# Accounting for gender research performance differences in ranking universities

*Giovanni Abramo* and Ciriaco Andrea D'Angelo*

*The literature on the theme of gender differences in research performance indicates a gap in favour of men over women. Beyond the understanding of the factors that could be at the basis of this phenomenon, it is worthwhile understanding if it would be appropriate to conduct the evaluation per population in a manner distinguished by gender. In fact, if there is some factor that structurally determines a penalization of performance by women researchers compared to men, then the comparative evaluation in the performance of organizations' that do not take gender into account will lead to an advantage for those that employ more men, under parity in the capacities of their staff. In this work we measure the differences in the performance and the rank of research institutions as observed when gender is taken into account compared to when it is ignored. The study population consists of all Italian universities and the performance measured in the hard sciences for the period 2006–10.*

**Keywords:** Bibliometrics, gender differences, ranking universities, research productivity.

THE representation of women in research varies across countries, institutions and disciplines. However, the fact of their underrepresentation is undeniable. Only four (Portugal, Estonia, Slovak Republic and Iceland) of the 28 OECD nations whose data are available show representation of women greater than 40%, with a maximum of 46% (ref. 1). This phenomenon is underlain by a mix of different factors, with different weights across countries: educational emancipation has come later for women, with consequent lesser numbers of potential candidates for academic positions; lesser interest among women for research activity; the scientific production of women tends to be lower than that for men, perhaps due to the social roles of women as wives and mothers, or from causes of gender discrimination; gender discrimination can also occur in recruitment processes. One way to reduce the underrepresentation of women in research is to control for the factors exogenous to scientific merit in all processes of comparative performance evaluation at the individual and institutional levels. By doing so one can avoid incorrect conclusions and choices that are harmful to women and to those institutions with greater female representation in their research staff. In this regard it is important to note that the so-called 'productivity gap' in favour of men is a documented fact. The lower research performance of women has been established in several studies from diverse countries and disciplines[2–6], although it is decreasing over time[7–12], and it is more visible in the early stages of career[13] and among top scientists[14,15]. Looking at productivity as indicated by patenting, women faculty members contribute about 40% compared to men[16].

There is equally substantial literature investigating the possible causes of the productivity gap, particularly the issues of the environmental and personal factors that can influence the performance of a researcher, beyond the personal merit of the individual[17]. Rossiter[18] indicates the particular case of the so called 'Matilda effect', which occurs when female scientists are not recognized in the bylines of the publications resulting from joint research. A separate concern is that in the career recruitment stages, the percentage of women applicants who are successful in selection procedures is generally low[19]. In the subsequent stages of entry to the academic environment, females generally evaluate their mentors as less satisfactory than do their male colleagues[20]. However, there is also no doubt that the changing personal conditions that the researchers experience over time also affect their productivity. In the late postdoctoral and early faculty years, many qualified women scientists stop applying for NIH grants[21]. During their careers, women present lower productivity in the intermediate levels of seniority[22]. In this phase, marriage and school-going children have negative effects on research productivity[23–25]. It has been demonstrated that research collaborations have a positive correlation

Giovanni Abramo is in the Laboratory for Studies of Research and Technology Transfer, Institute for System Analysis and Computer Science, National Research Council of Italy, Via del Taurini 19, 00185 Rome, Italy; Ciriaco Andrea D'Angelo is in the Department of Engineering and Management, University of Rome 'Tor Vergata', Via del Politecnico 1, 00133 Rome, Italy
*For correspondence. (e-mail: giovanni.abramo@uniroma2.it)

with scientific performance[26] and particularly international collaborations[27], but also that women register less international collaborations than men[28], possibly for reasons of women avoiding longer stays away from their families. Women tend to have more restricted collaboration networks[29–31], particularly in the first few years of their career[32,33], which limits their access to the resources and assets necessary for their research activity. Duch *et al*.[34] observe that academic research institutions tend not to support women with adequate financial resources, particularly in the hard sciences. According to Ceci and Williams[35], differential gender outcomes result exclusively from differences in resources.

However, the aim of the present study is not to investigate further into issues regarding whether or to what extent there is gender discrimination in the research sphere, or into the objective limitations on women due to their social roles. Rather we wish to determine if the separation of the measurement of research performance by gender produces detectably different results from measurement without gender distinction. In contexts where the potential of discrimination by gender is recognized, or where the family roles of women can condition the time, energy or personal concentration devoted to research, the conduct of comparative evaluation without distinction by gender would inevitably penalize the research organizations with a higher representation of women. The results of the analysis are thus of interest in all processes of comparative evaluation of institutions, such as for example national research assessment exercises, especially where these are intended for the efficient allocation of the available resources. The policy makers can then decide whether the extent of rank differences suggests gender distinction while conducting research evaluation exercises for institutions. We show evidences for Italian universities. We proceed by preparing two rankings of the research productivity of the universities for the period 2006–10: one obtained through the aggregation of individual performances with distinction by gender and the other where the aggregation is conducted without distinction, in order to examine the extent of the differences.

The next section presents the context, methodology adopted for finding the productivity of the universities' and the dataset used for the analyses. A section setting out the main results of the work follows. The final section presents the conclusions and offers several policy indications.

## Context, method and data

The Italian Ministry of Education, Universities and Research (MIUR) recognizes a total of 96 universities as having the authority to issue legally recognized degrees. Among these, 29 are small, private, special-focus universities, of which 13 offer only e-learning and 67 are public and generally multi-disciplinary universities scattered throughout Italy. In keeping with the Humboldtian model, there are no 'teaching-only' universities in Italy, as all professors are required to carry out both teaching and research. In the Italian university system, all professors are classified in one and only one field (named the scientific disciplinary sector (SDS), 370 in all). Fields are grouped into disciplines (named university disciplinary areas (UDAs), 14 in all). The overall staff system has over 58,000 professors, of which 95% is employed in public universities. Faculty members consist a majority of men, although the data since 1998 indicate a trend towards increasing presence of women. This is also shown in the representation of female assistant professors (45.3%), which is now much higher than that of full professors (20.7%). Female professors are in the majority only in the UDAs of ancient history, philology, literature, art history (55.2%) and biology (51.6%). The UDAs with the lowest presence of women are physics (19.6%), and industrial and information engineering (15.1%).

### *Measuring research productivity of universities*

To measure research productivity of universities, we adopt the approach described in Abramo and D'Angelo[36]. We begin by measuring research productivity at the individual level and then aggregate the individual measures for the evaluation of organizations. At the individual level, we adopt an indicator named fractional scientific strength (FSS) embedding both the number of publications produced, their standardized impact and the number of co-authors of each. The average yearly productivity of an individual, over a period of time, accounting for the cost of labour, is

$$\mathrm{FSS} = \frac{1}{w} * \frac{1}{t} \sum_{i=1}^{N} \frac{c_i}{\bar{c}} f_i, \qquad (1)$$

where $w$ is the average yearly wage of the professor; $t$ the number of years of work by the professor in the period under observation; $N$ the number of publications by the professor in the period under observation; $c_i$ the citations received by publication $i$; $\bar{c}$ the average of the distribution of citations received for all cited publications indexed in the same year and subject category of publication $i$; $f_i$ the fractional contribution of the researcher to publication $i$. The fractional contribution equals the inverse of the number of authors, except in life sciences, where the various contributions are weighted according to the order of the names in the byline[37].

Research productivity of universities is obtained by aggregating individual research productivity, according to the following formula

$$\mathrm{FSS_U} = \frac{1}{\mathrm{RS}} \sum_{i=1}^{\mathrm{RS}} \frac{\mathrm{FSS}_{R_i}}{\overline{\mathrm{FSS}_{R_l}}}, \qquad (2)$$

where RS represents the research staff of the university in the observed period, $\mathrm{FSS}_{R_i}$ the productivity of researcher *i* in the university and $\overline{\mathrm{FSS}_{R_i}}$ is the average productivity of all productive researchers in the same SDS of researcher *i*.

The choice of average productivity of all productive researchers as the optimum scaling factor to reduce distortions when comparing performance of heterogeneous research institutions is based on the results of a study by Abramo *et al*.[38]. This scaling factor could be calculated separately for the two subpopulations of gender, since Abramo *et al*.[39] have demonstrated that the relative distributions are significantly different. We ask what are the effects on the value of $\mathrm{FSS}_{\mathrm{U}}$ (and thus on the ranking lists) from the choice of whether or not to apply a scaling factor differentiated by gender. We will attempt to provide an answer to this question in a later section. However, first we will illustrate the dataset used in the analyses.

### *Data and sources*

Data on the research staff of each university, such as years of employment in the observed period, academic rank and their SDS classification, are extracted from the database on Italian university personnel maintained by the Ministry for Universities and Research (MIUR), Italy. Unfortunately, information on 'leave of absence' is not available and cannot be accounted for in the calculation of yearly productivity, to the disadvantage of women on maternity leave in the period of observation.

The bibliometric dataset for the analysis draws on the Observatory of Public Research (ORP), a database developed and maintained by the authors and derived under license from the *Web of Science* (*WoS*). Beginning from the raw data of Italian publications indexed in the *WoS*, and applying a complex algorithm for disambiguation of the true identity of the authors and their institutional affiliations (for details see D'Angelo *et al*.), each publication is attributed to the university professor who authored the same, with a harmonic average of precision and recall (*F*-measure) equal to 96 (error of 4%). Beginning from these data, we can calculate FSS for each Italian professor. We limit the field of observation to the hard sciences, i.e. the nine UDAs in Table 1. For the *WoS*-indexed publications to serve as a more robust proxy of overall output of a researcher, the field of observation is further limited to those SDSs (188 in all) where at least 50% of member scientists had at least one publication in the period 2006–10. For the purpose of the study and to ensure significant representation of both genders in each field, we then limit the analysis to those SDSs (99 in all) with at least 30 individuals of each gender. Finally, for a robust comparison of university ranks by UDA, we exclude those universities with less than 10 professors in the UDA. Table 1 shows the final dataset.

## Analysis and discussion

As an example of the preparation of the ranking lists, Table 2 shows the Italian universities active in chemistry ranked for productivity. As noted, for reasons of significance the construction of the ranking list considers only those universities (41 in chemistry) with at least 10 professors in the UDA. For each university, the table shows: the absolute values of productivity calculated as in eq. (2) with and without gender distinction, the relative positions in the ranking and the differences that emerge in terms of value and sign. Eight of the universities listed maintain the same position in the ranking; however, 33 show changes. Sixteen of the 33 move up in the rankings taking account of gender and among these, two (13 and 18) gain four positions, while another two (5 and 21) gain three places. On the opposite side we find three universities that lose four positions (3, 15 and 22) and two that lose three places (14 and 19).

Considering that the value of FSS for the UDA is given by the average of the individual values rescaled to the average of their SDS, we can apply the *t*-test for paired

**Table 1.** Dataset of the analysis: number of fields (SDSs), universities and professors in each university disciplinary area (UDA) under study

| UDA | No. of SDSs | Professors | | Universities | |
|---|---|---|---|---|---|
| | | Total | Female | Total | With at least 10 professors |
| Mathematics and computer science | 8 | 3,297 | 1,105 (33.5%) | 65 | 50 |
| Physics | 4 | 2,161 | 390 (18.0%) | 61 | 43 |
| Chemistry | 9 | 3,199 | 1,212 (37.9%) | 59 | 41 |
| Earth sciences | 4 | 534 | 176 (33.0%) | 41 | 22 |
| Biology | 19 | 5,338 | 2,591 (48.5%) | 66 | 50 |
| Medicine | 29 | 9,426 | 2,805 (29.8%) | 60 | 42 |
| Agricultural and veterinary sciences | 17 | 2,163 | 755 (34.9%) | 43 | 27 |
| Civil engineering | 3 | 828 | 130 (15.7%) | 49 | 31 |
| Industrial and information engineering | 6 | 2,051 | 298 (14.5%) | 64 | 42 |
| Total | 99 | 28,997 | 9,462 (32.6%) | 79 | 64 |

**Table 2.** Productivity rankings of Italian universities in chemistry (2006–10) with ($FSS_U^2$) and without ($FSS_U^1$) gender distinction

| | $FSS_U^1$ | | $FSS_U^2$ | | | | | $FSS_U^1$ | | $FSS_U^2$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| University | Abs. value | Rank | Abs. value | Rank | Sign | Rank diff. (abs) | University | Abs. value | Rank | Abs. value | Rank | Sign | Rank diff. (abs) |
| 1 | 1.724 | 1 | 1.737 | 1 | = | 0 | 22*** | 0.805 | 22 | 0.750 | 26 | – | 4 |
| 2 | 1.307 | 2 | 1.320 | 3 | – | 1 | 23** | 0.800 | 23 | 0.842 | 21 | + | 2 |
| 3** | 1.289 | 3 | 1.168 | 7 | – | 4 | 24 | 0.771 | 24 | 0.801 | 23 | + | 1 |
| 4 | 1.283 | 4 | 1.287 | 5 | – | 1 | 25 | 0.769 | 25 | 0.794 | 24 | + | 1 |
| 5 | 1.256 | 5 | 1.325 | 2 | + | 3 | 26* | 0.745 | 26 | 0.783 | 25 | + | 1 |
| 6* | 1.196 | 6 | 1.293 | 4 | + | 2 | 27 | 0.718 | 27 | 0.732 | 27 | = | 0 |
| 7 | 1.190 | 7 | 1.160 | 8 | – | 1 | 28 | 0.712 | 28 | 0.715 | 28 | = | 0 |
| 8 | 1.174 | 8 | 1.180 | 6 | + | 2 | 29 | 0.701 | 29 | 0.667 | 30 | – | 1 |
| 9 | 1.121 | 9 | 1.107 | 10 | – | 1 | 30 | 0.680 | 30 | 0.660 | 31 | – | 1 |
| 10 | 1.091 | 10 | 1.042 | 12 | – | 2 | 31 | 0.661 | 31 | 0.671 | 29 | + | 2 |
| 11 | 1.076 | 11 | 1.039 | 13 | – | 2 | 32* | 0.642 | 32 | 0.619 | 33 | – | 1 |
| 12 | 1.056 | 12 | 1.094 | 11 | + | 1 | 33 | 0.634 | 33 | 0.622 | 32 | + | 1 |
| 13 | 1.024 | 13 | 1.112 | 9 | + | 4 | 34 | 0.627 | 34 | 0.614 | 34 | = | 0 |
| 14* | 0.988 | 14 | 0.918 | 17 | – | 3 | 35*** | 0.613 | 35 | 0.584 | 35 | = | 0 |
| 15 | 0.985 | 15 | 0.877 | 19 | – | 4 | 36 | 0.565 | 36 | 0.527 | 37 | – | 1 |
| 16 | 0.984 | 16 | 0.938 | 16 | = | 0 | 37 | 0.540 | 37 | 0.496 | 38 | – | 1 |
| 17 | 0.976 | 17 | 0.956 | 15 | + | 2 | 38 | 0.540 | 38 | 0.553 | 36 | + | 2 |
| 18 | 0.953 | 18 | 0.991 | 14 | + | 4 | 39 | 0.489 | 39 | 0.431 | 40 | – | 1 |
| 19** | 0.928 | 19 | 0.835 | 22 | – | 3 | 40 | 0.459 | 40 | 0.487 | 39 | + | 1 |
| 20 | 0.882 | 20 | 0.863 | 20 | = | 0 | 41 | 0.255 | 41 | 0.261 | 41 | = | 0 |
| 21 | 0.862 | 21 | 0.890 | 18 | + | 3 | | | | | | | |

No. of observations: 41; sum of differences: 62; max of difference: 4; mean of differences: 1.561.
***Paired *t*-test: *P*-value < 0.01; ***P*-value < 0.05; **P*-value < 0.10.

**Table 3.** Descriptive statistics of rank differences between $FSS_U$ calculated with and without gender distinction

| UDA | No. of universities‡ | Shifting in rank | Max shift | Average shift | Spearman $\rho$ |
|---|---|---|---|---|---|
| Agricultural and veterinary sciences | 27 (2) | 12 (44.4%) | 2 | 0.519 | 0.995*** |
| Biology | 50 (7) | 42 (84.0%) | 7 | 1.960 | 0.984*** |
| Chemistry | 41 (11) | 33 (80.5%) | 4 | 1.561 | 0.986*** |
| Civil engineering | 31 (8) | 18 (58.1%) | 3 | 0.710 | 0.994*** |
| Earth sciences | 22 (5) | 15 (68.2%) | 5 | 1.091 | 0.968*** |
| Industrial and information engineering | 42 (8) | 21 (50.0%) | 4 | 0.857 | 0.994*** |
| Mathematics and computer sciences | 50 (10) | 39 (78.0%) | 9 | 2.200 | 0.978*** |
| Medicine | 42 (5) | 20 (47.6%) | 5 | 1.000 | 0.990*** |
| Physics | 43 (11) | 19 (44.2%) | 3 | 0.605 | 0.997*** |

Significance level: ****P* < 0.01; ***P* < 0.05; **P* < 0.10.
‡The number of universities with significant differences (*P*-value < 0.10) in $FSS_U$ with and without gender distinction is given within brackets.

samples in each university to evaluate the significance of any differences between the values of productivity. In formula, the *t*-test applied is

$$t\text{-test}_{\text{paired}} = \frac{FSS_U^1 - FSS_U^2}{s/\sqrt{n}}, \quad (3)$$

where $FSS_U^1$ is the university's productivity of the university without gender distinction; $FSS_U^2$ the productivity of the university with gender distinction; *s* the standard deviation of the difference between $FSS_U^1$ and $FSS_U^2$, and *n* is the number of researchers present in the UDA.

Table 2 presents the results of this test, with the asterisks in columns 1 and 8 indicating the universities that show significant differences in the two rankings. In the area of chemistry, 11 out of 41 universities show significant differences in productivity when distinguished for gender, with these differences having only partial effect on the variation in rank, as demonstrated by the high value of Spearman correlation (0.986) and the low average number of variations (1.561).

The above analyses are repeated for each of the remaining eight UDAs, with Table 3 showing the results. In comparing the rankings, it emerges that biology has the highest percentage of universities that change at least one position (84), while physics has the lowest percentage (44%). We may recall that physics is also the UDA with the least representation of women. Mathematics and

**Table 4.** Example of calculation of the indicator $R'$

| University | Rank#1 | Rank#2 | Abs. diff. \|Rank#1 – Rank#2\| | $Rank\#1_{inverted}$ | Abs. diff. $\|Rank\#1 - Rank\#1_{inverted}\|$ |
|---|---|---|---|---|---|
| ID1 | 1 | 2 | 1 | 5 | 4 |
| ID2 | 2 | 3 | 1 | 4 | 2 |
| ID3 | 3 | 4 | 1 | 3 | 0 |
| ID4 | 4 | 1 | 3 | 2 | 2 |
| ID5 | 5 | 5 | 0 | 1 | 4 |
| | | Total difference | 6 | | 12 => $R'$ = 6/12 = 50% |

computer science shows the cases of the highest individual shifts, with one university gaining nine positions in the productivity ranking by gender and another losing nine positions. This UDA also registers the highest average shift per university (2.2 positions). This UDA is in contrast with agricultural and veterinary sciences, where the maximum shift in position is 2 and the average shift is 0.519. The limited variations in rank are accompanied by Spearman correlation values that are consistently above 0.96, and all highly significant. Applying the *t*-test we observe another notable result: the highest number of universities with significant differences in productivity with and without gender distinction is seen in both chemistry (11 out of 41) and physics (11 out of 43). However, in the latter UDA the differences have minimal impact on the rankings.

As an alternative to the Spearman correlation coefficient, which measures the intensity and sign of the interdependence between the two ranking lists, we also consider another indicator ($R'$) that measures the potential of the rank differences. This is given by the sum of the absolute differences in rank registered in an area and the maximum sum of rank differences with reference to the theoretical situation of perfect inversion of the rankings. The indicator assumes a value of zero in case of identical ranking lists with and without gender distinction and 100 in the case of perfect inversion of the ranking lists. Thus

$$R' = \frac{\sum_{i=1}^{n} \left| d_{\text{rank}_i} \right|}{\max}, \quad (4)$$

where $d_{\text{rank}_i}$ is the difference in rank registered for university $i$, under the two evaluation methods, and $n$ is the number of universities active in the UDA.

$$\max = \begin{cases} \dfrac{n^2}{2}, \text{ for } n \text{ even number,} \\ (n-1) * \left( \dfrac{n-1}{2} + 1 \right), \text{ for } n \text{ odd number.} \end{cases}$$

Table 4 presents the example of a fictitious case of five universities. Comparing two hypothetical rankings (Rank#1 and Rank#2), we obtain the sum of the absolute differences in rank as 6. In the case of perfect inversion of the rankings ($Rank\#1_{inverted}$) the sum of the differences would be 12, from which we obtain an $R'$ value (ratio of 6 to 12) equal to 50%.

Figure 1 presents the $R'$ values for each UDA: in no case does the indicator exceed 10%. In physics the differences in rank are the lowest in comparison with the other UDAs ($R' = 2.8\%$). This UDA, and agricultural and veterinary sciences ($R' = 3.8\%$), form a first cluster with quite low values for shifts in ranking. A second group of UDAs with higher values of $R'$ but still lower than 5% is composed of industrial and information engineering, civil engineering and medicine. A third cluster, with values between 7% and 8% is composed of biology and chemistry. A final cluster, composed of mathematics and computer science, and earth sciences, shows values over 8%.

## Conclusion

Higher education institutions represent an important pillar of national research and innovation systems. Thus the policy agendas of many countries now place high priority on strengthening such institutions. An expression of this policy is the increasing diffusion of national research assessment exercises. Such assessments serve different goals, including a strategic one related to efficient resource allocation and stimulation of performance improvement. It follows that they must be conceived and executed with maximum methodological rigour.

Stimulated by a now well-consolidated literature that indicates the presence of a 'productivity gap' in favour of male researchers, in this study we verify the impact of gender aspects on the outcomes of bibliometric assessments carried out for research institutions.

In theory, in fact, if there is some factor that structurally determines a penalization of performance for women researchers compared to men, then a comparative evaluation of organizational performance that does not take gender into account will lead to an advantage for those that employ more men, under parity of capacities among the research staff.

The analyses conducted concerned Italian universities active in the hard sciences for the period 2006–10. Being

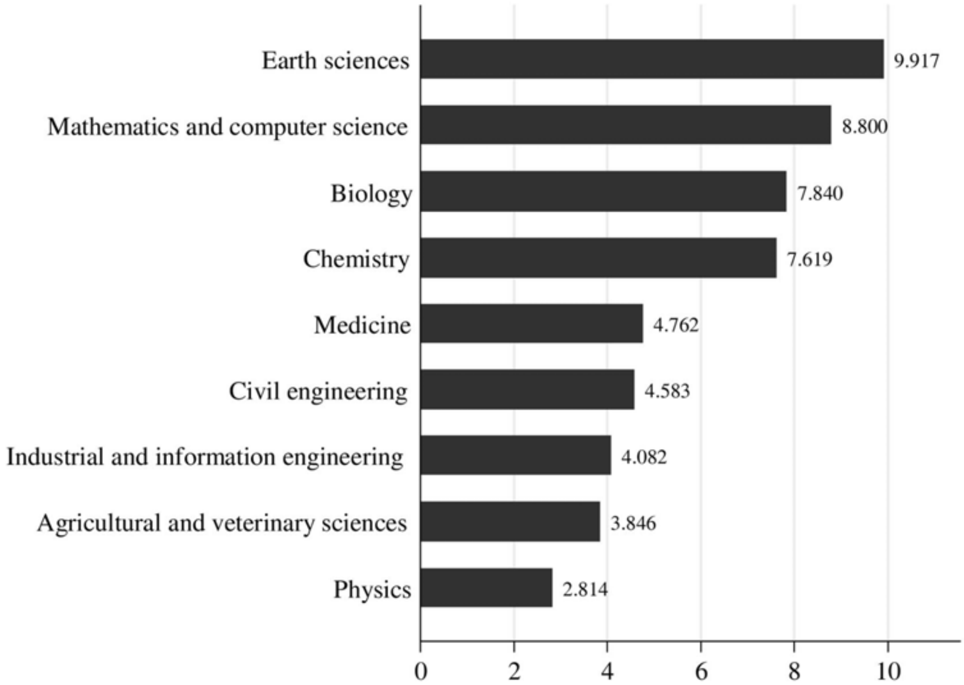


**Figure 1.** Values (%) of the indicator *R*′ per university disciplinary areas.

different from our previous study on the ranking of individual scientists[39], the results here show a strong correlation (never below 0.96) between the two ranking lists: one that did not distinguish by gender and the other that did. As could be expected, the gender productivity gap tends to have limited impact on the comparative aggregated performance of an entire organization, in part because at the level of entire disciplinary areas, the distribution of genders among the universities is not particularly heterogeneous.

Still, we should not ignore some of the shifts in performance observed at the level of the specific university disciplinary areas. For example, in mathematics and computer science, 10 of the 50 universities evaluated showed significantly different productivities under the two methods of evaluation, and two of these universities shifted nine positions. The shifts in positions in biology, and earth sciences are not negligible, whereas the generally high levels of correlation between the rankings also hide diverse and important outliers.

If the objectives of the national evaluation exercises are to stimulate improvement in the general performance in the system, to permit the users to make informed choices, and eventually to guide the allocation of resources (as in Italy and in a growing number of other countries), it is important that all factors exogenous to the true merit of the subjects evaluated be appropriately controlled for. This does not mean that we intend to issue a priori recommendations on the suitability of conducting comparative evaluation of research performance that would take gender into account. Our current objective is to measure to what extent the comparison of research performance of institutions with and without distinguishing by gender leads to rank positions that are detectably different. We leave it to the decision-makers to choose which approach to adopt, given the objectives of the evaluation and the conditions of the context. He/she should also consider that the choice to conduct evaluation exercises distinguished by gender may itself be interpreted by women scientists as a form of unnecessary and unwanted discrimination. Women researchers might in fact perceive the procedure as implying a cognitive gap in favour of men.

1. OECD, Main Science and Technology Indicators. OECD Science, Technology and R&D Statistics (database), 2015; DOI:10.1787/data-00182-en. Retrieved on-line from http://stats.oecd.org/Index.aspx?DataSetCode=MSTI_PUB (last accessed 6 November 2015).
2. Mauleón, E. and Bordons, M., Productivity, impact and publication habits by gender in the area of material science. *Scientometrics*, 2006, **66**(1), 199–218.
3. Lee, S. and Bozeman B., The impact of research collaboration on scientific productivity. *Soc. Stud. Sci.*, 2005, **35**(5), 673–702.
4. Xie, Y. and Shauman, K. A., Sex differences in research productivity: new evidence about an old puzzle. *Am. Sociol. Rev.*, 1998, **63**(6), 847–870.
5. Long, J. S., Measure of sex differences in scientific productivity. *Soc. Forces*, 1992, **71**(1), 159–178.
6. Fox, M. F., Publication productivity among scientists: a critical review. *Soc. Stud. Sci.*, 1983, **13**(2), 285–305.
7. Frietsch, R., Haller, I., Funken-Vrohlings, M. and Grupp, H., Gender-specific patterns in patenting and publishing. *Res. Policy*, 2009, **38**(4), 590–599.

8. Abramo, G., D'Angelo, C. A. and Caprasecca, A., Gender differences in research productivity: a bibliometric analysis of the Italian academic system. *Scientometrics*, 2009, **79**(3), 517–539.
9. Alonso-Arroyo, A., González-Alcaide, G., Valderrama-Zurián, J. C. and Aleixandre-Benavent, R., Gender analysis of papers published in *Actas Espanolas De Psiquiatria* (1999–2006). *Actas ESP Psiquiatr.*, 2007, **36**(6), 314–322.
10. Leahey, E., Gender differences in productivity: research specialization as a missing link. *Gender Soc.*, 2006, **20**(6), 754–780.
11. Xie, Y. and Shauman, K. A., Women in science: career processes and outcomes (review). *Soc. Forces*, 2004, **82**(4), 1669–1671.
12. Cole, J. R. and Zuckerman, H., The productivity puzzle: persistence and change in patterns in publication of men and women scientists. In *Advances in Motivation and Achievement* (eds Steinkamp, M. W. and Maehr, M. L.), Greenwich, Connecticut: JAI Press, 1984, vol. 2, pp. 217–258.
13. Xie, Y. and Shauman, K. A., Sex differences in research productivity: new evidence about an old puzzle. *Am. Sociol. Rev.*, 1998, **63**, 847–870.
14. Abramo, G., D'Angelo, C. A. and Caprasecca, A., The contribution of star scientists to sex differences in research productivity. *Scientometrics*, 2009, **81**(1), 137–156.
15. Bordons, M., Morillo, F., Fernández, M. T. and Gómez, I., One step further in the production of bibliometric indicators at the micro level: differences by gender and professional category of scientists. *Scientometrics*, 2003, **57**(2), 159–173.
16. Ding, W. W., Murray, F. and Stuart, T. E., Gender differences in patenting in the academic life sciences. *Science*, 2006, **313**(5787), 665–667.
17. Zainab, A. N., Personal, academic and departmental correlates of research productivity: a review of literature. *Malays. J. Libr. Inf. Sci.*, 1999, **4**(2), 73–110.
18. Rossiter, M. W., The Matilda effect in science. *Soc. Stud. Sci.*, 1993, **23**(2), 325–341.
19. Van Den Brink, M., Brouns, M. and Waslander, S., Does excellence have a gender? A national research study on recruitment and selection procedures for professorial appointments in the Netherlands. *Employee Relations*, 2006, **28**(6), 523–539.
20. Sambunjak, D., Straus, S. E. and Marusic, A., Mentoring in academic medicine: a systematic review. *J. Am. Med. Assoc.*, 2006, **296**(9), 1103–1115.
21. Ley, T. J. and Hamilton, B. H., The gender gap in NIH grant applications. *Science*, 2008, **322**(5907), 1472–1474.
22. Mauleón, E., Bordons, M. and Oppenheim, C., The effect of gender on research staff success in life sciences in the Spanish National Research Council. *Res. Eval.*, 2008, **17**(3), 213–225.
23. Fox, M. F., Gender, family characteristics, and publication productivity among scientists. *Soc. Stud. Sci.*, 2005, **35**(1), 131–150.
24. Stack, S., Gender, children and research productivity. *Res. Higher Educ.*, 2004, **45**(8), 891–920.
25. Kyvic, S. and Teigen, M., Child care, research collaboration, and gender differences in scientific productivity. *Sci., Technol., Human Values*, 1996, **21**(1), 54–71.
26. Bozeman, B. and Corley, E., Scientists' collaboration strategies: implications for scientific and technical human capital. *Res. Policy*, 2004, **33**(4), 599–616.
27. Abramo, G., D'Angelo, C. A. and Solazzi, M., The relationship between scientists' research performance and the degree of internationalization of their research. *Scientometrics*, 2011, **86**(3), 629–643.
28. Abramo, G., D'Angelo, C. A. and Murgia, G., Gender differences in research collaboration. *J. Informetr.*, 2013, **7**(4), 811–822.
29. Badar, K., Hite, J. H. and Badir, Y. F., Examining the relationship of co-authorship network centrality and gender on academic research performance: the case of chemistry researchers in Pakistan. *Scientometrics*, 2013¸ **94**(2), 755–775.
30. Larivière, V., Vignola-Gagné, E., Villeneuve, C., Gelinas, P. and Gingras, Y., Sex differences in research funding, productivity and impact: an analysis of Quebec university professors. *Scientometrics*, 2011, **87**(3), 483–498.
31. Kyvik, S. and Teigen, M., Child care, research collaboration, and gender differences in scientific productivity. *Sci. Technol. Human Values*, 1996, **21**(1), 54–71.
32. McDowell, J. M., Larry, D., Singell Jr, L. D. and Stater, M., Two to tango? Gender differences in the decisions to publish and coauthor. *Econ. Inquiry*, 2006, **44**(1), 153–168.
33. McDowell, J. M. and Smith, K. J., The effect of gender-sorting on propensity to coauthor: implications for academic promotions. *Econ. Inquiry*, 1992, **30**(1), 68–82.
34. Duch, J., Zeng, X. H. T., Sales-Pardo, M., Radicchi, F., Otis, S., Woodruff, T. K. and Amaral, L. A. N., The possible role of resource requirements and academic career-choice risk on gender differences in publication rate and impact. *PLoS ONE*, 2012, **7**(12), e51332.
35. Ceci, S. J. and Williams, W. M., Understanding current causes of women's underrepresentation in science. *Proc. Natl. Acad. Sci. USA*, 2011, **108**(8), 3157–3162.
36. Abramo, G. and D'Angelo, C. A., How do you define and measure research productivity? *Scientometrics*, 2014, **101**(2), 1129–1144.
37. Abramo, G., D'Angelo, C. A. and Rosati, F., The importance of accounting for the number of co-authors and their order when assessing research performance at the individual level in the life sciences. *J. Informetr.*, 2013, **7**(1), 198–208.
38. Abramo, G., Cicero, T. and D'Angelo, C. A., Individual research performance: a proposal for comparing apples to oranges. *J. Inform.*, 2013, **7**(2), 528–539.
39. Abramo, G., Cicero, T. and D'Angelo, C. A., Should the research performance of scientists be distinguished by gender? *J. Inform.*, 2015, **9**(1), 25–38.

ACKNOWLEDGEMENT. The work has received no financial support by third parties.

Received 24 April 2015; revised accepted 11 September 2015

doi: 10.18520/v109/i10/1783-1789