



The Brazilian scientific output published in journals: A study based on a large CV database



Marcelo S. Perlin^{a,*}, André A.P. Santos^c, Takeyoshi Imasato^a,
Denis Borenstein^{a,b}, Sergio Da Silva^c

^a School of Business Administration, Federal University of Rio Grande do Sul, Porto Alegre, R.S., 90010-460, Brazil

^b Prometeo Researcher, Facultad de Ciencias Económicas y Administrativas, University of Cuenca, Ecuador

^c Graduate Program in Economics, Federal University of Santa Catarina, Florianopolis, S.C., 88049-970, Brazil

ARTICLE INFO

Article history:

Received 30 May 2016

Received in revised form 27 October 2016

Accepted 27 October 2016

Available online 10 November 2016

Keywords:

Lattes platform

Scholarly publishing

Scientific productivity

Brazilian researchers

ABSTRACT

We assemble a massive sample of 180,000 CVs of Brazilian academic researchers of all disciplines from the *Lattes* platform. From the CVs we gather information on key variables related to the researchers and their publications. We find males are more productive in terms of quantity of publications, but the effect of gender in terms of research impact is mixed for individual groups of subject areas. For all fields of science, holding a PhD from abroad increases the chance for a researcher to publish in journals of higher impact. We also find that the more years a researcher takes to finish his or her doctorate, the more likely he or she will publish less thereafter, although in outlets of higher impact. The data also support the existence of an inverted U-shaped function relating research age and productivity.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

To keeping score of the scientific output of individual researchers and institutions, the Brazilian National Research Council (*CNPq*) implemented an individual CV database countrywide, integrated with a web-based query interface. This is called the *Lattes* platform (<http://lattes.cnpq.br/>). All researchers and institutions are incentivized to keep their records up to date through *Lattes*, recording all the main academic activities and publications. For this reason, *Lattes* can be used to gather good-quality information on individual researchers, as well as to help conduct the performance evaluation of academic and research institutions. Here, we take advantage of the massive database provided by *Lattes* and assemble a sample of 180,000 CVs. Our main interest is to collect the particular information conveyed in the CVs regarding the journal publication destination of Brazilian researchers.

Research activity usually involves a large amount of financial resources. World Bank data (<http://data.worldbank.org/>) show that Brazilian expenditures on research accounted for 1.2 percent of GDP from 2005 to 2012, which translates to approximately \$27 billion at current values.

Scholarly publishing is highly concentrated among a few authors (Lotka, 1926), and scientific research productivity is asymmetrically distributed (Chung & Cox, 1990; Ruiz-Castillo & Costas, 2014). Although what became known as Lotka's law has been disputed recently (Coile, 1977), there is a general agreement in literature that a relatively small proportion of scientists contribute most. Thus, "success breeds success" (Rorstad & Aksnes, 2015; Ruiz-Castillo & Costas, 2014).

* Corresponding author.

E-mail address: marcelo.perlin@ufrgs.br (M.S. Perlin).

Having this in mind, proper identification of the determinants of research productivity in different areas of knowledge can be not only useful for a better allocation of financial resources but also to improve the evaluation of researchers and academic institutions, providing new insights about the factors that directly explain the asymmetric productivity in one particular area and across areas.

Researcher productivity has been examined widely. Previous work points to key factors explaining it. Gender has been regarded as a factor that can explain research productivity asymmetry (Isfandyari-Moghaddam, Hasanzadeh, & Ghayoori, 2012; Leta & Lewison, 2003; Morganson, Jones, & Major, 2010; Prpi, 2002; Puuska, 2009). In one study, female research productivity accounted for only 70.6 percent of male productivity (Prpi, 2002). This result is reinforced as one looks at a sample of associate and full professors, a group heavily skewed toward males (Abramo, Cicero, & Dangelo, 2015; Barrios, Villarroja, & Borrego, 2013; Larivire, Vignola-Gagn, Villeneuve, Glinas, & Gingras, 2011; Rorstad & Aksnes, 2015). Moreover, the scant number of female Nobel laureates has less chance to marry and have children as compared to both their male counterparts and the general population (Charyton, Elliott, Rahman, Woodard, & DeDios, 2011). This posits a trade-off between career success and family life for female researchers, though this trade-off seems to be lessening for the current generation (van Arensbergen, van der Weijden, & van den Besselaar, 2012).

The effect of holding a PhD from a foreign university has also been considered in the literature regarding the Brazilian experience. Roos et al. (2014) show strong evidence that Brazilian researchers in the biological sciences with doctorate degrees from abroad are also those who publish less, though in journals of greater impact. Here, the student-supervisor relationship seems to play a role (Tuesta et al., 2015). Indeed, there is significant correlation between future productivity of a PhD candidate and the time spent in his or her interaction with the supervisor.

The previous literature also shows evidence of a non-linear impact of a researcher age on his or her productivity. This suggests the existence of an inverted U-shaped function. Soon after PhD completion, productivity rises, reaches a peak at a certain age, and then declines (Bonaccorsi & Daraio, 2003; Rorstad & Aksnes, 2015).

The main objective of this study is to provide an overview of the Brazilian scientific output of different subject areas. By using statistical models, we identify the factors that explain researcher productivity in terms of quantity and impact across the different areas. We replicate some findings from literature, contradict others, and further present new findings and insights based on a much larger sample. The massive database allows us to adopt not only a time series approach, but also a cross-section perspective. These make it possible to consider the evolution of the Brazilian research output from almost all subject areas.

Assessing the factors affecting research output using statistical models is not new. This perspective is already present in the studies of Roos et al. (2014), Rorstad and Aksnes (2015), and Diniz-Filho, Fioravanti, Bini, and Rangel (2016). However, our study is more comprehensive in that it considers a much larger database of researchers from different subject areas in higher education institutions. Roos et al. (2014) consider the same database as ours (Lattes), but their study covers only the biological sciences. Rorstad & Aksnes (2015) consider the Norwegian academic staff, which is ten times smaller than the Brazilian one, and Diniz-Filho et al. (2016) analyze data from Lattes for only one Brazilian university. By exploring this large dataset of heterogeneous areas, we contribute to the mapping of the profile of the Brazilian academic output, and also to identify the factors explaining volume and impact of research within each subject area.

The next section provides a description of the dataset and methods employed in this study. Section 3 presents and discusses the results. Section 4 concludes.

2. Materials and methods

2.1. Collecting the data

The Lattes platform was launched in August 1999. It replaced the prior paper-based and non-integrated electronic system, which conveyed information from individual academic researchers. In the subsequent two years, the number of CVs posted under Lattes grew more than 300 percent, from 35,000 CVs to more than 100,000. In 2010, Lattes provided high quality data on about 1.6 million researchers and about 4000 institutions (Lane, 2010).

An open, sound and consistent platform for measuring all the activities that make up academic productivity is a worldwide-acknowledged need. In this regard, a standardized platform such as Lattes is an asset because “metrics are data driven, so developing a reliable, joined-up infrastructure is a necessary first step” (Lane, 2010). Lattes is now internationally perceived as a powerful example of good academic practice.

Our study departs from others that use information from Lattes (such as Almeida & Guimarães, 2013; Leite, Mugnaini, & Leta, 2011) in at least one important aspect. We develop a software for mining big-scale data that renders us a sample significantly larger than any other previously considered in Lattes studies. The software, labeled *Nilrep-Lattes*, is written in Python and registered at the Brazilian National Institute of Industrial Property (code number BR512014000516-0). *Nilrep-Lattes* takes as an input a list of CVs from Lattes, such as <http://lattes.cnpq.br/x>, where x is a numeric value of 16 digits that is unique to each researcher. The software then collects information for every researcher, such as name, PhD origin, PhD completion time, years since PhD completion, current job address, and list of published papers, including publication year and journal ISSN.

Lattes is mostly used by academics, from undergrad students to full professors. While it provides a rich dataset, it also has shortcomings. The user supplies all information and updates are not automatic. Thus, some users may feed wrong

Download English Version:

<https://daneshyari.com/en/article/4968106>

Download Persian Version:

<https://daneshyari.com/article/4968106>

[Daneshyari.com](https://daneshyari.com)