1.1 additional written research output per annum). This is considered reasonable since the indicators are used to assess the permanence of researchers in Level 1 and their promotion to Level 2. Furthermore, since Level 1 entrants receive a monetary subsidy, one could argue that this is representative of a more significant incentive to increase production and rely less on the financial needs that may distract researchers from academic activities compared to those at the Candidate level. Also present are impacts on technical and written research outputs, due to the positive effect on female written research output, justifying that subsidies at this level appear to close the gap in this dimension. Lastly, a negative impact of Level 1 classification on the quality of publications by male entrants is determined, possibly because of the high amount of publications required by the program, thus encouraging researchers to opt for quantity over quality. There are no positive outcomes on the number of publications by male researchers at this stage.

As stated previously, caution is necessary when interpreting the Level 2 results due to the small sample size, especially in terms of gender-specific impacts. In any case, the outcome shows a negative effect on Level 2 females, when compared with their Level 1 female colleagues in terms of MA degrees. This is observed following two and three years of program implementation. A positive impact on the technical production of Level 2 male

[17] Results from using 1 neighbor matching are shown in Table A.2 in the Appendix. In the five-neighbor PSM each treated individual is compared with a weighted average (of the relevant variable) of the five closest neighbors (i.e., the five most similar individuals in the control group).

researchers (0.85 more outputs per annum) also takes place after three years, thus broadening the gender scientific production gap. Conversely, the negative effect established on theses supervision in process in men after two years (i.e., two fewer theses per annum) and the positive effect found for women after three years (i.e. one more thesis per annum) would also contribute to closing the gap. Similar conclusions may be drawn from the positive impact on Level 2 women with regard to publication quality.

In sum, the results suggest that the program has little impact at the Candidate level in terms of aggregate scientific production and of affecting the gender gap associated with such output. More consistent evidence of program impact on Level 1 is necessary, in particular, regarding the narrowing of the gender gap, which suggests that monetary incentives are important to alter the gender distribution of scientific production. Finally, results at Level 2 are more ambiguous, possibly due to robustness issues with respect to the small size of the sample.

Finally, some tests were performed to validate the identification strategy. Balancing tests were carried out on the means of the control variables selected for the estimation of the propensity scores between treated and controls. Balance of the sample is a key assumption of PSM strategy in terms of guaranteeing comparability between treatment and control groups (in the common support). Results of these tests were satisfactory and are reported in Tables A.3, Table A-4, Table A-5, and Table A-6 of the Appendix.

Table 6. Paraguay's National Researcher Incentive Program: Impact on Researcher Performance (Overall and by Gender)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	**MA**	**PhD**	**Theses directed (concluded)**	**Theses directed (in process)**	**Technical production**	**Written research production**	**Papers in scientific journals**	**Papers in Scopus**	**Quality of papers (mean SJR)**
Candidates									
ATT_All	-0.003	0.009	0.326	0.006	-0.245	0.348	0.452***	0.071	0.001
Obs_All	174	174	174	174	174	174	174	174	174
ATT_Women	0.089	-0.036	0.909	0.025	-0.410	-0.446	-0.131	0.000	-0.081
Obs_Women	77	77	77	77	77	77	77	77	77
ATT_Men	-0.169	0.069	-0.004	-0.165	0.250	1.096	0.473*	0.115	0.075
Obs_Men	82	82	82	82	82	82	82	82	82
Level I									
ATT_All	-0.070	0.038	0.752	0.786**	0.752**	1.103*	0.244	0.052	-0.021
Obs_All	162	162	162	162	162	162	162	162	162
ATT_Women	-0.057	0.057	0.683	0.380	0.971**	1.031*	0.454	0.206	0.099
Obs_Women	101	101	101	101	101	101	101	101	101
ATT_Men	-0.024	0.012	0.106	0.176	0.365	1.412	0.341	-0.053	-0.333**
Obs_Men	50	50	50	50	50	50	50	50	50

Level II-2 years									
ATT_All	0.001	0.068	0.705	-0.563	-0.186	-0.925	0.017	0.016	0.159
Obs_All	114	114	114	114	114	114	114	114	114
ATT_Women	-0.119*	0.133	0.056	-0.196	-0.152	1.330	1.013	0.865	0.279**
Obs_Women	65	65	65	65	65	65	65	65	65
ATT_Men	0.008	-0.008	-0.436	-2.056**	0.689	-0.388	-0.782	-0.057	0.074
Obs_Men	45	45	45	45	45	45	45	45	45
Level II-3 years									
ATT_All	-0.070*	0.104	-0.206	0.211	-0.242	1.606	0.630	0.102	0.222**
Obs_All	93	93	93	93	93	93	93	93	93
ATT_Women	-0.129**	0.143	0.105	0.993**	-0.043	1.045	0.905	0.109	0.150
Obs_Women	58	58	58	58	58	58	58	58	58
ATT_Men	0.010	-0.010	-0.115	-0.899	0.848*	0.269	-0.585	-0.178	0.011
Obs_Men	39	39	39	39	39	39	39	39	39

Source: Authors' elaboration based on CVPY.
* Difference in difference estimates with five-neighbor propensity score matching.
Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Dependent variable in columns 1 and 2 is the change in the accomplishment of a MA or PhD degree from 2011 to 2013. For Columns 3−9, the dependent variable is the difference between the mean production in 2010−11 and the mean production in 2012−13. The only exception is in the case of the evaluation of PRONII in three years (bottom panel), where a change is applied to the MA and PhD degree from 2011 to 2014, and changes in mean production from 2010−11 to 2012−14. Control variables used in each panel are the ones that result from Columns 5−8 in Table 5. ATT is the average treatment effect on treated. Obs is the number of observations.

Appendix Figure A.5, Figure A-6, and Figure A.7 allow for an analysis of compliance with the parallel trends assumption imposed by the DiD model, by illustrating the separate evolution of dependent variables in the treatment and control groups prior to 2011. As the figures indicate, some variables do not show a parallel pre-PRONII evolution. This poses the question of whether the time-invariant evolution of unobserved heterogeneity assumption is fulfilled, on which DiD strategy relies. The matching strategy proposed here, however, aims to address issues that might result from this by controlling for observed characteristics and matching each treated individual to a comparable one in the control group.

6. Conclusions and Policy Implications

This paper investigates three key issues that relate to a gender scientific production gap relating to researchers in Paraguay. Firstly, an analysis was made on whether there was a relevant gender gap prior to PRONII. Secondly, an investigation was carried out on whether or not the program did, in fact, discriminate at the selection stage against female researchers. Finally, an evaluation was conducted of the differential impacts of the program on productivity across genders.

The findings indicate that there was, indeed, a pre-existing gender gap in scientific production among PRONII researchers. Other aspects being equal, female researchers have a smaller number of written research outputs and papers that have been published in scientific journals. This implies that there is room for policy action to narrow the gap in academic achievement between male and female researchers with similar merits. This also calls for an examination of the impact of current policy actions.

There is no evidence of discrimination against female researchers at the selection stage of the program, confirming the notion that the program is gender-neutral in terms of evaluating female and male applicants equally. The program, however, does exert a certain degree of nongender-based discrimination based on age and science discipline, despite the fact that it is not reflected in the program's selection criteria.

Finally, results exhibit the program impact as not being homogeneous across genders or levels. In particular, PRONII actually contributes to closing the gender gap by increasing female researcher production relative to male researchers in terms of technical and written papers, as well as the quality of publications at Level 1 of PRONII. This also applies to the supervision of theses and quality of publications for Level 2 researchers. The program shows no impact on or contribution to the widening of the gap at other levels.

In sum, results of any PRONII impact at this stage on the gender gap are mixed, in that certain levels of the classification and/or certain scientific production variables appear to be associated with a narrowing of the gap, based on program implementation, compared to the gap widening or remaining the same for other levels/outputs. This ambiguity may be caused by idiosyncratic elements that lead to specific patterns in each case, which were not considered in the PRONII design, and are thus affected by program implementation. For example, previous evidence of the impact of policies specifically aimed at favoring women's STI careers shows that these initiatives have not been effective in dismantling institutional and cultural factors that lead to gender discrimination in STI activities (Muller et al., 2011). Moreover, given that PRONII is a new program and one that has no gender-specific objective, it appears reasonable that no impacts were found in some cases.

It is worth noting that since PRONII is a recently implemented program, the impacts identified are those associated with the very short term. Long-term impacts require further program maturity in order to be empirically identifiable. Additionally, there are issues with the sample size that hamper impact identification, which merit recognition, especially with regard to the higher program levels.[18] A further exercise, targeting a more extensive time period, could lead to more robust results and account for longer-term policy impact.

Some policy implications, however, can be observed from results obtained so far. Firstly, even though the program appears to have some positive effect on the gender gap, it is evident that additional efforts should be made to adequately tackle the issue more in depth. In addition, Level 2 results are not as conclusive in terms of PRONII's impact on the gender gap. It may be the case that at the more advanced stages in their academic career, women face greater barriers,[19] so that a more vigorous effort should be made to balance the scale. Finally, it may be an issue that calls for the design of more gender-specific measures (beyond the scope of PRONII, perhaps at the university level) to improve female scientific productivity and address the gender inequality in STI activities.

The research findings discussed previously have shed light on the magnitude of the gender gap in scientific researcher productivity in Paraguay; it should enable the introduction of policies to close it. In addition, the evidence of PRONII's gender-differentiated impact on scientific productivity indicates that gender-neutral programs can have nonneutral impacts, calling for additional research to better understand the mechanisms. This is a key to implementing, ex-post (on results), neutral incentives.

[18] More precisely, there are few resources available to identify relatively small impacts.
[19] This is asserted by a body of literature that implies that female academic development becomes harder the more advanced the career stage (Castillo, Grazzi, and Tacsir., 2014; Blickenstaff, 2005).

References

Angrist, J. D. and A.B. Krueger (1999). “Empirical Strategies in Labor Economics.” *Handbook of Labor Economics* 3: 1277−366

Abadie, A. and G. Imbens (2006). Large Sample Properties of Matching Estimators for Average Treatment Effects.” *Econometrica* 74 (1): 235–67.

Aboal, D. and E. Tacsir (2016). “The Impact of Subsidies to Researchers on Researcher’s Productivity: Evidence from a Developing Country.” UNU-Merit Working Paper #2016-019.

Baccini, A., L. Barabesi, M. Cioni, and C. Pisani (2014). “Crossing the Hurdle: th+e Determinants of Individual Scientific Performance.” *Scientometrics* 101(3): 2035–62.

Bagues, M., M. Sylos-Labini, and N. Zinovyeva (2017). Does the Gender Composition of Scientific Committees Matter?” *American Economic Review*, 107(4): 1207−1238.Bernal, R. and X. Peña (2011). “Guía Práctica para la Evaluación de Impacto.”, Facultad de Economía, Universidad de los Andes, Bogota: Ediciones Uniandes, p 336.

Blau, F.D. and L.M. Kahn (2013). “Female Labor Supply: Why Is the US Falling Behind?” NBER Working Paper 18702.

Blau, F.D. and L.M. Kahn (2007). “The Gender Pay Gap: Have Women Gone as Far as They Can?” *Academy of Management Perspectives* 21(1): 7−23.

Blickenstaff, J. (2005). “Women and Science Careers: Leaky Pipeline or Gender Filter?” *Gender and Education*, 17(4): 369−86.

Castillo, R., M. Grazzi, and E. Tacsir (2014). “Women in Science and Technology: What Does the Literature Say?” Technical Note No.IDB-TN-637. Washington, DC: Inter-American Development Bank.

Cattaneo, M., S. Galiani, P. Gertler, S. Martinez, and E. Titiunik, E. (2009). “Housing, Health, and Happiness.” *American Economic Journal*: *Economic Policy* 1(1): 75−105.

Ceci, S. and W. Williams, W. (2011). “Understanding Current Causes of Women's Underrepresentation in Science.” *Proceedings of the National Academy of Sciences* 108(8): 3157−3162.

Chabé-Ferret, S. (2015). “Analysis of the Bias of Matching and Difference-in-Difference under Alternative Earnings and Selection Processes.” *Journal of Econometrics* 185(1): 110–23.

Chabé-Ferret, S. (2014). “Bias of Causal Effect Estimators Using Pre-Policy Outcomes.” Unpublished.

CONACYT (2012). “Estadísticas e Indicadores de Ciencia y Tecnología de Paraguay – 2012”. Available at www.conacyt.gov.py.

Cuberes, D. and M. Teignier (2016). “Aggregate Effects of Gender Gaps in the Labor Market: A Quantitative Estimate.” *Journal of Human Capital* 10(1): 1−32.

Fox, F. M. (2005). “Gender, Family Characteristics, and Publication Productivity among Scientists.” *Social Studies of Science*, 35(1): 131−50.

Gertler, P., S. Martinez, P. Premand, L. Rawlings, and C. Vermeersch (2011). *Impact Evaluation in Practice*. 2nd Edition. Washington, DC: World Bank.

Goldin, C. (2014). "A Grand Gender Convergence: Its Last Chapter." *The American Economic Review* 104(4), 1091−119.

Gonzalez-Brambila, C. and F.M. Veloso (2007). "The Determinants of Research Output and Impact: A Study of Mexican Researchers." *Research Policy* 36(7): 1035−1051.

Guerrero-Bote, V. and F. Moya-Anegón (2012). "A Further Step Forward in Measuring Journals' Scientific Prestige: The SJR2 Indicator." *Journal of Informetrics* 6(4): 674–688.

Halbert, D. (2006). "Feminist Interpretations of Intellectual Property." *American University Journal of Gender, Social Policy & the Law* 14(3): 431−60.

Hausman, J., B. Hall, and Z. Griliches (1984). "Econometric Models for Count Data with an Application to the Patents-R&D Relationship." *Econometrica* 52(4): 909−38.

Heckman, J.J., H. Ichimura, J.A. Smith, and P.E. Todd (1998). "Characterizing Selection Bias Using Experimental Data." *Econometrica* 66, 1017–99.

Heckman, J.J., H. Ichimura, and P.E. Todd (1997). "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *Review of Economic Studies,* 64(4): 605–654.

Kelchtermans, K. and R. Veugelers (2011). "The Great Divide in Scientific Productivity: Why the Average Scientist Does Not Exist." *Industrial and Corporate Change* 20: 295−336.

Klasen, S. and F. Lamanna (2009). "The Impact of Gender Inequality in Education and Employment on Economic Growth: New Evidence for a Panel of Countries." *Feminist Economics* 15(3): 91−132.

Long, J. S. and J. Freese, J. (2014). *Regression Models for Categorical Dependent Variables using Stata*. Stata Press.

Lotka, A.J. (1926). "The Frequency Distribution of Scientific Productivity." *Journal of the Washington Academy of Science* 16: 317−323.

Morales, R. and D. Sifontes (2014). "Desigualdad de Género en Ciencia y Tecnología: Un Estudio para América Latina." *Observatorio Laboral Revista Venezolana*, 7(13): 95−110.

Moss-Racusin, C. A., J.F. Dovidio, V.L. Brescoll, M.J. Graham, and Handelsman (2012). "Science Faculty's Subtle Gender Biases Favor Male Students." *Proceedings of the National Academy of Sciences* 109(41): 16474−79.

Müller, J., C. Castano, A. Gonzalez, and R. Palmén. (2011). "Policy towards Gender Equality in Science and Research." *Brussels Economic Review* 54 (2/3): 295−316.

Nielsen, M. (2015). "Limits to Meritocracy? Gender in Academic Recruitment and Promotion Processes." *Science and Public Policy* 43(3): 386−99.

Olivetti, C, and B. Petrongolo (2008)."Unequal Pay or Unequal Employment? A Cross-country Analysis of Gender Gaps". *Journal of Labor Economics* 26(4): 621−54.

Olivetti, C. and B. Petrongolo (2014). “Gender Gaps across Countries and Skills: Demand, Supply and the Industry Structure”. *Review of Economic Dynamics* 17(4): 842−59.

Petersen, T., I. Saporta, and M-D.L. Seidel (2000). “Offering a Job: Meritocracy and Social Networks.” *American Journal of Sociology* 106: 763−816.

Rosenbaum, P. R., and D.B. Rubin (1983). “The Central Role of the Propensity Score in Observational Studies for Causal Effects.” *Biometrika* 70(1): 41−55.

Smith, J.A. and P.E. Todd (2005). “Does Matching Overcome LaLonde’s Critique of Non-experimental Estimators?” *Journal of Econometrics* 125(1–2): 305–353.

Steinpreis, R., K.A. Anders, and D. Ritzke, D. (1999). “The Impact of Gender on the Review of the Curricula Vitae of Job Applicants and Tenure Candidates: A National Empirical Study.” S*ex Roles*, 41 (7/8): 509−28.

Trix, F. and C. Psenka (2003). “Exploring the Color of Glass: Letters of Recommendation for Female and Male Medical Faculty.” *Discourse and Society* 14(2): 191−220.

UNCTAD (2011). “Applying a Gender Lens to Science, Technology and Innovation.” Unpublished.

UNESCO. (2007). *Science, Technology and Gender: An International Report*. Paris: United Nations Educational, Scientific and Cultural Organization.

Appendix

Table A.1. Definition of Indicators of Researcher Performance Applied in Empirical Exercises

Variable	Definition
1. Researcher performance indicators	
MA degree	Indicator variable =1 if the researcher's maximum education level is a MA degree
PhD degree	Indicator variable, =1 if the researcher's maximum education level is a doctorate degree
Supervised theses concluded	Number of concluded supervision of undergraduate and graduate theses per annum
Supervised theses directed in process	Number of ongoing supervision of undergraduate and graduate theses per annum
Technical production	Number of yearly technical outputs (including technical work, technological products, and new processes or techniques)
Written research production	Number of yearly written research publications (including papers in scientific and nonscientific publications, works published in events, publication of books and book chapters, and working papers)
Papers in scientific journals	Number of yearly papers published or accepted for publication in scientific journals
Papers in Scopus	Number of yearly papers published in Scopus journals
Mean SJR	Mean SJR of the journals in which the researcher has published in during the year
2. Area of science	
Discipline 1	Indicator variable =1 if the researcher's specialization is in Agricultural and Natural Sciences
Discipline 2	Indicator variable =1 if the researcher's specialization is in Engineering and Technology
Discipline 3	Indicator variable =1 if the researcher's specialization is in Health Sciences
Discipline 4	Indicator variable =1 if the researcher's specialization is in Social Sciences and the Humanities

Note: The pre-treatment values of performance indicators are applied when estimating the probit models for propensity scores. In such cases, the educational level attained by 2011 in the case of MA and PhD degrees is used, as are the mean values for 2010 and 2011 for the remaining variables. Conversely, the variables used for difference-of-difference impact evaluation are defined as the change in variables before and after Paraguay's National Researcher Incentive Program. As a result, the change in MA and PhD attainment between 2013 and 2011 is used, as is the change in mean production in 2012 and 2013 versus mean production in 2010 and 2011 for the remaining variables.

Figure A.1. Paraguay's National Researcher Incentive Program: Density of Accumulated Technical Production until 2011, by Gender

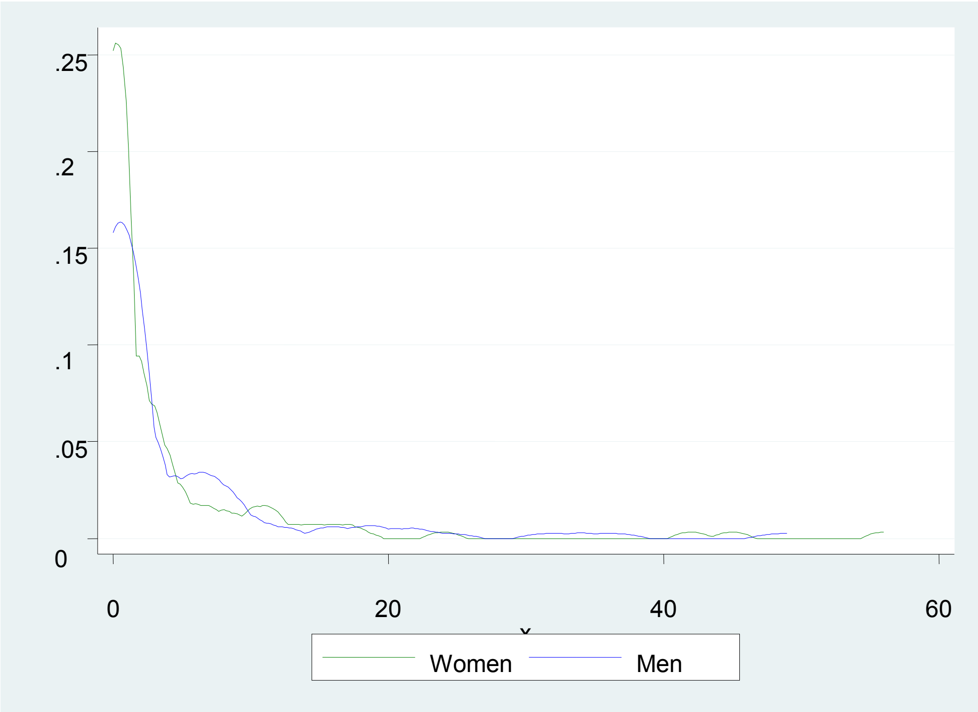


Figure A.2. Paraguay's National Researcher Incentive Program: Density of Accumulated Number of Publications in Scientific Journals until 2011, by Gender

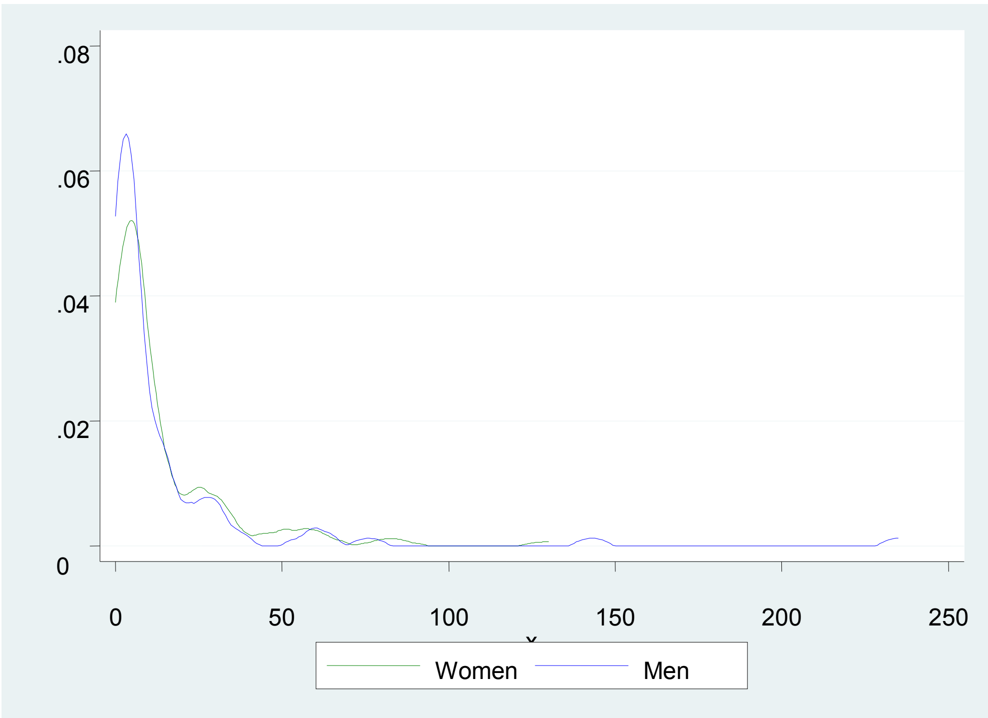


Figure A.3. Paraguay's National Researcher Incentive Program: Density of Accumulated Theses Supervised until 2011, by Gender

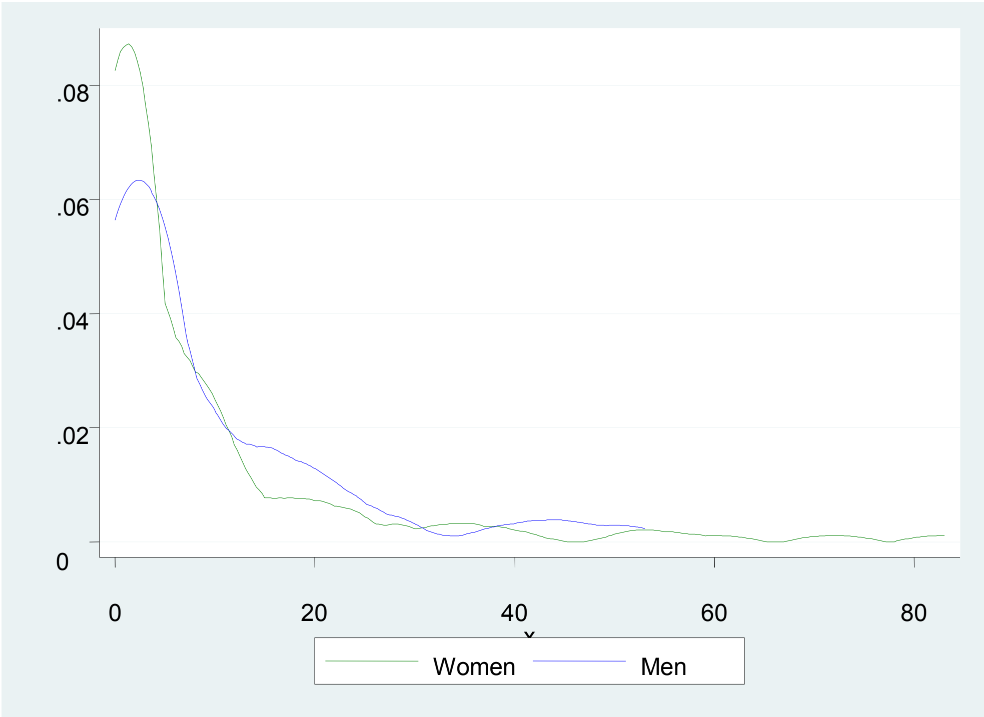


Figure A.4. Paraguay's National Researcher Incentive Program: Density of Accumulated Number of Publications in Scopus Journals until 2011, by Gender

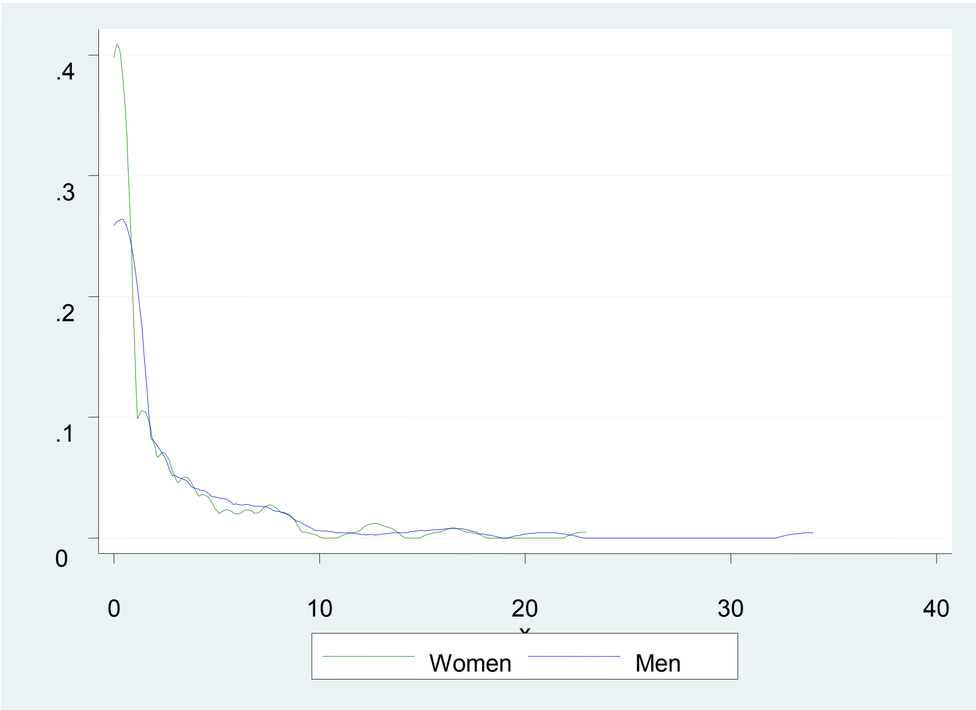


Table A.2. Paraguay’s National Researcher Incentive Program: Impact with One-Neighbor Matching

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	Master	PhD	Theses directed (conclud.)	Theses directed (in progress)	Technical production	Written research production	Papers in scientific journals	Scopus papers	Quality of papers (mean SJR)
Candidates									
ATT_All	0.066	-0.022	0.243	0.147	-0.408	0.29	0.460**	0.074**	0.001
Obs_All	174	174	174	174	174	174	174	174	174
ATT_Wom.	0.093	-0.019	1.296**	0.111	-0.769*	0.102	0.102	0.000	-0.084
Obs_Wom.	77	77	77	77	77	77	77	77	77
ATT_Men	-0.192	0.038	0.019	-0.096	0.154	1.423	0.481	0.115*	0.080
Obs_Men	82	82	82	82	82	82	82	82	82
Level I									
ATT_All	-0.175**	0.032	0.810	0.762	0.556	1.167	0.452	0.183*	0.168*
Obs_All	162	162	162	162	162	162	162	162	162
ATT_Wom.	-0.086	0.057	0.271	0.657	0.771	0.757	0.629*	0.286*	0.179
Obs_Wom.	101	101	101	101	101	101	101	101	101
ATT_Men	0.000	0.000	-0.382	0.176	0.176	0.618	0.000	-0.059	-0.341*
Obs_Men	50	50	50	50	50	50	50	50	50
Level II-2 years									
ATT_All	0.010	0.066	0.741	-0.594	-0.006	-0.090	0.071	0.060	0.102
Obs_All	114	114	114	114	114	114	114	114	114
ATT_Wom.	0.013	0.130	0.364	-0.608	-0.002	3.572**	2.281***	0.913	0.157
Obs_Wom.	65	65	65	65	65	65	65	65	65
ATT_Men	0.011	-0.011	-0.233	-2.011**	0.438	-0.801	-0.844	-0.065	0.157
Obs_Men	45	45	45	45	45	45	45	45	45
Level II-3 years									
ATT_All	-0.046	0.102	-0.387	0.312	-0.132	1.515	0.392	0.078	0.451***
Obs_All	93	93	93	93	93	93	93	93	93
ATT_Wom.	0.000	0.143	-0.041	0.953**	0.152	-0.450	0.292	0.000	0.157
Obs_Wom.	58	58	58	58	58	58	58	58	58
ATT_Men	0.014	-0.014	-0.127	-0.604	0.627	0.591	-0.625	-0.190	0.005
Obs_Men	39	39	39	39	39	39	39	39	39

Source: Authors’ elaboration, based on CVPY.
Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.Dependant variable in Columns 1 and 2 is the change in the accomplishment of an MA or PhD degree from 2011 to 2013. For Columns 3−9, the dependent variable is the difference between the mean production in 2010−11 and the mean production in 2012−13. The only exception is in the case of the evaluation of Paraguay’s National Researcher Incentive Program in three years (bottom panel), where a change in MA and PhD degree from 2011 to 2014 is used, as in mean production from 2010−11 to 2012−14. Control variables used in each panel are the ones that result from Columns 5−8 in Table 5. ATT is the average treatment effect on treated. Obs is the number of observations.

Table A.3. Paraguay's National Researcher Incentive Program: Balancing Tests for Matching Variables in Five Neighbors Propensity Score Matching Candidates

		Mean				t-test	
Variable	Unmatched/ Matched	Treated	Control	% bias	% reduct bias	t	p>t
	U	39.667	39.302	3.5		0.25	0.803
Age	M	38.574	40.526	-18.8	-435.4	-1.42	0.159
	U	1652.2	1679.9	-3.2		-0.22	0.823
Age^2	M	1568.1	1689.7	-13.8	-338.8	-1.07	0.286
Written research	U	2.4949	0.90094	74.6		5.32	0
production	M	1.75	1.2517	23.3	68.7	1.8	0.074
Papers in scientific	U	0.82323	0.08962	132.9		9.65	0
journals	M	0.41176	0.37328	7	94.8	0.55	0.586
	U	0.07071	0.20755	-40.1		-2.85	0.005
Field 2	M	0.10294	0.17598	-21.4	46.6	-1.23	0.222
	U	0.23232	0.46226	-49.5		-3.53	0.001
Field 4	M	0.26471	0.34118	-16.5	66.7	-0.97	0.336

Source: Authors' elaboration, based on CVPY.

Table A.4. Paraguay's National Researcher Incentive Program: Balancing Tests for Matching Variables in Five Neighbors- Propensity Score Matching Level 1 Researchers

		Mean				t-test	
Variable	Unmatched /Matched	Treated	Control	% bias	% reduct bias	t	p>t
	U	2.2701	0.73737	55.6		3.88	0
Age	M	1.5317	1.5429	-0.4	99.3	-0.03	0.979
	U	4.6897	2.4949	60.3		4.2	0
Master's	M	3.2222	2.8794	9.4	84.4	0.88	0.382
Theses directed	U	0.44828	0.11111	51.8		3.62	0
concluded	M	0.22222	0.14921	11.2	78.3	1.06	0.291
Written research	U	0.25287	0.40404	-32.4		-2.2	0.029
production	M	0.22222	0.16825	11.6	64.3	0.76	0.449
Papers in	U	46.322	39.667	71.9		4.91	0
Scopus	M	45.476	44.273	13	81.9	0.77	0.441
	U	0.32184	0.43434	-23.2		-1.58	0.116
Field 3	M	0.39683	0.33968	11.8	49.2	0.66	0.51

Source: Authors' elaboration, based on CVPY.

Table A.5. Paraguay's National Researcher Incentive Program: Balancing Tests for Matching Variables in Five Neighbors-Propensity Score Matching Level 2 Researchers (Two-Year Evaluation)

		Mean				t-test	
Variable	Unmatched/ Matched	Treated	Control	% bias	% reduct bias	t	p>t
	U	0.11538	0.32955	-52.8		-2.16	0.033
MA degree	M	0.11538	0.23077	-28.4	46.1	-1.09	0.281
	U	0.76923	0.30682	103.4		4.54	0
PhD degree	M	0.76923	0.73077	8.6	91.7	0.31	0.755
	U	0.48217	0.21919	56.1		2.7	0.008
Mean SJR	M	0.48217	0.51485	-7	87.6	-0.21	0.833
	U	0.42308	0.25	36.8		1.72	0.089
Discipline 3	M	0.42308	0.26154	34.3	6.7	1.22	0.228

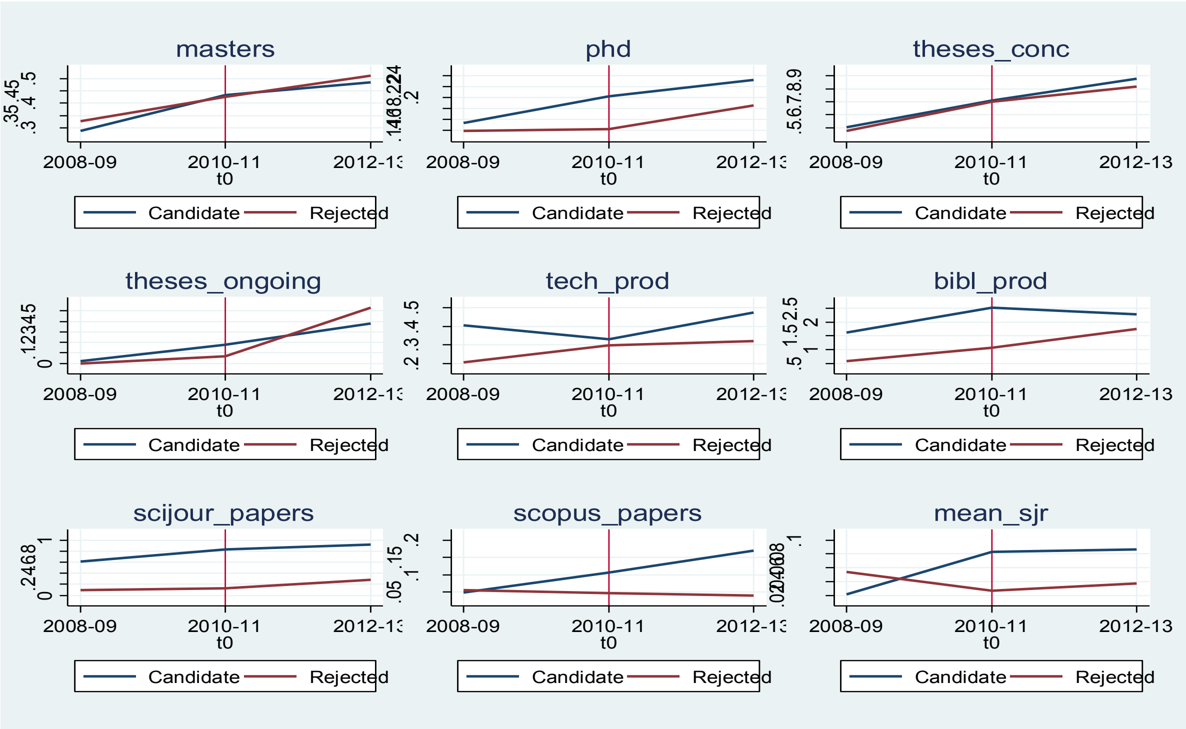
Source: Authors' elaboration, based on CVPY.

Table A.6. Paraguay's National Researcher Incentive Program: Balancing Tests for Control Variables in Five Neighbors Propensity Score Matching Level 2 Researchers (Three-Year Evaluation)

		Mean				t-test	
Variable	Unmatched /Matched	Treated	Control	% bias	% reduct bias	t	p>t
	U	0.11538	0.33333	-53.5		-2.17	0.033
MA degree	M	0.16667	0.33333	-40.9	23.5	-1.14	0.261
	U	0.76923	0.29333	107.1		4.63	0
PhD degree	M	0.66667	0.55556	25	76.7	0.67	0.508
	U	0.48217	0.21017	58.4		2.74	0.007
Mean SJR	M	0.38165	0.39342	-2.5	95.7	-0.06	0.952
	U	0.42308	0.25333	36		1.64	0.105
Discipline 3	M	0.22222	0.17778	9.4	73.8	0.32	0.748

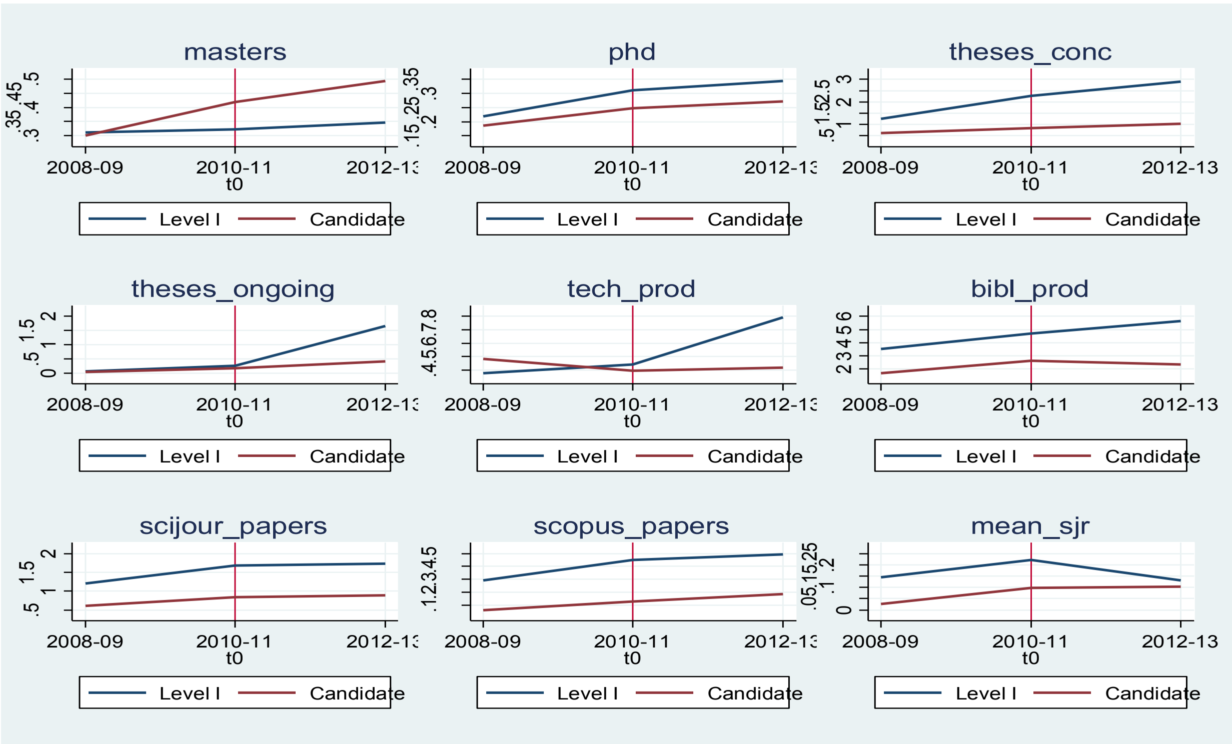
Source: Authors' elaboration, based on CVPY.

Figure A.5. Paraguay's National Researcher Incentive Program: Parallel Trend Analysis: Candidates and Applicants Rejected



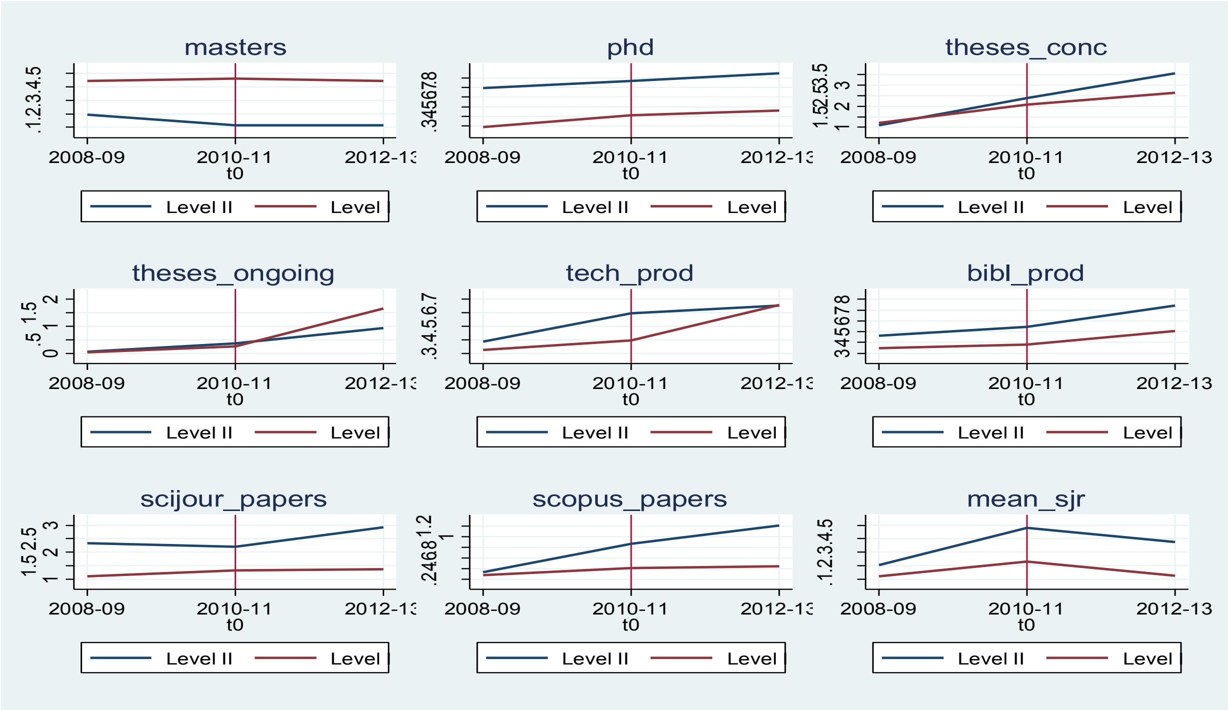
Source: Authors' elaboration, based on CVPY.

Figure A.6. Paraguay's National Researcher Incentive Program: Parallel Trend Analysis: Level 1 Researchers and Candidates



Source: Authors' elaboration, based on CVPY.

Figure A.7. Paraguay's National Researcher Incentive Program: Parallel Trend Analysis: Level 1 and Level 2 Researchers



*Source***:** Authors' elaboration, based on CVPY.