Journal of Informetrics 10 (2016) 151–161

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Drivers of academic performance in a Brazilian university under a government-restructuring program

José Alexandre F. Diniz-Filho[a,c,*], Maria Clorinda S. Fioravanti[b], Luis Mauricio Bini[c], Thiago Fernando Rangel[c]

[a] *Pró-Reitoria de Pós-Graduação, Universidade Federal de Goiás (UFG), Prédio da Reitoria, C.P. 131, 74001-970 Goiânia, GO, Brazil*
[b] *Pesquisa e Inovação, Escola de Veterinária e Zootecnia, UFG, Goiânia, Brazil*
[c] *Departamento de Ecologia, Instituto de Ciências Biológicas, UFG, Goiânia, Brazil*

ARTICLE INFO

Article history:
Received 19 April 2015
Received in revised form 1 December 2015
Accepted 4 December 2015
Available online 15 January 2016

Keywords:
Individual researcher productivity
Teaching load
Supervision
Scientific production
Brazilian universities

ABSTRACT

The search for correlates of scientific production is an important step toward the formulation of decision-making guidelines on academic and funding policy under a competitive system with continuously reduced budgets. Our goal here is to identify drivers of the scientific production of researchers working at the "Universidade Federal de Goiás" (UFG), a medium-to-large public Brazilian University, focusing on the effects of teaching load and supervision of graduate and undergraduate students on scientific production of faculty members. We analyzed data for 1487 faculty members of UFG, including the total number of papers published between 2011–2013, a weighted-index of scientific production and the number of publications in high-ranked journals (according to a Brazilian system of journal ranking in different areas). These variables were regressed on gender, teaching load at undergraduate and graduate levels, number of supervised undergraduate, Master and Doctoral students, self-declared amount of time dedicated to research and outreach, year of doctoral graduation, year of hiring and the scientific production before doctoral graduation. Several regression models were used to model scientific production, including ordinary least-square regression and Hurdle negative binomial models. Although there are some differences among research areas, the most important explanatory variable was the publication record of the researcher before doctoral graduation, reinforcing the role of a solid academic formation in terms of research experience. Undergraduate and graduate teaching loads were negatively and positively correlated with scientific production, respectively. However, the strength of the relationship was much higher for the second than for the first relationship. These correlates of scientific production provide guidelines for policy and management in universities, including criteria for balancing research and teaching loads, awarding fellowships and research grants, designing new policy for future hiring and creation of new graduate programs.

* Corresponding author at: Universidade Federal de Goiás (UFG), Pró-Reitoria de Pós-Graduação, Prédio da Reitoria, C.P. 131, 74001-970 Goiania, GO, Brazil. Tel.: +55 62 3512 14 80.
E-mail address: diniz@ufg.br (J.A.F. Diniz-Filho).

http://dx.doi.org/10.1016/j.joi.2015.12.004
1751-1577/

1. Introduction

Assessing the correlates of scientific research activity at several organizational levels, ranging from individual researchers up to institutes and universities, is a key goal of scientometric analysis (Carayol & Matt, 2006). However, there is an intense debate about the most appropriate metrics for evaluating research merit (e.g., Vieira & Gomes, 2010; Abramo, D'Angelo, & Di Costa, 2011a; Abramo, D'Angelo, & Solazzi, 2011b; Abramo, Cicero, & D'Angelo, 2011c; Franceschet & Costantini, 2011; McNutt, 2014; Abramo & D'Angelo, 2011; Hicks, Wouters, Waltman, de Rijcke, & Rafols, 2015) and how to take into account inherent differences among research areas (Abramo, Cicero, & D'Angelo, 2013a; Ruocco & Daraio, 2013). Even so, under the current scenario of strong competition for limited funds, scientometric analyses have been proven to be an important tool to find correlates of scientific productivity, providing guidelines for decision-making in, for instance, faculty hiring, affirmative actions and funding (e.g., Abramo & D'Angelo, 2011; Akhmat, Zaman, Shukui, Javed, & Khan, 2014). According to these analyses, factors at individual level (e.g., age, gender, seniority, job stability, involvement in collaborative networks, and working conditions) are important predictors of variation in scientific production among researchers (e.g., Carayol & Matt, 2006; Rice, Venables, & Patacchini, 2006; Gonzalez-Brambila & Veloso, 2007; De Witte & Rogge, 2010; Cruz-Castro & Sanz-Menéndez, 2010; Perianes-Rodríguez, Chinchilla-Rodríguez, Vargas-Quesada, Gómez, & Moya-Anegón, 2009; Abramo, D'Angelo, & Di Costa, 2009; Abramo et al., 2011a, 2011b; Abramo, D'Angelo, & Rosati, 2014; Pachi, Yamamoto, Costa, & Lopez, 2012; McCarty, Jawitz, Hopkins, & Goldman, 2013; Bauer, Schui, Eye, & Krampen, 2013; Miller, Coble, & Lusk, 2013; Torrisi, 2013; Çokgezen, 2013; Silaghi-Dumitrescu & Sabau, 2014; Baccini, Barabesi, Cioni, & Pisani, 2014). However, the relative importance of these potential correlates varies from study to study, and the underlying reasons for the importance of these correlates are also open to debate (see Baccini et al., 2014 for a recent review). It is clear, therefore, that more empirical studies in different systems and in countries with different research traditions are needed to better understand the generality and relative roles of drivers of scientific production.

There are now about 6000 universities in Brazil, in different categories (public and private universities), with about 37,500 undergraduate courses and almost 4000 graduate (Master and Doctoral) programs (see http://emec.mec.gov.br/). In general, public (state or federal) universities offer top-ranked undergraduate courses and graduate programs. Starting in 2007–2008, under the auspicious of the Supporting Program for the Restructuring and Expansion of Federal Universities (REUNI, its acronym in Portuguese), the Federal System of Public Universities increased more than two-fold in faculty size and number of graduate students, with 65 federal universities and about 300 campuses now widespread throughout the country. The expansion of the system involved funding for building new infrastructure, a massive hiring of new faculty members, staff, and the admission of an increased number of students. As a consequence, new graduate programs were also created because young, productive doctors were hired and wanted to start their own programs focused in new research lines. Research budgets were also increased to support this growing research and graduate system, within the context of a, at least until up to 2014, favorable economic scenario (see Escobar, 2015) This expansion has been important to increase the proportion of Brazilians with higher education degrees and to foster the scientific production in the country (see Regalado, 2010; Almeida & Guimarães, 2013).

Brazil has two main research funding agencies at the national level. First, the National Counsel of Technological and Scientific Development (CNPq), under the Ministry of Science, Technology and Innovation, mainly funds individual researchers through grants and scholarships. Second, the Coordination for the Improvement of Higher Level Personnel (CAPES), under the Ministry of Education and Culture funds and evaluate graduate programs. Since 2007, the mission of CAPES was expanded to include issues related to formation of teachers at all levels (Myers, 2011). In general, proposal ratings by both CNPq and CAPES are based on peer-reviewing and scientometric-based assessments. However, universities seldom use comparable metrics to evaluate their own faculty members or to promote incentives for scientific publication and innovation. One of the most controversial issues discussed in Brazilian universities is the distribution of teaching load among faculty members, which usually does not take into account other activities related to research. The discussion is how (and if) scientific productivity is constrained by other academic activities, especially teaching and administration. Evidence-based answers to this question would lead to better definition of academic policies. Thus, it is important to identify the underlying drivers of the variation in production, providing criteria to balance teaching load, research and administrative activities.

Our goal here is to identify correlates of scientific production of researchers in Brazil. For this task, we use data obtained from scholars working at the Federal University of Goiás (UFG), as a case study. We hypothesized that variation in scientific production can be explained by a set of predictors related to different academic activities (e.g., teaching load, number of supervised students, outreach and administrative activities) and to some faculty members-specific characteristics (gender, years of doctoral degree and admission).

We are aware that our analysis is restricted to a single university (similarly to the studies of De Witte & Rogge, 2010; Silaghi-Dumitrescu & Sabau, 2014; Baccini et al., 2014) and, beforehand, we suggest that further studies are needed to gain generalization. However, considering the similarities between our results and those from previous studies, we believe that our findings are consistent enough to be generalized to other similar-sized universities in Brazil. This is so because similar statutes, mainly the Law of Directives and Bases of National Education (LDB), regulate all the Federal System of Universities in Brazil. For instance, under LDB, the minimum teaching load in public universities is 8 h/week (Article n. 57). Also, the main Brazilian funding agencies (CNPq and CAPES) have unified criteria at national level, within research areas, to evaluate research proposals, homogenizing to some extent the perception of the level of scientific productivity needed for successful applications.

2. Methods

The UFG is a medium-to-large institution (see www.ufg.br). It was formally created in 1960, by the fusion of several independent colleges. Currently, UFG has about 25,000 students attending to ca. 120 undergraduate courses and 70 graduate (MSc and Doctoral) programs. These continuously growing numbers are very recent, as UFG almost doubled in terms of faculty size and number of students within the last 5 years. During this time span, several new faculty members were hired (about 52% of the 2341 members in December 2014 were hired after 2009) and many new graduate and undergraduate courses were created. The recent and rapid expansion of UFG generated enormous heterogeneity among faculty members, which is useful for an exploratory analysis of the drivers of scientific activity. For our analyses we identified the 1487 faculty members of UFG with a Doctoral or PhD degree in 2013, excluding temporary lecturers and those involved in rectory administration. We think that the exclusion of temporary lecturers from our dataset is justifiable because it is not expected that they contribute to the scientific productivity of the university as they are, besides temporary (maximum of 2 years), part-time workers. A similar reasoning can be applied to justify the exclusion of full-time scholars that are involved in administrative tasks.

2.1. Data

We obtained data on scientific production and student supervision from the Lattes System, a widespread online Brazilian repository of CVs, maintained by CNPq (www.cnpq.br/lattes; named after the Brazilian physicist Césare Mansueto Giulio Lattes). The CNPq created the Lattes System in 1999 aiming to facilitate the comparison among scholars productivity when applying for research grants, the selection of reviewers for research grants and the generation of statistics regarding the temporal trends and spatial distribution of the Brazilian scientific production. In short, the Lattes System is an online application where scholars are able to create their Curriculum Vitae, including several relevant information for research assessment (e.g., basic personal information, list of projects, membership in editorial board of scientific journals, list of journals serving as reviewer, bibliometric indicators, list of articles and chapters, and list of supervised students; see http://buscatextual.cnpq.br/buscatextual/visualizacv.do?id=K4727587J2 for an example). The software was freely licensed for several countries in Latin America (Colombia, Ecuador, Chile, Peru, and Argentina).

The following variables were retrieved from each of the 1487 Lattes CVs from UFG faculty members between 2011 and 2013: (1) total number of papers, (2) number of top-ranked papers (see below) and (3) rank-weighted number of papers (see below). We also used the Lattes System and classified faculty members as belonging to 8 general areas of knowledge recognized by CNPq (see http://www.memoria.cnpq.br/areasconhecimento/index.htm): (1) Agrarian Sciences (e.g., veterinary, animal husbandry, and agronomy), (2) Arts and Linguistics (e.g., Portuguese, comparative literature, dramaturgy, and classic idioms), (3) Biological Sciences (e.g., ecology, genetics, zoology, botany, immunology, and parasitology), (4) Engineering Sciences, (5) Exact and Earth Sciences (e.g., physics, chemistry, computer sciences, and geology), (6) Humanities Sciences (e.g., sociology, education, history, and geography), (7) Health Sciences (e.g., medicine, odontology, and nursing), and (8) Applied Social Sciences (e.g., economy, administration, and law). Of course, there is some overlap among these areas and it is common that a faculty member classifies oneself into more than one of these areas (e.g., a computer scientist working with molecular genetics and bioinformatics). In these circumstances, we assigned faculty members to research areas similar to the department or institute where they are working. For instance, in the example above, the faculty member was considered as belonging to the Biological Sciences if he/she is working in a biological science institute or as belonging to the Exact and Earth Sciences, if he/she is working in a computer sciences department. Although these decisions adds a certain level of uncertainty, the classification was used to test and control for general differences among areas.

Scientific papers published by UFG faculty members were ranked according to the Qualis system developed by CAPES (see http://qualis.capes.gov.br/webqualis). This system was created to rank journals within areas in order to compare Brazilian graduate programs, taking into account the intrinsic differences among areas of knowledge (rather than simply comparing all journals in all areas by a single metric, such as JCR Impact Factor). Currently, the Qualis system classifies journals into 7 levels, from A1 (highest rank) to A2, B1, B2, B3, B4, B5 and C. The rank of each journal is defined independently by 48 committees of different knowledge areas, which are composed by researchers from all over the country (see http://www.capes.gov.br/avaliacao/sobre-as-areas-de-avaliacao). Committees are free to use any ranking criteria for the journals, which vary substantially among areas of knowledge. For example, while most committees use bibliometric indicators (i.e., impact factors) or the fact of a journal being or not ISI, Scopus or Scielo listed to rank journals, humanities usually use more subjective criteria. This ensures that each area is free to define what is a "very good" publication in its own area, and so this system could be used, in principle, to compare the scientific production in different areas. For instance, if a sociologist and a biologist publish ten A1 papers, then they can be considered equally "productive" in their own research areas. Because different committees may rank the same journal differently, each publication was scored according to the best Qualis score for the journal among the 48 committees. We believe that our analytical strategy is suitable to compare researchers from different fields of knowledge because we used a rank system defined by field-specific committees with freedom to define the most adequate criteria to rank quality of scientific publication within their area of knowledge.

In addition to (1) total number of papers and (2) number of papers in top-ranked (A1) journals, we also calculated a (3) rank-weighted number of papers. Assuming that a publication in a top-ranked (A1) journal weights 1 unit, the weight assigned to papers published in journals ranked in each level below decreases in steps of 0.15 (e.g. one paper published in a B5 journal receives 10% of the weight attributed to a paper published in a A1 journal). These differences in 0.15 of weight is

Table 1
Explanatory variables used in this study to model scientific production of UFG faculty.

Explanatory Variables	Definition
UNG	Teaching loading (hours per year) in the undergraduate courses
GRA	Teaching loading (hours per year) in the graduate (Master and Doctoral) courses
SDOC	Number of Doctoral students supervised (finished dissertations) between 2011–2013
SMAS	Number of Master students supervised (finished thesis) between 2011–2013
SIC	Number of Scientific Initiation (undergraduate) students supervised between 2011–2013
PFU	Self-declared number of hours dedicated to research projects with financial support
PNF	Self-declared number of hours dedicated to research projects without financial support
OUT	Self-declared number of hours dedicated to outreach (talks, short courses, interaction with society) activities
BEF	Scientific production (in the same units as the response variable) before finishing PhD or Doctoral
YDOC	Year in which the Doctoral or PhD was obtained
YUFG	Year starting as permanent faculty in UFG
GEN	Gender

the minimum difference between Qualis levels suggested by the CAPES' Superior Technical Council (CTC). It is then possible to calculate, for each researcher, the total number of "A1-equivalent" papers (A1E hereafter). For example, if a researcher published 5 papers, including 2 papers in A1 journals, 2 papers in A2 journals and 1 paper in a B1 journal, the rank-weighted number of papers is given by:

$$\text{A1E} = (2 \times 1) + (2 \times 0.85) + (1 \times 0.70) = 4.4$$

Because A1E is a weighted score, different combinations of publication "quality" can lead to similar A1E values. For example, one researcher may publish few papers on A1 journals whereas another researcher publishes twice as many papers in B2 journals. Thus, to differentiate between quantity and quality of publication, we also analyzed the number of papers published only in top-ranked A1 journals.

To avoid underestimating scientific production of researchers in humanities, which traditionally publish substantially more books (and eventually book chapters), we assigned an A2 rank for each book and B2 rank for each book chapter. The Pearson's correlation coefficient between the number of A1E papers (excluding books) and the number of works published (including books) was $r = 0.91$, indicating that accounting for publications in books does not affect the results, as most scientific production was published in journals.

Several variables were used as explanatory variables for our response variables (i.e., total number of papers, number of A1 papers, and A1E) (see Table 1). For supervision of students, we counted the number of Doctoral, Master and "Scientific Initiation" (undergraduate students) supervised between 2011 and 2013. The Scientific Initiation program of fellowships is a very successful initiative of CNPq that started more than 20 years ago, with the goal of supervising undergraduate students, even for freshman-level students.

We also obtained from the university administration database, the average annual teaching load of each researcher (at undergraduate and graduate levels), self-declared amounts of time (in hours) spent annually in scientific research (funded and non-funded research projects), technological innovation and outreach activities. Finally, we also recorded researcher's gender (as a dummy variable: female: 0 and male: 1), year when a researcher was hired at UFG and year of Doctoral degree.

Finally, to account for the effects of level of academic formation and research tradition, we also measured the A1E scientific production up to the year that each research obtained his/her PhD or Doctoral (including all papers published before the year in which the research finished his/her PhD or Doctoral). We hypothesize that researchers that were formed in a more productive research group and published earlier in their career keep producing and this fact enhanced their scientific activities in the future, and to evaluate this we counted all papers produced before the Doctoral or PhD and accumulated the production (using Qualis weights described above).

2.2. Analyses

We modeled the total number of papers, number of A1 papers and A1E papers as response variables using different regression models. For A1E, which is actually a continuous variable, we log-transformed the response to normalize the distribution of residuals and fitted a linear model with all the previously defined explanatory variables (e.g., Yu, Yu, Li, & Wang, 2014). However, due to the presence of many zeros in the data, we also fitted a series of count models after rounding A1E to the nearest integer (also generating a model directly comparable with the ones used for the other two response variables, A1 and papers). These hurdle models were also used for A1 papers and total number of papers, following closely the methods described in Zeileis, Kleiber, and Jackman (2008) (see also Baccini et al., 2014) and using several routines in R (R Core Team, 2015), as described below.

We started with Quasi-Poisson and Negative Binomial models using *glm* and *glm.nb* functions, allowing estimating the θ dispersal parameter for the last model (Zeileis et al. (2008). In addition to these overdispersed models, we also used two-component models, which have been found effective in handling highly right-skewed data with a high proportion of zeros (Zeileis et al., 2008; Zuur, Ieno, Walker, Saveliev, & Smith, 2009). We fitted hurdle models (HNB), using Poisson, geometric, and negative binomial link functions to model both positive counts and zero versus larger counts, and compared these alternative models with Akaike Information Criterion. In this context, we also fitted a Zero-Inflated Negative Binomial (ZINB) model. In both HNB and ZINB models, Wald tests were used to evaluate if addition of hurdle or zero-inflated components significantly improved model fit. HNB and ZINB models were fitted using the *pscl* package for R (Zeileis et al., 2008).

For models with likelihood estimates, fit was evaluated with Akaike Information Criterion (see Burnham & Anderson, 2012), and compared with an intercept-only model. To provide descriptive statistics allowing comparison of models fit, a pseudo-R^2 was computed by correlating observed and fitted values for each response variable. Explanatory variables were standardized before the analyses in order to make the partial coefficients comparable to each other.

We also incorporated the areas of knowledge as a categorical (dummy) variable in these models, so the effects of other explanatory variables must be interpreted as partial coefficients in respect to area. Models with and without the effect of area were also compared using AIC. Finally, after comparing the results of all models described above and defining which model better describes the data, we fitted separate models for each area, to obtain further insights about the relative importance of explanatory variables.

3. Results

3.1. Overall description of scientific production

Researchers at UFG published on average 4.3 papers between 2011 and 2013 (1.43 per year), ranging from 0 to 63 (Fig. 1a). About 30% of researchers did not have any publications in the last three years. Researchers published, on average, 2.87 A1E papers (ranging from 0 to 51.5), and 0.989 papers in A1 journals, ranging from 0 to 32 (Fig. 1b and c). Only 18.7% of the papers

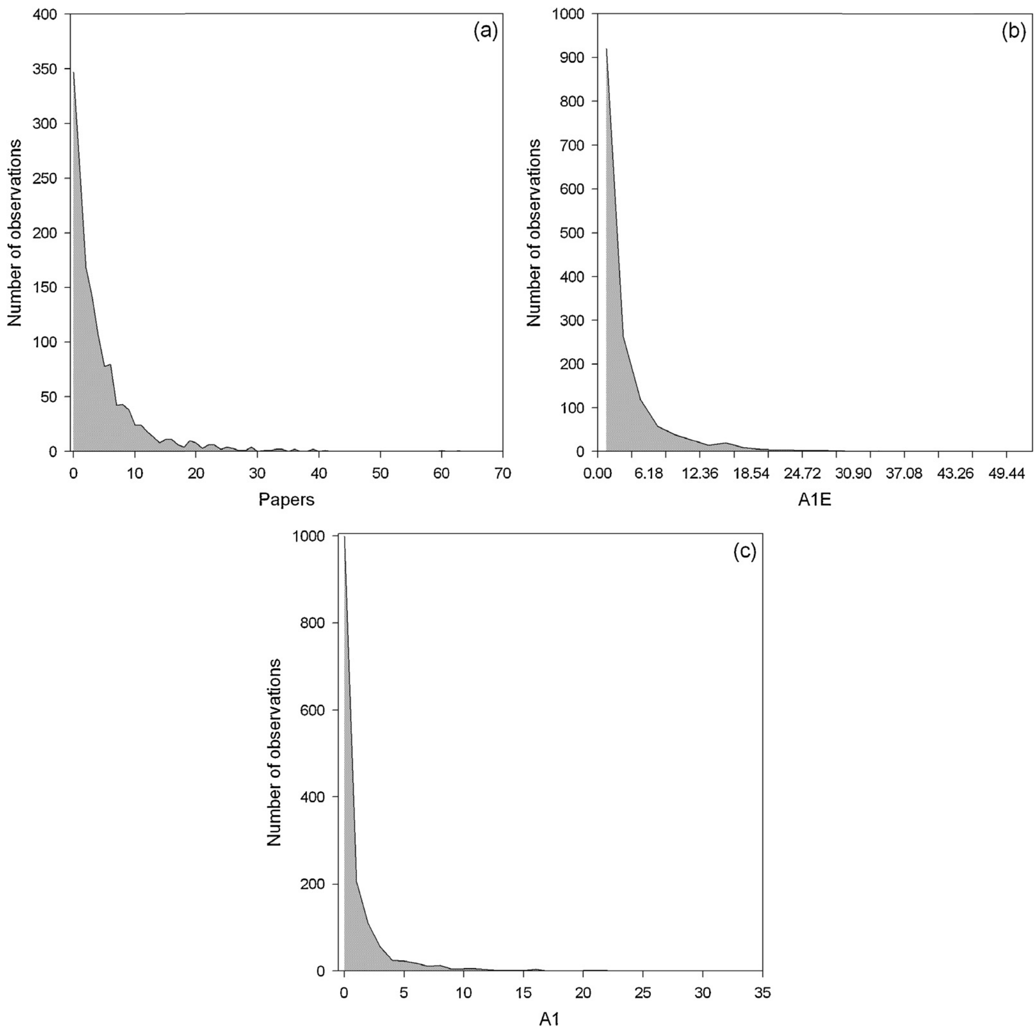

Fig. 1. Frequency distribution of number of papers (a), A1E (b) and A1 (c) papers by the 1487 faculties of Universidade Federal de Goiás (UFG) between 2011–2013.

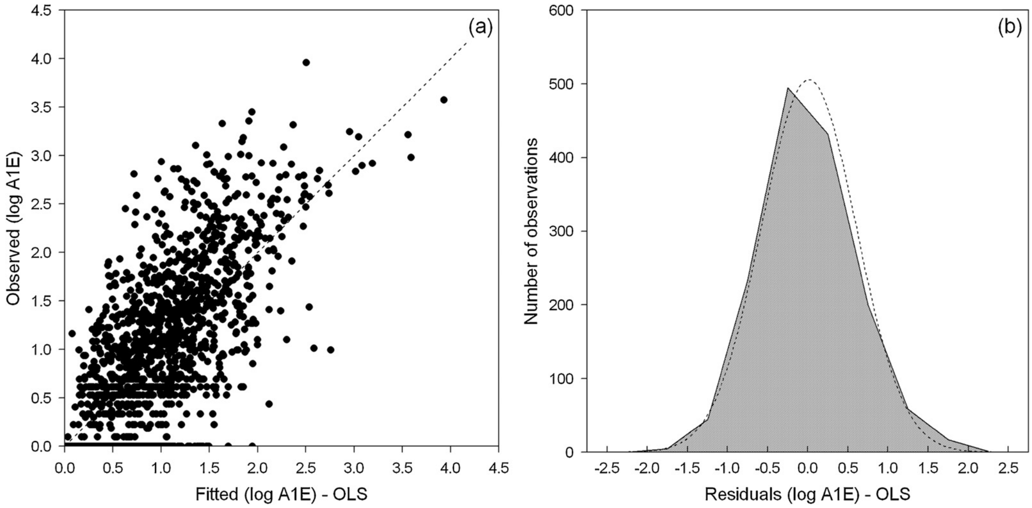

Fig. 2. Fitted versus estimated values (a) and the residuals distribution (b) of the OLS model for A1E papers. The dashed line in (b) indicates the expected normal distribution.

were published in A1 journals. Indeed, the total number of papers published is highly correlated with A1E score ($r = 0.95$) and number of A1 papers ($r = 0.75$). As expected by Lotka's law, these metrics have highly skewed distributions. For instance, the frequency distribution of the response variables reveals that few faculty members published many papers, whereas most faculty members published very few ones (Fig. 1).

3.2. Models for A1E

For the response variable A1E, the coefficient of determination of the OLS multiple regression model was equal to 48.7% ($F = 75.46$; $P > 0.01$; Fig. 2a), after incorporating the effect of area of knowledge (which improved model fit; ΔAIC = 132.56). Residuals from this model were nearly normally distributed (Fig. 2b). However, due to the large number of zero counts, we also fitted the count models after rounding A1E to the nearest integer.

Quasi-Poisson and Negative Binomial, as well as count components of Hurdle Negative Binomial (HNB) and zero-inflated negative binomial (ZINB) models for A1E, indicated relationships similar to the ones previously described by OLS (see below) and reconstruct relatively well the distribution of A1E. However, the parameter θ of the negative binomial variance function was relatively high (equal to 2.74), so the two-component models, HNB and ZINB, are expected to improve model fit. Hurdle models using a negative binomial function fitted data better than ZINB models (ΔAIC = 38.61), so only detailed results for HBN will be presented. For HNB, about 41% (pseudo-R^2) of the variance in A1E among researchers can be explained by the 12 explanatory variables combined, correctly approximating the observed number of zero values in A1E (393 modeled zeros, against 464 observed zeros).

Partial regression coefficients from OLS and HNB for A1E are similar and reveal some general trends (Table 2). For instance, undergraduate teaching load was negatively related to A1E. Conversely, graduate teaching load had a positive relationship with scientific production. In addition, a positive relationship with the number of supervised doctoral students, as well as with the number of supervised masters and undergraduate students, was also observed. We found no relationship between scientific production measured by A1E and time dedicated to non-funded research project. Comparatively, we found a stronger positive relationship between research production and self-declared time dedicated to funded research projects. In general, researchers dedicated to outreach have smallest scientific production (in terms of A1E and total number of papers). The year of doctoral graduation was negatively correlated with A1E, indicating that young researchers are generally less productive than senior researchers. Conversely, the year of hiring was positively correlated with scientific production, indicating that researchers hired more recently tend to be more productive. There is a significant effect of gender in the count component of HNB, different from OLS. Again, we highlight that scientific production before doctoral graduation was the best predictor of scientific production during the three years period (2011–2013).

In addition to the main trends described above for OLS and count components of HNB, Wald test revealed a significant improvement in model fit for the Hurdle component ($\chi^2 = 64.3$; $P < 0.01$). The coefficients of the hurdle component, describing the difference between zero and positive counts, are not conspicuously different between each other. However, in the zero component there is a clearer contrast between the negative and positive effects of undergraduate and graduate teaching load, respectively, a higher effect of the supervision of doctoral students (in comparison to the production before the doctoral in the count component) and the number of years since the researcher started in the UFG as a permanent professor. A significant negative relationship with time dedicated to outreach was also observed, differently from the non-significant

Table 2
Effects of the explanatory models in the three different metrics for scientific production of UFG faculty, expressed as partial regression coefficients (to save space, intercept were omitted). Models for A1E include an OLS model of log-transformed values, and the count and hurdle components of Hurdle Negative Binomial (HNB) models. Coefficients for HNB for total number of papers (PAPERS) and A1 papers are also shown. All effects are partial in respect to the areas of knowledge. Bold-faced coefficients indicate $P < 0.05$. R^2 for each model is also shown (calculated as a pseudo-R^2 for Hurdle models). All models possess very high ΔAIC when compared with intercept-only model.

Models	OLS	Hurdle					
Responses	A1E	A1E (Rounded)		PAPERS		A1	
Components		Count	Zero	Count	Zero	Count	Zero
UNG	**−0.074**	**−0.107**	**−0.266**	−0.052	**−0.232**	−0.098	**−0.141**
GRA	**0.131**	**0.149**	**0.564**	**0.130**	**0.537**	0.109	**0.288**
SDOC	**0.129**	**0.120**	**1.070**	**0.131**	**0.846**	**0.136**	**0.202**
SMAS	**0.133**	**0.178**	**0.229**	**0.179**	**0.264**	**0.162**	0.139
SIC	**0.118**	**0.106**	**0.538**	**0.153**	**0.505**	**−0.177**	0.002
PFU	**0.128**	**0.156**	**0.377**	**0.137**	**0.230**	**0.220**	**0.506**
PNF	0.015	−0.023	0.155	−0.017	0.137	−0.124	0.039
OUT	−0.020	0.066	**−0.205**	**0.119**	**−0.186**	−0.073	−0.001
BEF	**0.339**	**0.447**	**1.109**	**0.536**	**1.301**	**0.587**	0.851
YDOC	**−0.092**	**−0.192**	−0.207	**−0.242**	−0.197	**−0.193**	**−0.282**
YUFG	**0.118**	**0.099**	**0.510**	0.077	**0.345**	0.034	**0.406**
GEND	0.008	**0.070**	−0.015	**0.060**	−0.085	−0.024	−0.089
R^2	0.490	0.417		0.322		0.311	

effect observed for the count component. The effect of gender, which is significant in the count component, is not significant in the zero component.

3.3. Models for A1 papers and total number of papers

Results for modeling total number of papers and A1 papers, which are count variables, were again better modeled with HNB (Table 2), but pseudo-R^2 are smaller than those found for OLS and HNB. Results are qualitatively similar to those previously described for A1E. However, some differences were detected. For instance, the count component of HNB model for the total number of papers was positively and significantly related to outreach, but not to year of hiring or undergraduate teaching load. The effect of gender is marginally significant in the count component ($P = 0.0499$).

The count component for A1 was positively and significantly related to the number of graduate students, time spent working on funded projects and the scientific production before doctoral graduation. Negative and significant relationships were found for the number of supervised undergraduate students and year of doctoral graduation. For the zero component, we also found significant positive relationships between A1 papers and graduate teaching load, number of supervised doctoral students, time spent working on funded projects and year of hiring. Negative and significant relationships were found for undergraduate teaching load and year of doctoral graduation (Table 2). It is interesting to highlight that some of the effects previously described for A1E and total number of papers, such as the contrast between negative and positive teaching loading in undergraduate and graduate courses, as well as time dedicated to funded projects, appear more clearly in the zero component. Effect of gender is not significant neither in count or zero component.

3.4. Models for areas

The overall patterns described above varied slightly among different areas of knowledge, but some consistent results can be highlighted. To save space and considering the previous results for the overall dataset, only models for A1E using OLS will be presented here. The adjusted R^2 of the models ranged from 29% (for Engineering, Humanities, and Applied Social Sciences) up and 53% (for Biological Sciences), all indicating improvement over a null model according to F-statistics (Fig. 3).

The signs of the coefficients are in general similar to the ones obtained with the full dataset, although in some cases the coefficients are not statistically significant. The importance of scientific production before doctoral degree in explaining the scientific production was consistent among all areas, with significant coefficients. Negative effects of undergraduate teaching load were significant only for Biological and Earth/Exact Sciences. Positive effects of graduate teaching load were detected for Biological, Engineering, Earth/Exact, Health and Applied Social Sciences. Supervision of students in different levels was significantly related to A1E in different areas, as well as time dedicated to funded projects. In line with the results obtained with the full dataset, the regression coefficients associated to time dedicated to non-funded projects and gender were not statistically significant across all areas. For outreach, there is only a negatively significant coefficient for biological sciences.

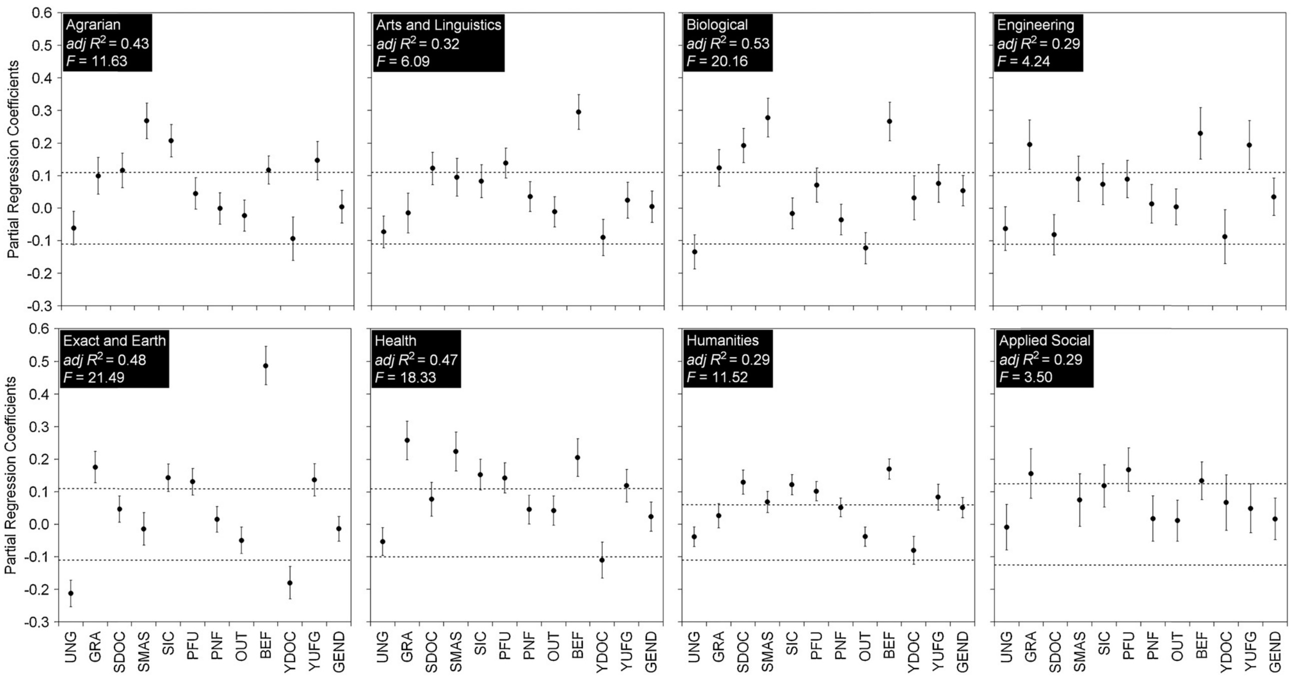

Fig. 3. Partial regression coefficients (±SE) from ordinary least squares (OLS) models regressing A1E papers against different explanatory variables (see Table 1 for codes). The intercepts of the models were omitted to facilitate visualization. The adjusted R^2 of the overall regression models, including all explanatory variables, are shown. Coefficients that are not within the limits indicated by the dashed lines are significant ($P < 0.05$).

4. Discussion

Our main goal was to understand how scientific production varies among researchers in a medium-to-large Federal University in Brazil. Similar studies were done elsewhere (e.g., Carayol & Matt, 2006; Çokgezen, 2013; Silaghi-Dumitrescu & Sabau, 2014; Baccini et al., 2014). We hypothesized that scientific production of researchers is influenced by a number of factors, including their teaching load, amount of time dedicated to research, number of undergraduate and graduate students, "experience" (time since doctoral graduation and time since hiring), and gender. We found that scientific production before the doctoral graduation was the best predictor of scientific production when the full dataset was analyzed. Also, it was the most consistent predictor (i.e., always significant) when the areas of knowledge were analyzed separately. Although it may be difficult to establish causal relationships based on correlative approaches, we believe that our models provide interesting results, which can be useful to suggest guidelines for improving management policies in the University.

Our findings reinforce the role of a solid academic formation, as approximated by scientific production before doctoral graduation (see Abramo et al., 2011a). In addition, our findings could be used to set new guidelines for human resource administration, such as faculty hiring policies and sabbatical leaves (see Abramo et al., 2014). In general, it would be important for the university to hire productive young doctors, which is now possible due to the consolidation, in the last 20 years, of the graduate system in Brazil supported by CAPES and other agencies from Federal and State Governments. In some institutes, a good hiring process happened during the last 5 years within the context of the REUNI program, by considering the positive effect of the year of hiring in our models. This important effect is not a mathematical artifact because many faculty members have recently completed the PhD. If we exclude from the sample the faculty members that completed the PhD after 2010 (thus excluding the overlap between the time spans used to calculate A1E), the model with 1140 faculty members has a similar fit (45% for the OLS model) and the production before PhD or Doctoral is still the most important predictor. The hiring process in most Brazilian federal universities is based on the assessment of applicants' CV, a written exam, a teaching event and in a defense of their scientific production. In general, teaching is the component of the selection process with the highest weight to decide which applicant will be hired. However, considering our results and the clear need to increase the quality and quantity of Brazilian research production, we think that the teaching component is overrated in the selection process, besides being the more subjective among the four components. Thus, as a first recommendation, we suggest rethinking the weight given to this component, so traditional in Brazilian universities (for a similar discussion in Italy, see Abramo, Cicero, & D'Angelo, 2013b). In a broader context, the importance of production before PhD or Doctoral in predicting future scientific production may raise controversial issues. For instance, the use of simple metrics in academic evaluation has been, in our opinion, convincingly criticized. In a nutshell, they "must not substitute for informed judgment" (Hicks et al., 2015). However, considering our results and the usually large number of candidates for an academic position, the use of these metrics, at least to produce an initial short-list (which would them be subjected to qualitative judgment), seems defensible. We are aware that even this suggestion is not without criticism. Again, according to Hicks et al. (2015): "Reading and judging a researcher's work is much more appropriate than relying on one number. Even when comparing large numbers of researchers, an approach that considers more information about an individual's expertise, experience, activities and influence is best". However, in a country without a consolidated research tradition, we think that it may be hard to evaluate research potential without a clear previous experience, and "slow science" will not emerge if researchers are not encouraged to produce more and became productive, in such a way that "quality" may gradually emerge from "quantity" (see Loyola, Diniz-Filho, & Bini, 2012 for a discussion). Nevertheless, it is important to highlight that using A1E as a response variable tries to brings "quality into quantity", despite potential problems due to wide variation in how distinct committees create their Qualis classification.

Scientific production was negatively and positively correlated with undergraduate and graduate teaching load, respectively, when all areas of knowledge are considered together (but this result varies among the areas, with more clear positive effects for teaching in the graduate courses). These antagonistic effects for graduate and undergraduate courses may help explaining why, as recently found by Baccini et al. (2014), the total teaching load does not have a significant effect on scientific production. In the overall model, it is also noticeable that the strength of the (positive) relationship between scientific production and graduate teaching load was higher than the strength of the (negative) relationship between scientific production and undergraduate teaching load. Thus, although a high graduate teaching load may foster scientific production, it is unlikely that low scientific production can only be attributed to a high undergraduate teaching load. The underlying reason for a positive relationship between scientific production and graduate teaching load may be accounted for because usually, but not always, only research-productive faculty members are allowed to teach graduate courses in Brazilian graduation programs (see below). This result can also be substantive in terms of university policy by indicating that more productive faculty members are more engaged in the mentoring and training of future scientists. In addition, it may also indicate the importance of graduate students for the scientific production in Brazil.

Faculty members involved in graduate programs, both in teaching and supervision (especially graduate supervision), have higher A1E (see also Baccini et al., 2014). Probably, this is so because CAPES evaluates all graduate programs around the country, lowering the grade of the graduate programs that enlist researchers with low scientific production. On the other hand, the number of graduate programs in UFG also doubled during the last 3 years due to the new programs created by recently hired researches. This occurred because young faculty member were not immediately incorporated as faculty members in the existing programs due to low scientific production (in comparison to that required by a specific graduate program), political factors within graduate programs or simply because the existing programs were out of their areas of expertise. To successfully implementing the new programs, researchers must have a minimum amount of scientific production to support the program and be certified by CAPES, which is also part of the explanation for the positive relationship between teaching load (and supervision of Master and Doctoral Students) and scientific production.

The positive and significant relationship between scientific production and number of supervised students is a remarkable result. This finding suggests a potential mutual benefit for highly productive researchers and their students. More interestingly, however, the number of undergraduate students is, in general, more correlated with scientific production than the number of master students. If the different areas of knowledge are analyzed separately (see Fig. 3), the importance of undergraduate students is even emphasized because, for some areas, supervision of undergraduates was more important than supervision of master and doctoral students. We believe that this result highlights the importance of Scientific Initiation program and that academic administration should increase incentives for the most productive researchers to increase the number of supervised undergraduate students.

Self-declared time dedicated to research is correlated (although not very high) with scientific production only in the case of funded research. This finding suggest that less productive researchers are using non-funded (i.e. not peer-evaluated) research projects to justify their salaries, but without concrete results. Because the total number of research projects (regardless whether they are funded or not) has been widely used in Brazil as part of the distribution of resources to the universities by the Ministry of Education, our finding has important implications for revising this misleading policy. This policy may have negative implications, as pointed out by Abramo and D'Angelo (2011), who showed the negative impacts of investing in scientific research in a non-competitive and complacent environment.

Gender and outreach activities were, in general, not correlated with scientific production. Despite all discussion about the effect of gender (e.g., Arensbergen, Weijden, & Besselaar, 2012; Abramo, Cicero, & D'Angelo, 2015; Leslie, Cimpian, Meyer, & Freeland, 2015), our analyses (using OLS and A1E as a response variable) support some new findings showing that this effect is slowing down in the recent generations of researchers, as also suggested by Baccini et al. (2014). However, the results from the Hurdle models (see results for the count components for A1E and Papers in Table 2) suggest that gender inequality in terms of scientific production can still be detected (see Larivière, Ni, Gingras, Cronin, & Sugimoto, 2013 for a recent global analysis).

The patterns discussed above varied among areas, especially in terms of the statistical significance of the regression coefficients (but the relative magnitude and sign of the coefficients were, in general, maintained). This similarity between models allowing or not for areas may be due to the fact that the use of A1E already accounts for differences in terms of journal qualification among areas. However, the Qualis system does not allow incorporating different levels of productivity among areas, because although A1 are considered leading journals in these areas, it does not account for the fact that the expected number of A1 published papers to characterize a top researcher would be different. A further investigation could use the expected number of publications to characterize a graduate program as "excellent" and, then, standardize the production of faculty members by this expectation. Unfortunately, these criteria are usually not available and may change quickly through time.

Supervision of undergraduate students (IC), with fellowships granted by the Scientific Initiation program from CNPq, was a significant correlate of scientific production in some areas of knowledge (Agrarian, Humanities, Health, Exact and Earth Sciences). This result highlights again the importance of this relatively low-cost program to increase the scientific production in areas that, at least in Brazil, are not among the most productive.

5. Concluding remarks

Our study is a first step to explain scientific production in a relatively young federal university in Brazil after substantial recent expansion. We used a dataset from a single university as a model system, representing a typical medium-to-large university in Brazil. However, our results should hold for most universities of similar size in the country. There is a wide variation in the levels of scientific production among faculty members, and we suggest at least three main guidelines for improving both scientific production and better management of research activities.

The new faculty members hired by the University tend to have higher production, and the main explanation for this is a good formation (in terms of learning how to publish or being involved in productive research groups). Following the same reasoning, granting paid sabbatical leaves for faculty members with low scientific production may be an inefficient way to improve scientific production in the University and should be evaluated with caution. Finally, it is important to increase the involvement of faculty in graduate programs, especially at Doctoral levels (as long as their scientific production support these activities), improving the publication level of the University.

Acknowledgments

We thank Ludo Waltman and three anonymous reviewers for many suggestions that improved the original manuscript. Our work has been continuously supported by Productivity (PQ) grants from CNPq. This paper originates from UFG Project 42093 and we thank Pedro Cruz and Brasmeire Pereira de Freitas for data assessment and discussions on scientific production in the University.

References

Abramo, G., & D'Angelo, C. A. (2011). National-scale research performance assessment at the individual level. *Scientometrics*, *86*, 347–364.

Abramo, G., D'Angelo, C. A., & Di Costa, F. (2009). Research collaboration and productivity: Is there a correlation? *Higher Education*, *57*, 155–171.

Abramo, G., D'Angelo, C. A., & Di Costa, F. (2011). Research productivity: Are higher academic ranks more productive than lower ones? *Scientometrics*, *88*, 915–928.

Abramo, G., D'Angelo, C. A., & Solazzi, M. (2011). The relationship between scientist's research performance and the degree of internationalization of their research. *Scientometrics*, *86*, 629–643.

Abramo, G., Cicero, T., & D'Angelo, C. A. (2011). Assessing the varying level of impact measurement accuracy as a function of the citation window length. *Journal of Informetrics*, *5*(4), 659–667.

Abramo, G., Cicero, T., & D'Angelo, C. A. (2013a). Individual research performance: A proposal for comparing apples to oranges. *Journal of Informetrics*, *7*(2), 528–539.

Abramo, G., Cicero, T., & D'Angelo, C. A. (2013b). The impact of unproductive and top researchers on overall university research performance. *Journal of Informetrics*, *7*(1), 166–175.

Abramo, G., D'Angelo, C. A., & Rosati, F. (2014). Career advancement and scientific performance in universities. *Scientometrics*, *98*, 891–907.

Abramo, G., Cicero, T., & D'Angelo, C. A. (2015). Should the research performance of scientists be distinguished by gender? *Journal of Informetrics*, *9*(1), 25–38.

Akhmat, G., Zaman, K., Shukui, T., Javed, Y., & Khan, M. M. (2014). Relationship between educational indicators and research outcomes in a panel of top twenty nations: Windows of opportunity. *Journal of Informetrics*, *8*(2), 349–361.

Almeida, E. C. E., & Guimarães, J. A. (2013). Brazil's growing production of scientific articles—How are we doing with review articles and other quality indicators? *Scientometrics*, *97*, 287–315.

Arensbergen, P. V., Weijden, I. V. D., & Besselaar, P. V. D. (2012). Gender differences in scientific productivity: A persisting phenomenon? *Scientometrics*, *93*, 857–868.

Baccini, A., Barabesi, L., Cioni, M., & Pisani, C. (2014). Crossing the hurdle: The determinants of individual scientific performance. *Scientometrics*, *101*, 2035–2062.

Bauer, H. P. W., Schui, G., Eye, A. V., & Krampen, G. (2013). How does scientific success relate to individual and organizational characteristics? A scientometric study of psychology researchers in the German-speaking countries. *Scientometrics*, *94*, 523–539.

Burnham, K. P., & Anderson, D. R. (2012). *Model selection and multi-model inference*. New York, NY: Springer.

Carayol, N., & Matt, M. (2006). Individual and collective determinants of academic scientists' productivity. *Information Economics and Policy*, *18*, 55–72.

Çokgezen, M. (2013). Some determinants of publication performance: A study of Turkish economics academicians. *Journal of Scientometric Research*, *2*, 85–91.

Cruz-Castro, L., & Sanz-Menéndez, L. (2010). Mobility versus job stability: Assessing tenure and productivity outcomes. *Research Policy*, *39*, 27–38.

De Witte, K., & Rogge, N. (2010). To publish or not to publish? On the aggregation and drivers of research performance. *Scientometrics*, *85*, 657–680.

Escobar, H. (2015). Fiscal crisis has Brazilian scientists scrambling. *Science*, *349*, 909–910.

Franceschet, M., & Costantini, A. (2011). The first Italian research assessment exercise: A bibliometric perspective. *Journal of Informetrics*, *5*(2), 275–291.

Gonzalez-Brambila, C., & Veloso, F. M. (2007). The determinants of research output and impact: A study of Mexican researchers. *Research Policy*, *36*, 1035–1051.

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature*, *520*, 429–431.

Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. (2013). Bibliometrics: Global gender disparities in science. *Nature*, *204*, 211–213.

Leslie, S. J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, *347*(6219), 262–265.

Loyola, R. D., Diniz-Filho, J. A. F., & Bini, L. M. (2012). Obsession with quantity: A view from the south. *Trends in Ecology and Evolution*, *27*(11), 585.

McCarty, C., Jawitz, J. W., Hopkins, A., & Goldman, A. (2013). Predicting author h-index using characteristics of the co-author network. *Scientometrics*, *96*, 467–483.

McNutt, M. (2014). The measure of research merit. *Science*, *346*, 1155.

Miller, J. C., Coble, K. H., & Lusk, J. L. (2013). Evaluating top faculty researchers and the incentives that motivate them. *Scientometrics*, *97*, 519–533.

Myers, D. (2011). Selling to the BRIC: Brazil—An awakening tiger. *Learned Publishing*, *24*(4), 261–265.

Pachi, C. G. F., Yamamoto, J. F., Costa, A. P. A., & Lopez, L. F. (2012). Relationship between connectivity and academic productivity. *Scientometrics*, *93*, 265–278.

Perianes-Rodríguez, A., Chinchilla-Rodríguez, Z., Vargas-Quesada, B., Gómez, C. O., & Moya-Anegón, F. (2009). Synthetic hybrid indicators based on scientific collaboration to quantify and evaluate individual research results. *Journal of Informetrics*, *3*(2), 91–101.

R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing., lang;http://www.R-project.org/⟩

Regalado, A. (2010). Brazilian Science: Riding a gusher. *Science*, *330*, 1306–1312.

Rice, P., Venables, A. J., & Patacchini, E. (2006). Spatial determinants of productivity: Analysis for the regions of Great Britain. *Regional Science and Urban Economics*, *36*, 727–752.

Ruocco, G., & Daraio, C. (2013). An empirical approach to compare the performance of heterogeneous academic fields. *Scientometrics*, *97*, 601–625.

Silaghi-Dumitrescu, R., & Sabau, A. (2014). Scientometric analysis of relative performance in a key university of Romania. *Scientometrics*, *99*, 463–474.

Torrisi, B. (2013). Academic productivity correlated with well-being at work. *Scientometrics*, *94*, 801–815.

Vieira, E. S., & Gomes, J. A. (2010). A research impact indicator for institutions. *Journal of Informetrics*, *4*(4), 581–590.

Yu, T., Yu, G., Li, P.-Y., & Wang, L. (2014). Citation impact prediction for scientific papers using stepwise regression analysis. *Scientometrics*, *101*, 1233–1252.

Zeileis, A., Kleiber, C., & Jackman, S. (2008). Regression models for count data in R. *Journal of Statistical Software*, *27*, 1–25.

Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., & Smith, G. M. (2009). *Mixed effects models and extensions in ecology with R*. New York, NY: Springer.