



# The Role of Early-Career University Prestige Stratification on the Future Academic Performance of Scholars

Mario González-Sauri<sup>1</sup> · Giulia Rossello<sup>1,2</sup>

Received: 19 December 2020 / Accepted: 25 January 2022 / Published online: 30 April 2022  
© The Author(s) 2022

## Abstract

This paper investigates the effect of university prestige stratification on scholars' career achievements. We focus on 766 STEM PhD graduates hired by Mexican universities between 1992 and 2016. We rank university according to their prestige based on the pairwise assessment of quality contained in the PhD hiring networks. Further, we use a quasi-experimental design matching pairs of individuals with the same characteristics, PhD training or first job experience. Our results challenge the positive association between prestige and academic performance as predicted by the 'Matthew effect'. Scholars hired internally sustain higher performance over their careers in comparison to those who move up or down the prestige hierarchy. Further, we find a positive (negative) relation between downward (upward) prestige mobility and performance that relates to the "big-fish-little-pond" effect (BFLPE). The evidence of a BFLPE-like effect has policy implications because hinders the knowledge flows throughout the science system and individual achievements.

**Keywords** University prestige · Academic inbreeding · Early-career researchers · Faculty hiring network · Academic labor market · Developing countries

## Introduction

For universities, forming and hiring PhD graduates is essential for their competitiveness as young scholars will become lecturers, senior researchers and professors in their faculties. However, for early-career scholars, the transition from PhD and first job is competitive, stressful and can have consequences for their future careers (Bazeley, 2003). At the same time, the academic labor market is well-known for exhibiting a Matthew effect, where

---

Financial support was provided through The National Council on Science and Technology (CONACYT) of Mexico. We gratefully acknowledge the comments and suggestions of Prof. Paula Stephan, Prof. Robin Cowan, Dr. Neil Foster-McGregor, and Beatriz Calzada Olivera.

---

✉ Mario González-Sauri  
gonzalez@merit.unu.edu

<sup>1</sup> UNU-MERIT, Maastricht University, Boschstraat 24, 6211AX Maastricht, The Netherlands

<sup>2</sup> Institute of Economics, Sant'Anna School of Advanced Studies, Martiri della libertà 33, 56127 Pisa, Italy

prestigious positions in hierarchical networks give advantages to early-career scholars across their careers (Bol et al., 2018; Teplitskiy et al., 2020; Horta et al., 2018).

The Matthew effect predicts that ‘early career’ gains in prestige confer researchers advantages that over time render higher research performance. The mechanism behind this success is a process of cumulative advantage in which early career prestige attracts resources (Merton, 1968; Long, 1978; Merton, 1988), increases visibility (de Solla Price, 1965; Wang, 2014; Farys and Wolbring, 2021) and collaborators (Perc, 2014). This positive feedback loop reinforces the prestige of the authors (as well as universities) and leads to higher research productivity (Allison and Stewart, 1974). This thesis has been tested extensively from sociology, economics, and studies of science and technology, presenting mixed results. One group of studies asserts that university prestige is positively associated with individual performance (Fox, 1983; Headworth and Freese, 2016; Su, 2011). For instance, the study of Bedeian and Feild (1980) concludes that institutional prestige from the PhD is more important than publication productivity prior to the first job appointment in the academic labor market. However, when the endogeneity of university prestige is accounted, the literature shows mixed results (Allison and Long, 1987; Williamson and Cable, 2003; Bair, 2003; Miller et al., 2005; Laurance et al., 2013; Appelt et al., 2015), leaving several gaps that are areas of opportunity for the present study.

The first gap that we address is the potential selection bias in the estimation of prestige and research performance. Most studies, disregard, that researchers are sorted into prestigious appointments by self-selection. To address this problem, we take advantage of a quasi-experimental design. Our treatment is the change in prestige that naturally occurs to scholars in the transition between PhD graduation and first job appointment. This change in prestige is a partially random assignment, given the exogenous choice of hiring committees. This shock allocates scholars into three groups depending on their change in prestige (treatment condition). One group exhibit a positive change or upward prestige mobility when moves from a less prestigious PhD university to a more prestigious first job appointment. Conversely, a second group displays a negative change in prestige (downward mobility) and the last group experience no change in prestige as this group is hired by their faculty.<sup>1</sup> We address the selection bias, by matching early career researchers in these three groups keeping constant their individual characteristics, PhD training or first job experience. Fixing these determinants of research performance allows to observe the effect of early-prestige in the short, medium and long-run career.

A second gap in the literature is that results are typically bias towards top-tier institutions located mostly in North America and Western Europe (Clauset et al., 2015). Those universities usually have a long history and are well known and integrated internationally (Demeter and Toth, 2020). In contrast, less mature higher education systems are younger, less consolidated and operate under resource constraints. Therefore, this paper examines the subject in a large emerging economy, namely Mexico.

More importantly, the prestige of academic institutions in those contexts is less known and can be more volatile compared to more mature university systems. In developing university systems, prestige may change according to modes, tastes and new information. For example, reputation can decrease if frauds and misbehavior are found within a department. In the same way, if a department wins a large National grant, its reputation will improve. Similarly, the academic mobility of star scholars may change substantially the prestige and

<sup>1</sup> Sometimes called in the literature Academic Inbreeding (Horta et al., 2010).

output of an entire department. Indeed, most of the US-based literature on prestige and hiring has found at the departmental level, that the correlation between hiring centrality and survey based measures of prestige is higher than the correlation between bibliometric production and survey based measures of prestige (Clauset et al., 2015; Burris, 2004). Additionally, in the case of an emerging economy many institutions are not listed in formal university rankings<sup>2</sup> and non-English literature is not present in bibliometric databases. Thus, all the mentioned reasons make the use of survey-based or bibliometric-based measures not only unsuitable for measuring prestige in this context but also practically harder. A third gap to address is then to develop a measurement of prestige suitable for less developed university systems, robust against prestige variation over time, taking into account the job market dynamics, (see Oyer 2006).

To overcome the aforementioned problem of measurement error, we estimate the institutional prestige using a ranking algorithm based on university hiring networks of PhDs. The university hiring network has been extensively studied and contains explicit hiring flows from one university to the other (Barnett et al., 2010; Clauset et al., 2015; Lang et al., 2019). Past research on these flows indicates that university hiring networks contains an implicit hierarchical order of prestige among institutions (Barnett et al., 2010; Mai et al., 2015). This hierarchy of prestige comes into existence because hiring decisions in the academic labor market are pairwise evaluations of “quality” between candidates and universities. We use the information of each pairwise assessment of quality to measure university prestige, applying the algorithm developed in our previous study (Cowan and Rossello, 2018). We argue that this measurement of prestige based on a pairwise peer review assessments goes beyond university ranks based on bibliometric indicators and subjective surveys. In particular, in this paper, we propose a dynamic estimation of the algorithm that overcomes the potential volatility of prestige in less mature higher education systems.

Using our dynamic measurement of prestige, this paper asks how university prestige stratification affects academic performance and the labor market mobility of scholars in Mexico. We resolve the endogeneity of university prestige on academic performance from several sources. To overcome the problem of reverse causality, we assess how early-career institutional prestige affects future academic performance of scholars with longitudinal data. In line with the literature, changes in prestige during early-career affect future research performance, given that initial conditions in academia confers cumulative advantages/disadvantages with long-run consequences (Bazeley, 2003; Bol et al., 2018; Lee, 2019). To overcome the problem of bias from omitted variables correlated with individual capabilities, we take advantage of a quasi-experimental design and a re-sampling technique. For our three treatment groups (Upward, Downward, or Unchanged prestige change) we control for the PhD training, the postdoctoral experience and the individual characteristics because these variables also affect academic performance. Our re-sampling technique matches pairs of scholars with similar characteristics assigned to different treatments. Firstly, we compare pairs with similar characteristics that receive the training and prestige from the same university and graduate window. Secondly, we compare pairs with equivalent characteristics, similar networks, experience and prestige from their first job appointment. By holding constant these initial conditions for similar individuals allows us to compare the effect of changing prestige (treatment) on future research performance.

---

<sup>2</sup> As for example Times Higher Education and Shanghai University Ranking.

Our results are based on a sample of 766 Mexican PhD graduates in STEM between 1992 and 2016 and show a highly stratified university system. At the macro level, a selected group of 10 institutions graduate and hire the vast majority of early career scholars. This stratification is consistent with more mature university systems (Burris, 2004; Barnett et al., 2010; Clauset et al., 2015). This centralization of research could promote higher levels of specialization and targeted allocation of resources. But also can reveal a structural problem of “lock-in”, hindering mobility and the flows of knowledge through the national science system.

Focusing on the individual level, our results reveal a more nuanced and complex pattern than the one predicted by the Matthew effect. We find that moving up the prestige ladder does not necessarily correlate with a higher academic performance. Controlling for the PhD training and the experience gained from their first job, our results confirm that scholars (with similar characteristics) hired internally exhibit on average higher levels of academic performance. Interestingly, early-career scholars with similar characteristic and initial conditions who experience downward prestige mobility perform on average better than scholars who exhibit upward prestige mobility. We explain our results using the analogy from the big-fish-little-pond effect (BFLPE), where individual performance relates to the average performance of her/his peers. The effect that prestige has on individual performance may relate to the “profile of prestige” an individual is accustomed to. Agents might perform “better” (“worse”) in an environment with a “lower” (“higher”) competition where the perceived average ability of peers is “lower” (“higher”). Testing these psychological mechanisms is beyond the scope of this work, but we highlight a non-linear association between prestige and performance that can be conducive of future research.

## Prestige, Mobility and Performance

### University Prestige

The main approaches for ranking universities by prestige are input-output (Debackere and Rappa, 1995; Chan et al., 2002; Kalaitzidakis et al., 2003; Oyer, 2008; Buela-Casal et al., 2012), survey (Abbott and Barlow, 1972; Cyrenne and Grant, 2009; Moodie, 2009; Olcay and Bulu, 2017), and network-based measures (Barnett et al., 2010; Cowan, and Rossello 2018; Zhu and Yan 2017; Nevin 2019). Some of the most popular input-output based measures use bibliometric indicators. For instance, Debackere and Rappa (1995) use bibliographic records to calculate the prestige rankings for universities in the top-twenty departments of neuroscience by adding their citations. Similarly, Oyer (2008) uses the methodology proposed by Kalaitzidakis et al. (2003) to rank universities based on the contribution of universities to the top thirty journals in economics selected by their normalized citation index. Chan et al. (2002) take a similar approach and use the number of pages produced by a university faculty in top journals to rank universities. However, bibliometric based measures focus on research outputs, disregarding inputs such as graduates production and other relevant measurements of academic excellence. Buela-Casal et al. (2012) assess the higher education system in Spain using a mix of input-output indicators including journal publications, number of full-time researchers, number of R&D projects, PhD graduates, scholarships and patents.

Another way to rank universities is through surveys. This method aims at capturing the peers’ perception of academic prestige that includes niches within a discipline. Abbott and

Barlow (1972) use a survey of graduate faculty to rank universities in 29 disciplines, with an ordinal response scale of five levels. For Fogarty and Saftner (1993) the main drawbacks of survey methods are that they assume that scholars are unbiased of their current and previous affiliations. An additional issue with survey measures is that scholars tend to have sticky and localized information about other universities—they know better institutions of similar status that are those which are in direct competition with them (Cowan and Rossello, 2018). For instance, institutions that are closer in rank are likely to compete for the same resources (projects and grants) and publish articles in similarly ranked academic journals.

Network-based measures are a new framework to rank universities that examine a social process in which scholars recognize quality in their work. This evaluation is cross-validated by the interactions between institutions. Following these lines, the core idea of our ranking approach is that the PhD hiring networks contain information about how scholars evaluate each-other quality (Clauzet et al., 2015). Finally, the methodology of Cowan and Rossello (2018), applied in the university system of South Africa, exploits the information contained in the movements in the labor market to approximate the distribution of prestige. The advantage of their method is that it directly uses pairwise assessments of quality between PhD graduates and hiring committees to rank universities. As a further expansion of their methodology, our ranking algorithm is dynamic across time. Thus, we consider the job market of variable sizes that consequentially changes the prestige of the institutions involved. A dynamic perspective to rank universities is more suited in a large, developing setting where universities are young and the distribution of prestige is potentially more volatile.

## The Link Between Mobility and Performance

The literature studies the relationship between university prestige and mobility to understand how the prestige of the PhD-granting institution affects the placement and subsequent labor market outcomes. In the PhD job market, both individuals and universities have incentives to make accurate choices. On the one hand, for the job market candidate, the transition from the PhD and to her/his first academic job might affect future career perspectives (Laudel and Gläser, 2008). On the other hand, for the university, each new hire is a strategic asset that influences its human capital and competitiveness (Cowan and Rossello, 2018). However, hiring PhD graduates could be a difficult task, since early-career scientists often have few research records. In such a case, universities have little information about the inherent ability of young candidates, especially when they graduated in other universities.

Related to this, Oyer (2006) examines the job market for economists in the US 1979–2004 and finds that the job market conditions, that influence the first job placement of scholars, contains a large element of randomness. In particular, he finds that the initial job market ‘luck’ (i.e. favorable market conditions) affects top positions in academia that later drive research productivity between neighboring cohorts of graduates. The role of job market luck in affecting placement operates due to the asymmetry of information. The latter makes the prestige of the PhD-granting institution a signal for the unobserved skills and abilities of applicants. Indeed, past literature shows a positive relationship between the prestige of the PhD granting institution and future employment (Crane, 1970; Debackere and Rappa, 1995; Oyer, 2006; Bedeian et al., 2010; Appelt et al., 2015; Pinheiro et al., 2017; Headworth and Freese, 2016). PhD graduates from prestigious institutions tend to

get “better” job compared to graduates from lower-tier universities. Similar results suggest that a prestigious PhD is as a key mechanism to alleviate the asymmetry of information between candidates and hiring faculties. Moreover, since a “good” affiliation provides opportunities, networks, and resources, this first sorting of the job market has potential long-run consequences on career achievements (Oyer, 2008; Bedeian et al., 2010). A similar early-career advantage is potentially problematic when quality and prestige do not entirely overlap. In this respect, some studies argue that institutional prestige from the PhD granting institution is more important than researchers “quality” for obtaining a first academic job (Long et al., 1979; Allison and Long, 1990; Baldi, 1995; Gerhards et al., 2018).

Another group of studies pays attention to how changes of institutional prestige relate to academic performance and academic achievement of scholars. Oyer (2008) use a longitudinal sample to estimate how changes in institutional prestige affect the academic performance of economists. He shows that, even after controlling for proxies of individual-level ability, that early academic prestige positively correlates with academic performance measured by publication productivity. Moreover, he shows that scholars generally move down the prestige ranking over their careers. He argues that this is because high ranked universities produce a significant percentage of the total graduates that later move to lower-ranked universities. Chan et al. (2002) examine the mobility of scholars publishing in 16 top financial journals. They find that upward ranking mobility is rare and that scholars who experienced it produce twice as many publications compared with average production of scholars from destination universities. They further show that after controlling for ability using publication productivity, the rank of the PhD grading institution predicts upward ranking mobility through their academic careers. Azoulay et al. (2014) take an alternative approach, comparing academic performance of scholars before and after upward mobility given by a prestigious academic recognition. They find that gains from upward ranking mobility have a lower effect on scholars who have above average citations than on scholars with low or below average citations. In general, these studies suggest that upward ranking mobility is associated with higher academic performance, but this is not always the case. Cowan and Rossello (2018) offer a closer look at how prestige differentials from the PhD to the first job relate to academic performance. They use a quasi-experimental methodology based on matching pairs and examine a sample of 1011 South African job market candidates in STEM. They show that scholars hired internally (maintaining their place in the prestige hierarchy) exhibit on average higher performance compared with scholars who move up in the prestige rank. Their work underlines that the link between prestige changes and performance might be complex.

## The BFLPE and Similar Mechanisms

The complex relation between prestige changes and performance can relate with the relation that individuals have with their new peers and the working environment. When individuals change institution they compare themselves with new peers, therefore, their view of themselves may be odd because they have little information on the new environment.

Psychology of education examines how the social comparison affects individual performance by looking at how the average achievement of peers affects the individual academic self-concept (Marsh and Hau, 2003; Marsh et al., 2008). The main hypothesis in this literature is the big-fish-little-pond effect (BFLPE). The hypothesis suggests that a student will have a lower academic self-concept (and thus performance) in an academically selective school, where the average achievement of peers is high, than in a non-selective one (Astin,

1969; Marsh and Hau, 2003; Marsh et al., 2008; Salchegger, 2016; Rosman et al., 2020; Keyserlingk et al., 2020).

The empirical evidence on the BFLPE shows mixed results and focuses on children or adolescents at school age (Salchegger, 2016). Two recent contributions test the BFLPE looking at first-year university students. Rosman et al. (2020) find no support for the BFLPE examining 115 first-year undergraduate psychology students at the Leibniz Institute in Germany. In contrast, in a larger and more representative study, Keyserlingk et al. (2020) find strong support for the BFLPE in a sample of German students in the transition from high school to universities. However, in higher education, competition and the need of collaborating with peers is higher. In similar circumstances, additional mechanism of social comparison might influence beliefs and achievements. The additional mechanisms in the literature are peer effects and the *what does not kill me makes me stronger* effect.

In general terms, the literature on peer effects in academia studies whether the social comparison generates learning that in turn affects performance. Most of this literature focuses on positive peer effects. For example, Slavova et al. (2016) study whether hiring a new scientist affects the scientific performance of the incumbents in the hiring department. They examine 94 U.S. chemical engineering departments, finding that a new hire generates positive peer effects in the performance of colleagues with a recent tenure. As a limitation, their analysis does not consider which is the impact of changing department for the newcomer. Related to this, past research suggests that, depending on the level of the competition, peer effect can also be negative (Stapel and Koomen, 2005). When resources are scarce, the level of competition to access them increases and in extreme cases a newcomer can be perceived by the group as a treat. In this case, the effect of the social comparison may negatively affect the performance. An extreme case of thereof are bullying episodes or misbehavior in academia (McKay et al., 2008; Keashly and Neuman, 2010; Giorgi, 2012).<sup>3</sup>

The way in which people are integrated or promoted in a new workplace can affect individual self-esteem and academic self-concept as well. In general, the literature associates high self-esteem and self-concept to high career outcomes. However, there are cases where this association appears to be negative (Whelpley and McDaniel, 2016; Sherf and Morrison, 2019; Li et al., 2020; Weiss and Knight, 1980).

For example, in a recent contribution, Wang et al. (2019) find support for the *what does not kill me makes me stronger* effect. They compare publication and citation records of 561 narrow wins and 623 near miss scientists who applied for the NIH grant<sup>4</sup> They find that despite an early setback, individuals with near miss proposal systematically outperform those with narrow wins in the longer run. This result could be consistent with a BFLPE in early career scientists, where an initial promotion or confirmation of abilities may be counter-productive for their future performance. The complex mechanisms associating individual performance and the comparison with peers motivates us to study how prestige changes affect research performance in the transition into the first academic job.

<sup>3</sup> See Henning et al. (2017) for a review on the topic. However, most of this research is limited to voluntary surveys and small samples.

<sup>4</sup> NIH (National Institutes of Health) is the largest public funder of biomedical research in the world.



## Data

Data originates from the *Mexican National Council of Science and Technology* (CONACYT). Data were collected through the most extensive science policy of the country, the *National System of Researchers* (NSR), whose aim is to increase the productivity, quality and competitiveness of Mexican researchers (Gras, 2018). NSR was implemented in 1985 when the primary motivation behind the policy was the raising concern about technological capabilities and performance of the Mexican science system under the threat of inflation and budget cuts. Reyes and Suriñach (2015) describe how the policy evolved across the years, but in general, its structure is substantially unchanged.

We focus on STEM graduates, excluding from the analysis those in Social Sciences and Humanities to reduce the potential influence that schools of thought have on the PhD job market.<sup>5</sup> Our sample spans 25 years representing 766 PhD job market candidates hired in a Mexican university between 1992 and 2016. We present the summary statistics of the panel in Table 1. We include 36 Mexican institutions<sup>67</sup> and longitudinal records for each scholar of academic performance (*NSR rating*) and individual level controls as *gender*, *discipline*, *graduation year* and *evaluation year*.

## NSR Rating

Our dependent variable is the *NSR rating*, which measures the academic performance of researchers using 5 ordered categories. The general framework of the NSR rating process is summarized in Fig. 1 and works as follows. Each researcher applies to the NSR submitting her/his curriculum vitae and publications. The submitted publications comprise not only scientific articles, but also books, chapters in books, patents, and technological developments and transfers. Each application is assigned to one of seven different research disciplines. Every three years, for each discipline, the NSR forms the evaluation commission which comprises 14 prominent researchers called to rate the applications. The evaluation commission works as follows. Each member of the commission evaluates the performance of the applicants evaluating all the submitted material, this following a peer-review process that ends with a grade. The CONACYT authorities supervise the quality and independence of the evaluation commissions (more details are in “[NSR Disciplines and Evaluation Procedure](#)”).<sup>8</sup> In contrast to bibliometric measures, a peer-review evaluation has the advantage of including a holistic evaluation taking into account the validation practices of each discipline within the country. In particular, a similar assessment of research performance considers seniority, the quality of publications, the individual contribution to co-authored works, and above all (sub-)field differences. The evaluation process ends with a rating that

<sup>5</sup> CONACYT Research Areas included are I, II, III, VI and VII. Respectively, Physics-Mathematics and Earth Sciences, Biology Chemistry and Life Sciences, Medicine and Health Sciences, Biotechnology and Agricultural Sciences, and Engineering. See Appendix “[NSR Disciplines and Evaluation Procedure](#)”.

<sup>6</sup> The list of institutions is presented in the Appendix “[A. Faculty Hiring Matrix Names](#)”.

<sup>7</sup> The sample comprises the 36 Mexican universities who provide doctoral education in STEM subjects. In Mexico, universities specialized in teaching do not grant doctorates. In Mexico, there are 1250 institutions of Higher Education, including public universities, technological institutes, technological universities, private institutions, teacher training colleges, and other public institutions. Among those, universities are 213. However, 50% of the research and 58% of the students concentrates in only 45 public universities which are mostly located near Mexico City and in other large cities.

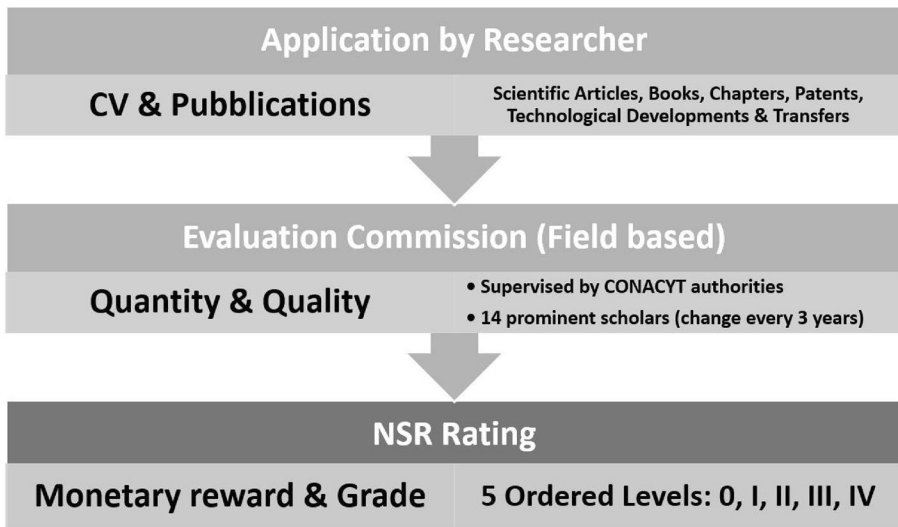
<sup>8</sup> For more details on the NSR system see (Gras, 2018).



**Table 1** Summary statistics, the panel of 766 researchers (1992–2016)

	Gender	Age	Pr-job	Pr-PhD	$\Delta Pr$	$\Delta Pr^*$	NSR rating
Mean	0.47	44.13	22.73	15.79	1.62	- 6.94	2.42
SD	0.50	5.40	14.68	12.05	0.63	13.32	0.86
Median	0.00	44.00	20.20	11.89	2.00	0.00	3.00
Min	1.00	31.00	1.29	1.29	1.00	- 62.31	1.00
Max	0.00	75.00	75.08	75.08	3.00	65.24	5.00

*Gender* 0 is for male researchers, *Pr-job* is the prestige rank of the university of researchers' first job. The lower the score, the higher the prestige, *Pr-PhD* is the prestige rank of the university of researchers' PhD. The lower the score, the higher the prestige,  $\Delta Pr$  is the change in prestige from PhD to first job. The Down group is equal to one, the group who Stayed is two, and the Up group is three.  $\Delta Pr^*$  is the continuous change of prestige. The negative means signify that, on average, scholars move down. That is because the lower the rank, the higher the prestige. *NSR rating* is the dependent variable that scores research performance from one to five (the best performance)

**Fig. 1** NRS Rating system. Details of the evaluation procedure

systematizes the academic performance of researchers in 5 ordered categories. In the paper, we use those categories to measure the academic performance of individuals.

## Interactive Prestige Ranking

In this section, we describe the variation of the prestige ranking algorithm developed by Cowan and Rossello (2018) for a dynamic setting. The algorithm assumes that that movements in the academic job market contain information about how universities and PhD graduates perceive each other's quality. The input of the algorithm is the hiring network defined as  $G = (V, E)$ . The vertices  $V$  of the network are the universities participating in

the job market and edges  $E$  represent movements of PhD graduates from one university to another. The network  $G$  is represented by a weighted directed adjacency matrix  $A$  that captures the flows of graduates from PhD to their first job institution.  $A$ 's off diagonal elements ( $a_{ij}$  with  $i \neq j$ ) show the number of scholars that graduated from university  $i$  and were hired by university  $j$  within 5 years after doctoral graduation. Conversely, the diagonal elements ( $a_{ii}$ ) are the PhD graduates hired internally (trained and hired by their faculty).

The ranking algorithm is based on two key assumptions. The first concerns universities, that is, they try to improve their status and quality and in pursuing it they try to hire from universities "better" than themselves. The second assumption considers scholars, that is, they want to be hired by the most prominent institution. If both academics and institutions satisfy those desires fully, in  $A$  exists a unique order of university names (rows/columns names) such that PhD graduates only move down the hierarchy. In other words, under this assumption, rows and columns of the adjacency matrix  $A$  can be rearranged in an *upper triangular matrix* such that all entries below the diagonal are equal to zero. We define this unique order  $o^*$ , where the sum of rows have a global maximum score equal to  $s^*$ .

$$s^* = \max \sum_i \sum_{j>i} A_{ij} \tag{1}$$

Geographic location and other recruiting criteria imply that the PhD job market often departs from this strict assumptions. Thus, empirically the order is not unique, and  $A$  it is not a perfect *upper triangular matrix*. However, since prestige is an important selection criteria for university and scholars, we apply the heuristic algorithm proposed in Cowan and Rossello (2018) to find the set of orders that gets as closer as possible to  $s^*$  and have the minimum number of violations from an upper triangular matrix configuration. To approach the underlying prestige hierarchy, the algorithm works as follows. The algorithm for  $k = 10000$  times starts assigning to  $A$  a random order of rows, then it computes the score  $s_k$  such that

$$s_k = \sum_i \sum_{j>i} A_{ij} \tag{2}$$

For 100 times the algorithm tries to improve the score  $s_k$  in the following way. For each iteration, two nodes (both rows and columns) are randomly selected and swapped. If the swap does not decrease the score, we keep it, otherwise we reject it. After this 100 searches, the obtained order  $o_k$  and score  $s_k$  are recorded, obtaining a set of  $n$ -tuple  $O = \{o_1, o_2, \dots, o_k\}$  orders and their associated scores  $S = \{s_1, s_2, \dots, s_k\}$ .

From these two sets  $O$  and  $S$ , the algorithm selects the set  $Q = \{o_m \in O \mid s_m \approx s^*\}$  of orders that reach the highest score. Each  $n$ -tuple of the set  $Q = \{q_1, \dots, q_m\}$  contains a possible university rank  $R(v) = \{r_1, r_2, \dots, r_m\}$  Then for each university the algorithm computes its prestige score according to the formula

$$R_v = m^{-1} \sum_{i=1}^m r_i \tag{3}$$

which is in other words the mean of its ranks in the set of orders with maximum score  $Q$ .<sup>9</sup> The prestige score of each university provides a natural ordering or ranking of universities that is our measure of prestige.

<sup>9</sup> Further details of the procedure are described in the Appendix.

A key assumption of this algorithm is that a single adjacency matrix  $A$  captures the underlying hierarchy of prestige. This implies that all universities (nodes) and scholars participate in the labor market and the size of the market is fixed. However, movements between universities over an interval of time can be constrained by various forces.

We relax the assumption that the hiring network is fixed across time, adopting a *dynamic* computation of the algorithm. The proposed variation iterates the previous algorithm over closed intervals,  $t = [y - \Delta, y + \Delta]$ , of time centered around the PhD graduation year  $y$ , with fixed windows of  $\Delta = 3$  years. This implies that the hiring network and the scholars and universities involved are different for each time window.<sup>10</sup> However, not all institutions are present across  $t$  intervals, for instance, more recent universities are not listed in the early years of the sample. Hence, the final scores of our *Interactive Prestige Ranking* is the average score of each university  $i$  over  $t$  intervals of time.

After the computation of the ranking, we distinguish between three groups of scholars, *Up*, *Down* and *Stay*, that exhibited different changes in the prestige during the transition between PhD graduation and first job in the following way. For each researcher in the sample, we calculate the difference between the PhD prestige rank and the one of her/his first job institution. The difference is positive (negative) for the group of scholars who move *Up* (*Down*) which experience upward (downward) prestige mobility—they are hired by a university more (less) prestigious than their PhD. The difference is equal to zero for scholars who *Stay* experiencing internal hiring—those hired by their PhD institution.

## Prestige Ranking Results

Table 2 shows the ranking of Mexican universities using the dynamic ranking with a 3 years time window. The rank follows from the average of the university rank computed by our algorithm across all time windows. The lower the average indicates that the university has occupied the higher positions more often across time periods.

Table 3 shows the stratification of prestige in the Mexican university system. Looking at the movements in the prestige hierarchy, Table 3 highlights a high level of stratification in the Mexican university system. Where the 10 most prestigious Mexican universities produce the 68% of PhD graduates and nearly half of them are hired as a first job in those institutions. This stratification is also geographical, as mostly all top universities are located near Mexico City. Thus Mexican academia operates in a highly stratified system where public and private research funds are mostly centralized.

## Descriptive Statistics

Table 4 presents the correlation between the dynamic ( $d$ ) and static ( $s$ ) prestige ranking (Pr) with individual level variables. In addition of the measurement of productivity from the NSR we compute bibliometric indicators using Science Citation Index data of *Web Of Science* (WOS) for the period 1992–2016. First, we compare the stability of prestige over time

<sup>10</sup> The number of universities is very small in early years of the sample, their analogous adjacency matrices yield trivial orders of small length. To overcome this, we aggregated the first years of the sample up to the year 2000.

by comparing the Spearman correlation coefficient from the static and dynamic ranks. The high correlation between the prestige from PhD and first job, 0.64 and 0.83 respectively, indicate that prestige does vary over time but not largely. The difference in prestige ( $\Delta Pr$ ) is positive (0.182) and negatively ( $-0.649$ ) correlated with the prestige from the PhD and first job, respectively. This pattern shows that moving up the ranking (positive difference) is associated with lower prestige from PhD university (higher in rank). Conversely, moving down the ranking (negative difference) is associated with a lower prestige from first job.

Next, we draw attention to the peer-review research productivity variable from NSR and the prestige variables. The results show that static ( $s$ ) and dynamic ( $d$ ) measures of prestige are negatively correlated with productivity (NSR). This is expected, since lower rank signifies higher university prestige. Interestingly, the correlation between  $\Delta Pr$ , and all the measurements of productivity is close to zero but positive and significant. A difference is equal to zero, indicates that scholars were hired by university after PhD graduation (stayed). Lastly, we show the correlation between the *NSR rating* and other bibliometric variables. In general, the NSR rating is positively correlated with bibliometric measurements of productivity. However, this correlation is not higher than 0.40, which suggest that the NSR performance variable takes into account local research and other products of research such as patents (See “[NSR Disciplines and Evaluation Procedure](#)” and the previous section for a description of the *NSR rating*).

## Quasi-Experimental Design

In this section, we examine how movements in the prestige hierarchy in the transition from the PhD to the first job affect scholars’ academic performance. We take advantage of a quasi-experimental design that naturally occurs in the academic labor market of early-career scholars. After PhD graduation, early-career researchers are self-selected into academic positions. However, the choice of hiring committees is a quasi-random assignment that naturally clusters PhD graduates into three groups according to prestige differentials: *Up*, *Down*, and *Stay*. The allocation is a partially random assignment given that early-career researchers typically have thin publication records after PhD graduation, such that ability and research skills are mostly unobserved. Given the asymmetric information in the labor market and its dynamics, the choice of hiring committees evaluating early-career researchers contains a large element of randomness (Oyer, 2006). In this setting, PhD prestige plays an important role as a signal mechanism of candidates’ quality.

To deal with the endogeneity of university prestige related to training, experience and individual characteristics, we use a bootstrap matching pairs technique. To assess the effect of early career university prestige on future academic performance, we compare the research performance of the bootstrap matches pairs of individuals with similar characteristics but different treatment (*Up*, *Down*, *Stay*). When individuals are paired holding constant their PhD institution, we test how changes in prestige relate to academic performance irrespective of training. Similarly, the comparison done matching scholars by their first job examines how prestige movements affect scholars performance for agents with the same first academic job. The Matthew effect of the academic labor market may amplify the role of early-career prestige on long-run academic performance. Therefore, we replicate our analysis comparing the performance of matched pairs scholars in the *short* (up to 2 years), *medium* (3–5 years) and *long-run* (6–25 years) after PhD graduation.

**Table 2** Ranking Mexican Universities 1992–2016 for STEM Sciences

Tier 1	Tier 2	Tier 3	Tier 4
1. CINVESTAV <sup>a</sup> (2.5294)	10. UAM (2.684)	19. BUAP (3.203)	28. TEC (4.8494)
2. CICESE (2.6497)	11. UAEM (4.0295)	20. UASLP (3.5339)	29. UDG (3.9136)
3. UNAM (1.3432)	12. INECOL (4.3667)	21. CIMAV (4.7517)	30. UGTO (3.8549)
4. CINVESTAV <sup>b</sup> (2.0916)	13. IBERO (7.3052)	22. IPN (3.0192)	31. IMSS (3.7836)
5. IMP (2.4086)	14. CICY (4.641)	23. UV (4.778)	32. UNISON (5.8237)
6. ECOSUR (3.0022)	15. CIO (4.8308)	24. UADY (6.7742)	33. UMich (4.8774)
7. IPICYT (4.5405)	16. COLPOS (4.937)	25. UAEM (3.8454)	34. UAEH (4.6305)
8. CIBINOR (4.8267)	17. CIQA (6.7092)	26. UCOL (4.6986)	35. UANL (3.8814)
9. INAOE (4.3645)	18. CIAD (6.5083)	27. UAQ (6.3504)	36. TECNAM (3.2274)

Full names of universities are listed in Appendix “A. Faculty Hiring Matrix Names”

\*Brackets correspond to standard deviations

<sup>a</sup> The research unit of CINVESTAV located in Mexico City (Zacatenco)

<sup>b</sup> All the other research units of CINVESTAV located in other provinces of Mexico. We consider important to distinguish from (<sup>a</sup>) from (<sup>b</sup>) as research in Mexico in highly centralized

**Table 3** Prestige movements per university prestige 1992–2016

	Down	Stay	Up	Tot (%)
PhD Tier 1	304	283	17	68
PhD Tier 2	54	49	22	14
PhD Tier 3	38	48	12	10
PhD Tier 4	10	47	14	8
Job Tier 1	49	283	46	42
Job Tier 2	64	49	13	14
Job Tier 3	144	48	2	22
Job Tier 4	149	47	4	22

The Table presents how many scholars in each university prestige ranking-tier moved, considering the PhD (top) and first job (bottom) institutions

Our quasi-experimental design follows Cowan and Rossello (2018) and Way et al. (2019). The basic idea is that career movements from the PhD to the first job is a quasi-random assignment made by the job market. Individuals with the same PhD training are placed in different institutions, likewise, individuals with different training (PhD) are

**Table 4** Correlation between Prestige Ranking and bibliometric measures

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Pr-PhD <sup>d</sup>	1												
2. Pr-Job <sup>d</sup>	0.533***	1											
3. Δ Pr	0.182***	-0.649***	1										
4. Pr-PhD <sup>s</sup>	0.640***	0.342***	0.181***	1									
5. Pr-Job <sup>s</sup>	0.309***	0.829***	-0.650***	0.413***	1								
6. Age	-0.119***	-0.140***	0.063***	-0.089***	-0.121***	1							
7. Gender	0.014	0.093***	-0.093***	0.053**	0.098***	-0.084***	1						
8. NSR	-0.141***	-0.183***	0.066***	-0.106***	-0.179***	-0.059***	0.130***	1					
9. Cites	-0.045**	-0.085***	0.035*	-0.033*	-0.103***	0.019	-0.017	0.195***	1				
10. CNCI	-0.114***	-0.092***	0.004	-0.113***	-0.091***	0.001	-0.043**	0.217***	0.163***	1			
11. Pubs	-0.017	-0.064***	0.042*	0.040*	-0.053**	0.046**	0.085***	0.295***	0.415***	0.145***	1		
12. Frac. Pubs	-0.098***	-0.118***	0.049**	-0.068***	-0.114***	-0.009	0.125***	0.366***	0.014	0.256***	0.205***	1	
13. H-Index	-0.109***	-0.194***	0.100***	-0.074***	-0.209***	0.004	-0.001	0.379***	0.664***	0.272***	0.316***	0.130***	1

Pr corresponds to the continuous prestige for the PhD and first job. We compare the dynamic (*d*) and the static (*s*) computation of the ranking algorithm. For the static prestige rank, we use the full sample represented in one adjacency matrix that yields an atemporal ranking

Δ Pr is the variable of interest that measures the discrete change in prestige from PhD to first job. A negative change means that a scholar moved down the prestige ladder, and vice versa

NSR Rating is the dependent variable that scores research performance from one to five (the best performance)

Cites are the total citations matched from WOS 1992–2016.

CNCI are Category Normalised Citation Impact computed for each researcher (1992–2016). CNCI is computed for each scholar, dividing the actual citations by the expected citation rate for similar documents, year of publication and subject area. For documents with more subject areas, an average of the ratios of the actual to expected citations is used. The CNCI of a university is the average of the CNCI values for all the papers produced a scholar

Pubs are the total publications of each scholar matched from WOS 1992–2016

Frac. Pubs corresponds to the number of publications divided by the number of co-authors

H-Index is the maximum number of *h* papers have received at least *h* citations

\*p<.05, \*\*p<.01, \*\*\*p<.001

placed in the same first job. Thus, our strategy compares matched couples that either received the same training (PhD) or are exposed by the same working environment (same first job) but experienced different prestige movements (treatments groups). The additional variables on which we match the pairs are: gender, age, discipline, and graduation year.

However, we should remark two potential limitations and the scope for future research. The first limitation is that individuals with the same PhD training (or first job) might have slightly different productivity or performance (broadly defined). Unfortunately, we do not have prior students' data to control for that. Besides this limitation, we must highlight that our method partially implicitly controls for performance. Each PhD program has rules and quality standards with minimum requirements, both in admission and in promotion decisions. Thus, all graduates from a PhD program met at least those minimum requirements. Internal rules at universities make the performance of individuals in the same PhD batch comparable. The same reasoning applies to the first job. Internal university rules make hiring committees accountable for their decisions. This implies that job candidates must meet minimum "quality" requirements to be considered for the job. The latter makes new hires of a university similar in terms of prior performance. Thus, our strategy might mitigate this potential data limitation.

A second limitation is that our technique discretized prestige movements in three treatment groups, rather than considering them as a numeric variable. Our method might have the disadvantage of considering movements from the first to the second ranked institution as movements from the first to the bottom ranked one. Besides this potential drawback, we should remark that the way in which we operationalized our ranking algorithm limits this possibility. The situation in the example above is very unlikely. Most hiring patterns from one university to other are stratified and clustered around universities of comparable prestige. Very distant prestige movements represents a strong violation of the basic assumptions of our ranking algorithm. The ranking algorithm therefore minimises that possibility. Even if it is still possible for an individual to move from the first to the bottom ranked institution, this situation is rare, since a university that has many individuals moving this way will be penalised by the ranking algorithm. The rank of a university is higher as much more students it is able to place in better ranked first job.

For each comparison *Up vs. Stay*, *Down vs. Stay*, and *Up vs. Down* we generate  $n = 10000$  bootstrap samples of the group on the left-hand side (the smaller) of its same size  $s$ . For each of the 10000 samples of size  $s$ , we create matched pairs of scholars matching on *gender*, *age*, *discipline*, *graduation year* and PhD (or first job) university. In order to compare their performance, in each sample we estimate the proportion of pairs,  $p^* = (p_1, p_2, \dots, p_n)$ , in which one group  $g^\alpha$  have higher performance than the other (group)  $g^\beta$ . Such as

$$p_n^\alpha = s^{-1} \sum_{i=1}^s I(g_i^\beta < g_i^\alpha) \quad (4)$$

$$I(g_i^\beta < g_i^\alpha) = \begin{cases} 1, & \text{if } g_i^\beta < g_i^\alpha \\ 0, & \text{Otherwise.} \end{cases}$$

Where academic performance  $g$  is the individual NSR research rating.<sup>11</sup> For each group in the comparisons *Up vs. Stay*, *Down vs. Stay* and *Up vs. Down*, we estimate the two  $p^*$  and

<sup>11</sup> NSR rating is one of 5 ordered categories.



construct their  $F(p^*)$  cumulative empirical distribution function (CEDF). To assess the performance of one group over the other, we test for first order stochastic dominance (Levy, 1992). This test implies higher performance of  $g^\alpha$  over  $g^\beta$  if  $F(p^\alpha) \leq F(p^\beta)$  for all  $p^*$ .<sup>12</sup> We compare the two CEDFs running a *two-sided* and a *one-sided* Kolmogorov-Smirnov test (KS test).

The null hypothesis of a *two-sided* test is  $H_{01} : F(p^\alpha) = F(p^\beta)$ —the two CEDF are drawn from the same distribution. Rejecting the null hypothesis  $H_{01}$  implies that the academic performance is statistically different between the two groups. The null hypothesis of the *one-sided* test is  $H_{02} : F(p^\alpha) \geq F(p^\beta)$ . Rejecting the null hypothesis implies that  $F(p^\alpha)$  stochastically dominates  $F(p^\beta)$ , in other words, that the increase in academic performance associated with a change of prestige from group  $g^\alpha$  is statistically different and greater than  $g^\beta$ .

## Matched Pairs Results

In this section, we compare scholars performance of *Up vs. Stay*, *Down vs. Stay*, and *Up vs. Down* in the short, medium and long-run. In particular, we examine the CEDF of the proportion in which one group received a higher NSR rating than the other. *Up*, *Down*, and *Stay* represent prestige changes from the PhD to the first job where prestige is measured with the dynamic ranking algorithm with a moving time window of 3 years.<sup>13,14</sup> Results of the KS-tests of  $H_{01}$  and  $H_{02}$  in Table 5 show for every comparison that the CEDFs are different and one group stochastically dominates the other.

Figure 2 compares the NSR research performance of matched pairs of scholars who *Stay* and move *Up* the hierarchy. In all figures, scholars match if they have the same *gender*, *age*, *discipline*, and *graduation year*. Additionally, figures on the left match scholars with the same PhD while those on the right-hand side match those with the first job institution. The matching procedure allows us to compare scholar with same PhD (or first-job) and characteristics but experiencing different prestige movements. In both cases, the CEDF of *Stay > Up* (solid lines), is located below that of *Up > Stay* (dashed lines) implying that the *Stay* group stochastically dominates the *Up* group. The implication of the results is the following. On the one hand, looking at scholars with the same (PhD) training (left-plots) we find that those hired internally have on average a better research performance than those experiencing upward prestige mobility (hired into a university more prestigious than their PhD). On the other hand, comparing scholars with the same first job (right-plots) but different training (PhD) we find that internal hired perform better than those coming from less prestigious PhDs (*Up*). These results suggest that scholars who manage to secure positions at their faculty after graduation demonstrate higher NSR levels of performance than those who migrate to upper ranked institutions.

Results for the comparison between the *Down* and *Stay* groups are in Fig. 3. In this case, we compare scholars who take academic positions in their faculties after graduation and PhD graduates experiencing downward prestige mobility. Results are the same matching the pairs on the PhD or the first job institution—the CEDF *Stay > Down* stochastically

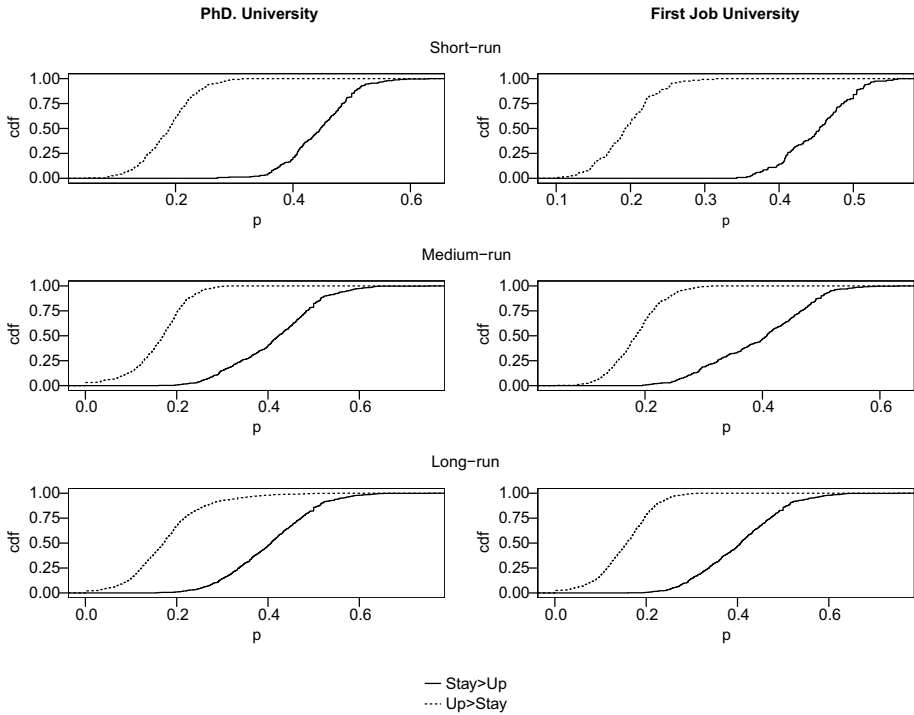
<sup>12</sup> Graphically, this is inspected if  $F(p^\alpha)$  lies below and to the right of  $F(p^\beta)$ .

<sup>13</sup> Details on the dynamic rank computation are in “[Interactive Prestige Ranking](#)” section.

<sup>14</sup> We execute the algorithm with a time window of  $\Delta = 3$  and run robustness checks using other time ranges ( $\Delta = 5$  and  $\Delta = 8$ ) and our results are consistent.

**Table 5** KS-test results for Phd and first job matches

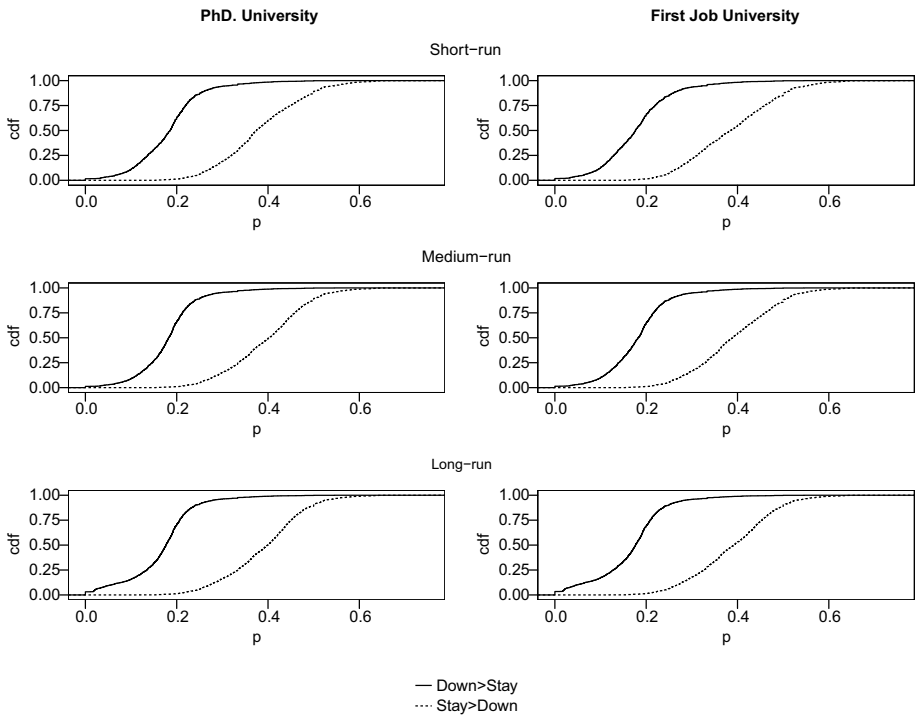
$H_0$	Short-run	Med-run	Long-run
<b>PhD Matches</b>			
$F(p^{Up}) = F(p^{Stay})$	D = 0.9874; p value = 0	D = 0.9097; p value = 0	D = 0.8165; p value = 0
$F(p^{Up}) \leq F(p^{Stay})$	D = 0; p value = 1	D = 0; p value = 1	D = 0; p value = 1
$F(p^{Up}) \geq F(p^{Stay})$	D = 0.9874; p value = 0	D = 0.9097; p value = 0	D = 0.8165; p value = 0
$F(p^{Down}) = F(p^{Stay})$	D = 0.812; p value = 0	D = 0.8487; p value = 0	D = 0.8564; p value = 0
$F(p^{Down}) \leq F(p^{Stay})$	D = 0; p value = 1	D = 0; p value = 1	D = 0; p value = 1
$F(p^{Down}) \geq F(p^{Stay})$	D = 0.812; p value = 0	D = 0.8487; p value = 0	D = 0.8564; p value = 0
$F(p^{Up}) = F(p^{Down})$	D = 0.7353; p value = 0	D = 0.6077; p value = 0	D = 0.5924; p value = 0
$F(p^{Up}) \leq F(p^{Down})$	D = 0; p value = 1	D = 0; p value = 1	D = 0; p value = 1
$F(p^{Up}) \geq F(p^{Down})$	D = 0.7353; p value = 0	D = 0.6077; p value = 0	D = 0.5924; p value = 0
<b>First Job Matches</b>			
$F(p^{Up}) = F(p^{Stay})$	D = 0.9227; p value = 0	D = 0.888; p value = 0	D = 0.9227; p value = 0
$F(p^{Up}) \leq F(p^{Stay})$	D = 0; p value = 1	D = 0; p value = 1	D = 0; p value = 1
$F(p^{Up}) \geq F(p^{Stay})$	D = 0.9227; p value = 0	D = 0.888; p value = 0	D = 0.9227; p value = 0
$F(p^{Down}) = F(p^{Stay})$	D = 0.8434; p value = 0	D = 0.8319; p value = 0	D = 0.8434; p value = 0
$F(p^{Down}) \leq F(p^{Stay})$	D = 0; p value = 1	D = 0; p value = 1	D = 0; p value = 1
$F(p^{Down}) \geq F(p^{Stay})$	D = 0.8434; p value = 0	D = 0.8319; p value = 0	D = 0.8434; p value = 0
$F(p^{Up}) = F(p^{Down})$	D = 0.6273; p value = 0	D = 0.6835; p value = 0	D = 0.6273; p value = 0
$F(p^{Up}) \leq F(p^{Down})$	D = 0; p value = 1	D = 0; p value = 1	D = 0; p value = 1
$F(p^{Up}) \geq F(p^{Down})$	D = 0.6273; p value = 0	D = 0.6835; p value = 0	D = 0.6273; p value = 0



**Fig. 2** *Up* versus *Stay* comparison. The solid curves are CEDFs of the proportion of pairs in which  $R_{Stay} > R_{Up}$ . Dotted curves are CEDFs for  $R_{Up} > R_{Stay}$ . Pairs matched by gender, age, discipline, graduation years, and same PhD university (left), or same first job university (right). From top to bottom: short-run (*Up* to 2 years), medium-run (3–5 years), and long-run (6–25 years) after PhD graduation

dominates *Down > Stay*. In line with our previous results, plots on the left-hand side show that scholars with the same PhD training moving (*Down*) to a less prestigious institution in their first job tend to have a lower NSR rating than those hired internally (*Stay*). Similarly, plots on the right-hand side compare scholars with the same first job and indicate that those hired internally (*Stay*) have a higher performance than those moving down the hierarchy.

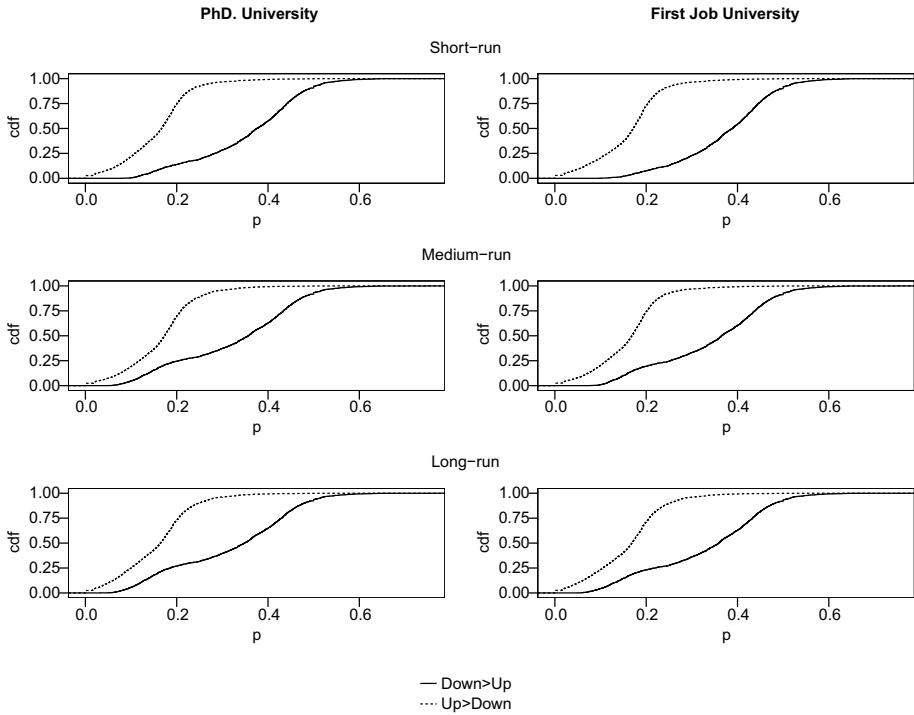
The last comparison in Fig. 4 examines performance differences between scholars who experience upward and downward prestige mobility. Results show that the CEDF of the proportion of pairs of scholars where the performance of *Down > Up* stochastically dominates *Up > Down*. The stochastic dominance of one over the other implies that scholars who experience downward prestige mobility sustain higher performance over their career than those experiencing upward mobility in their early career. These results are consistent both matching pairs, keeping fixed the (PhD) training (left-plots) or the first job (right-plots) institution. In the first case, comparing scholars with the same (PhD) training, we find that those moving down the hierarchy have higher performance than those moving up to more prestigious first job institutions. In the second, pairing scholars with the same first job but different PhDs institution, we find that those coming from more prestigious PhDs (*Down*) have a higher NSR rating on average than those moving up from less prestigious PhD institutions.



**Fig. 3** Down versus Stay comparison. The solid curves are CDFs of the proportion of pairs in which  $R_{Down} > R_{Stay}$ . Dotted curves are CDFs for  $R_{Stay} > R_{Down}$ . Pairs matched by gender, age, discipline, graduation years, and same PhD university (left), or same first job university (right). From top to bottom: short-run (Up to 2 years), medium-run (3–5 years), and long-run (6–25 years) after PhD graduation

These results seem counter-intuitive at first glance, since most studies have associated upward ranking mobility with higher academic performance.<sup>15</sup> In particular, the first result of a negative impact of upward prestige mobility comparing scholar with the same training contradicts the previous results of Chan et al. (2002). However, their analysis is slightly different. They use a longitudinal analysis in one sub-field of economics, and their sample is limited to scholars with publications in 16 top journals in finance. In contrast to us, they find that scholars who experience upward prestige mobility publish twice as many as their colleagues. What is most interesting is that what we found is a pattern through the career of scholars and for both dynamic (Fig. 4) and static ranking estimations (Appendix Fig. 9). Nevertheless, these findings require further research that we discuss in the “Discussion” section.

<sup>15</sup> Moreover, we should consider that those that start high in the hierarchy have few options to move up the ranking, and they are more likely to move down, while the opposite is true for those low in the prestige hierarchy. This implies that those moving up (down) are more likely to have PhDs in less (more) prestigious institutions.



**Fig. 4** Down versus Stay comparison. The solid curves are CEDFs of the proportion of pairs in which  $R_{Down} > R_{Stay}$ . Dotted curves are CEDFs for  $R_{Stay} > R_{Down}$ . Pairs matched by gender, age, discipline, graduation years, and same PhD university (left), or same first job university (right). From top to bottom: short-run (Up to 2 years), medium-run (3–5 years), and long-run (6–25 years) after PhD graduation

### Robustness Checks

Our methodology might be prone to potential weaknesses that we discuss in this section, providing additional robustness checks.

### Static and Dynamic Ranking Results Comparison

The first potential drawback stands in the different time-frame between the categorization of prestige movements (*Up*, *Down*, *Stay*) which follows from the dynamic prestige ranking and the dependent variable *NSR rating*. This might mean that people move to the hierarchy in the PhD to the first job transition, and at the same time universities might change their ranking position relative to other institutions. A simple way to overcome this limitation is to run the same analysis using the static prestige ranking. In this case university prestige is assumed to be constant over the period of analysis and thus movements in prestige are computed on the basis of this aggregation of the data. Results using the static prestige ranking, in Appendix Figs. 7, 8, 9, are consistent with the previous. Additionally, we should remark the issue above does not apply to the short run estimation (top panels of Figures 1, 2 and 3). Since the dynamic computation of the ranking is done using windows of times that overlaps with the comparison of the research performance. In particular, the evaluation of the NSR ratings (our dependent variable) overlaps with the period of the transition up/

**Table 6** Ordinal logistic regression: changes of prestige on academic performance

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Pr-PhD	-0.014*** (0.002)			-0.009*** (0.002)	-0.018*** (0.002)	-0.011*** (0.003)	-0.011** (0.003)
Pr-Job		-0.013*** (0.002)		-0.009*** (0.002)		-0.006* (0.003)	-0.006 (0.003)
Δ Pr			0.004 (0.002)		0.009*** (0.002)		
Δ Pt=Down						0.144*** (0.029)	-0.250*** (0.075)
Δ Pt=Stay						0.394*** (0.038)	
Δ Pt=Up							-0.395*** (0.106)
Gender	0.328*** (0.052)	0.356*** (0.052)	0.339*** (0.052)	0.347*** (0.052)	0.347*** (0.052)	0.362*** (0.052)	0.362*** (0.052)
Discipline II	-0.184** (0.068)	-0.186** (0.068)	-0.181** (0.068)	-0.187** (0.068)	-0.187** (0.068)	-0.163* (0.068)	-0.163* (0.068)
Discipline III	-0.194 (0.109)	-0.200 (0.108)	-0.287** (0.108)	-0.168 (0.109)	-0.168 (0.109)	-0.216* (0.109)	-0.216* (0.109)
Discipline VI	-0.576*** (0.090)	-0.623*** (0.090)	-0.635*** (0.090)	-0.594*** (0.090)	-0.594*** (0.090)	-0.587*** (0.090)	-0.587*** (0.090)
Discipline VII	-0.067 (0.078)	-0.037 (0.079)	-0.097 (0.078)	-0.032 (0.079)	-0.032 (0.079)	-0.029 (0.079)	-0.029 (0.079)
Age	0.110*** (0.007)	0.120*** (0.007)	0.118*** (0.006)	0.115*** (0.007)	0.115*** (0.007)	0.120*** (0.007)	0.120*** (0.007)
Age <sup>2</sup>	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)

**Table 6** (continued)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
AIC	12850.649	12842.576	12891.727	12832.456	12832.456	12809.343	12809.343
BIC	12930.017	12921.944	12971.094	12918.438	12918.438	12908.552	12908.552
Log Likelihood	-6413.325	-6409.288	-6433.863	-6403.228	-6403.228	-6389.672	-6389.672
Deviance	12826.649	12818.576	12867.727	12806.456	12806.456	12779.343	12779.343
Num. obs.	5508	5508	5508	5508	5508	5508	5508

\*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$



down/stay from the PhD to the first job. In more general terms, comparing the results using the static and the dynamic ranking. The main results show consistency.

## Regression Analysis

In this sub-section, we further assess the robustness of our previous results using an *Ordinal Logistic Regression* model. In Table 6 the dependent variable is the *NSR rating* of research performance with five ordered categories (See Section “Data”). We include the control variables that we use in the stochastic pair analysis.

Models 1 to 3, on Table 6, explore the effect of university prestige from PhD, first job and change in prestige (Pr-PhD, Pr-Job,  $\Delta$ Pr). As expected, the models show negative coefficients on prestige of PhD and first job given that higher rank (lower prestige) decreases the likelihood of achieving higher academic performance. These results are also consistent in Model 4, that shows a similar effect of prestige from PhD and first job. Model 3 and Model 5 incorporate the continuous change in prestige ( $\Delta$ Pr), which is significant but closer to zero. Similarly to the stochastic pair analysis, these results suggest that researchers who are hired by their faculty ( $\Delta$ Pr = 0) have higher odds of achieving higher research performance. To better understand the average change of moving *Up*, *Down* or *Stay* in the ranking we estimate Models 6 and 7 using the categorical variable of change in prestige similar to the previous section. We argue that the discrete change in prestige is more suitable to analyze the BFLPE, as it captures the shock in prestige internalized by researchers.

As shown in Table 1, scholars move on average only 6 places ( $\Delta$ Pr). Henceforth, it is more likely that a graduate move up or down of few position in the rank. For example, A graduate who moved from the most prestigious university in the country to the second most prestigious is more common than moving from the most prestigious university to the least prestigious for their first job.

However, both could be labelled by their peers as researchers moving down the ladder and, so, they can also internalize their experience psychologically thinking that they have been “downgraded” and perceive a sense of failure (Waters and Leung, 2017).

Indeed, scholars are typically only aware of the universities and departments from whom they work and compete. Thus, they operate with asymmetric information and are myopic with respect to their precise place in the distribution of prestige. Henceforth, moving up/down within their league (average 6), has other potential costs and psychological effects (See “Discussion” section). And, the effect of changing prestige may be internalized irrespectively of the change in the number of places in the prestige ladder.

Model 6, estimates the average effect of moving *Down* and *Stay* with respect to the baseline category *Up*. The results are consistent with our previous results, Model 6 shows that the researchers who experience downward mobility reported 115% ( $\exp(0.144)$ ) odds of achieving higher levels of academic performance in comparison to the *Up* group. In the same model, the group of researchers hired by their faculty after PhD graduation reports 158% higher odds of achieving higher performance rates than the group who stay. Model 7, has the group *Stay* as baseline and henceforth the log-odds of moving *Up* or *Down* are negative. Figures 5 and 6, reports the predicted probability for each group, clearly the probability of achieving higher levels of research performance is more loaded in the stay group. The predicted probability of achieving an NSR level II or higher is approximately 10% larger for the group that stayed in comparison to the group that went up. These results are in line with the previous section.

## Discussion

Our main findings underline a more complex relation between prestige movements and performance than the common positive association. Our first result is that scholars experiencing downward prestige mobility show higher performance compared with colleagues with similar characteristics moving upward in the prestige hierarchy. This result challenges past studies that often associates upward prestige movements with above average academic performance (Chan et al., 2002). We interpret this finding in the light of the literature of the BFLPE and related mechanisms presented in “[The BFLPE and Similar Mechanisms](#)” section.

Becoming a researcher in academia has career phases where the social comparison might matter. The BFLPE hypothesis in psychology claims that individual performance can be affected by how individuals perceive themselves in comparison with their peers. The transition from the PhD to the first academic job is a stressful event in a scholar career, where the social comparison might matter more. Indeed, the psychological literature predicts that the BFLPE mechanism takes place when individuals change their environment.<sup>16</sup> When individuals change institution, or advance in their career, their peers also changes. For instance, after PhD graduation, mobile scholars change their place in the prestige hierarchy. A scholar moving down the hierarchy has higher prestige and academic self concept relative to their incumbent peers with lower prestige. The higher prestige and academic self-concept translates into higher competitiveness, visibility and resources than scholars moving upward.

Following the BFLPE mechanism, we can interpret the results of the Mexican PhD job market as follows. On the one hand, scholars who move down might think they are “*big shots*” relative to their new peers, and this is beneficial to their performance. Or in terms of positive peer effects, their new colleagues might think they have made “*a catch*” and give them more resources. In addition, an initial “*failure*”—moving down in the hierarchy—might lead to an effort to “*regain the previous prestige*”. Our results might also relate to how co-workers see the new hired, and this might generate different peer effects. Depending on how a new researcher is integrated in the new department, the peer effect can positively or negatively affect individual performance.

The explanation mentioned above is consistent with the apparently paradoxical result in this paper. To accurately test the BFLPE mechanism will require psychological tests on self-esteem or academic self-concept. However, the latter is out of the scope of our work and should motivate further studies. Still, we consider our results to be relevant for the policy debate of prestige stratification and mobility in the academic market, specially in developing settings. In light of the consistency of our findings, further studies should operationalize a standardised psychological test aimed at measuring the change in academic self-concept.

Our second result suggests a positive association between internal hiring and research performance in comparison to scholars moving in the hierarchy. This result, in conjunction with our analysis of the stratification of the Mexican system, suggests that PhD graduates are not moving to the peripheral areas of Mexico. The market is operating quite efficiently, however, identifying talented PhDs that are hired internally. This also suggests a negative effect of mobility, as internal hiring is associated with higher research performance.

<sup>16</sup> See Marsh, 1991 for a discussion.

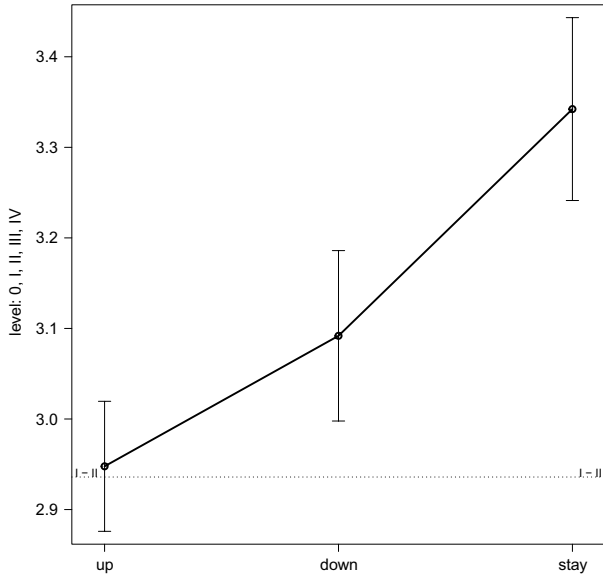
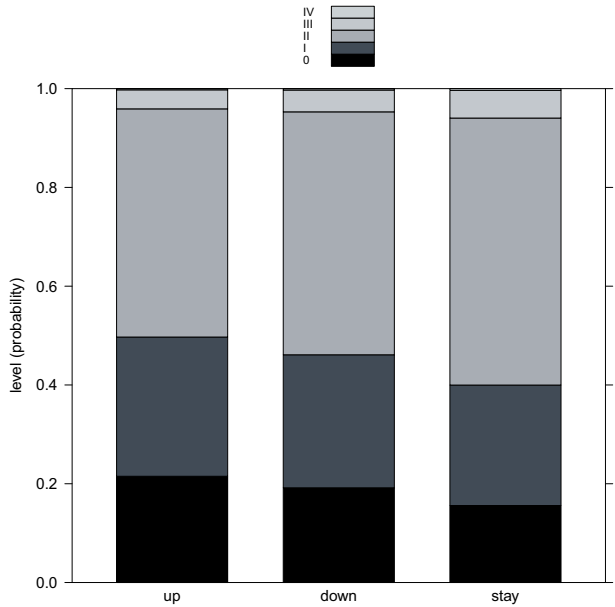


Fig. 5 Marginal effect Model 6 Table 6

Fig. 6 Predicted probability Model 6 Table 6



The negative association between early career mobility and performance is not surprising. Past literature on academic inbreeding<sup>17</sup> finds that its relationship with individual performance is ambiguous and depends on the sample, country, era, field, career stages, and measures which are used (Gorelova and Yudkevich, 2015; Capponi and Frenken, 2021). For example, Cruz-Castro and Sanz-Menéndez (2010), controlling for the type of mobility experienced by researchers, find that in Spain inbreeding scholars obtain a tenure earlier and with a higher number of publications than mobile colleagues. While, Horta et al. (2010) with Mexican data find a negative relationship between inbreeding and the number of publications. However, our work differs to theirs in many respects. First, we use a more nuanced measure of performance, the NSR rating which includes the local knowledge production and the quality evaluation of independent scholars specialized in the fields. Second, we give a more narrow definition of mobility focusing on the transition from the PhD to the first job and excluding international mobility.

The positive relation between inbreeding and performance in early-career is not surprising in Mexico, where inbreeding is pretty common. The Mexican university system is highly stratified and geographically centralized. The centralization of science and research is a well-known problem in Mexico, that is changing slowly through investing in research infrastructure in peripheral areas (Lopez-Olmedo et al., 2017). However, our results suggest that science policies in Mexico should continue working on increasing the flows of specialized human capital to aid the development of regional capabilities. With these lenses, our results might highlight that inbreed scholars specialized in areas germane to their faculty. A similar thing might be beneficial for the career of individuals but not necessarily good for the system as a whole.

More generally, Oyer (2006) highlights that the dynamics of the job market might impact the career in the long run. Possible mechanisms are university-specific capabilities and norms, co-worker behavior, and university turnover. The latter in particular implies that the time of graduation and information about the job market might give individuals an advantage/disadvantage depending on the job market conditions. In a system with few resources and low mobility, two main mechanisms might explain why those internally hired are more successful than the others. On the one hand, resource constraints make the job market less predictable because available positions at universities might be subjected to economic fluctuations and budget cuts. On the other hand, low mobility levels make that information about universities and their job market sticky and localized. In a similar case, those able to secure their initial positions at universities will have better information of the job dynamics as well as about norms and routines that might help them to become more germane to their institutions compared to those trained elsewhere.

We should remark a limitation of our analysis, our work considers Mexico as a closed system. We do not have data on foreign PhD graduates returning to the country. However, given the ties between Mexico and North America, it is likely that many Mexican trained abroad return to their country. For example, Finn (2010) estimates that only 40% of Mexican PhD graduates, trained in the US, are hired in a US university.<sup>18</sup> In the same line, Rivero and Peña (2020) show that repatriation policies have contributed to keeping the rate of

<sup>17</sup> Academic inbreeding focuses on the other side of the coin with respect to mobility: immobility and inertia. Academic inbreeding is defined when universities hire their own PhD graduates (Gorelova and Yudkevich, 2015).

<sup>18</sup> See <https://orise.orau.gov/stem/reports/stay-rates-foreign-doctorate-recipients-2011.pdf>; Last access December 2020.

return of Mexican researchers to 60% and 83% from Europe and the US, respectively. This suggests that a large proportion of PhD graduates returns to Mexico. This sub-sample of foreign-trained PhDs might behave differently following different career paths. However, besides this limitation, we concentrate on the Mexican university system, and we consider our findings relevant for science policy in emerging contexts.

## Conclusion

This work is the first studying how prestige stratification affects scientific performance of early career scientists in the Mexican higher education system. The majority of comparable analyses looked at university systems in developed economies, especially in North America, where mobility after the PhD tends to be high and systems are more integrated. In the U.S. for example, there are often hiring practices that prevent universities from hiring their own graduates immediately after the PhD. Studying these mechanisms in less mature settings with higher resource constraints has policy implication because the gap between prestige and performance can be larger than in other settings.

Our findings in general suggest that there is a negative relation between mobility during early-career and academic performance. Moreover, when we decompose mobility looking at prestige differentials between PhD and first job institution, we find that scholars who *Stay* or move *Down* the hierarchy remain mostly in first-tier (top 10) institutions.<sup>19</sup> Our results of the matched pair analysis provide evidence of the same association of prestige movements and performance in the short-, medium- and long-run. Further, comparing those moving up with those moving down the hierarchy, we find that those moving down have sustained higher performance than those moving up.

The reasons why promising scholars experiencing upward prestige mobility have lower performance than their colleagues (with the same PhD) that stay or move downwards requires further investigation, as highlighted in the previous section. Similarly to the higher education system in the U.S. and other developed economies, we find a large stratification in the Mexican university system but low mobility (around 50% of PhD graduates are hired by their faculty), with a few prestigious institutions (around 10) producing the majority of PhD graduates that are subsequently mostly hired in these same institutions. A high concentration of prominent scholars in a few academic institutions reveals large inequalities in the distribution of prestige. On the one hand, the stratification of higher education could promote higher levels of specialization with a targeted allocation of resources. On the other hand, it can also reveal a structural problem in the science system, a “lock-in”, where researchers are trained and hired by elite institutions and flows of knowledge are reduced throughout the national science system. In the case of Mexico, a structural “lock-in” could be additionally reinforced by the negative association between mobility (upward prestige mobility in particular) and performance.

---

<sup>19</sup> Those moving *Up* the hierarchy, get their PhD degree mostly from second-tier (bottom 30) institutions, but move to first-tier institutions with their first job. This is an expected result, in the light that those graduating from prestigious universities have fewer possibilities to move higher in the hierarchy.

## Appendix

### A. Faculty Hiring Matrix Names

	ABBREVIATION	NAMES
1	CINVESTAV(ZACATENCO)	CENTRO DE INVESTIGACION Y DE ESTUDIOS AVANZADOS DEL I.P.N. - UNIDAD ZACATENCO
2	CICESE	CENTRO DE INVESTIGACION CIENTIFICA Y DE EDUCACION SUPERIOR DE ENSENADA
3	UNAM	UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
4	CINVESTAV(OTHERS)	CENTRO DE INVESTIGACION Y DE ESTUDIOS AVANZADOS DEL INSTITUTO POLITECNICO NACIONAL
5	IMP	INSTITUTO MEXICANO DEL PETROLEO
6	ECOSUR	EL COLEGIO DE LA FRONTERA SUR
7	IPICYT	INSTITUTO POTOSINO DE INVESTIGACION CIENTIFICA Y TECNOLÓGICA AC
8	CIBINOR	CENTRO DE INVESTIGACIONES BIOLÓGICAS DEL NOROESTE SC
9	INAOE	INSTITUTO NACIONAL DE ASTROFISICA OPTICA Y ELECTRONICA
10	UAM	UNIVERSIDAD AUTONOMA METROPOLITANA
11	UAEM	UNIVERSIDAD AUTONOMA DEL ESTADO DE MORELOS
12	INECOL	INSTITUTO DE ECOLOGIA, A.C
13	IBERO	UNIVERSIDAD IBEROAMERICANA AC
14	CICY	CENTRO DE INVESTIGACION CIENTIFICA DE YUCATAN AC
15	CIO	CENTRO DE INVESTIGACIONES EN OPTICA AC
16	COLPOS	COLEGIO DE POSTGRADUADOS
17	CIQA	CENTRO DE INVESTIGACION EN QUIMICA APLICADA, A.C
18	CIAD	CENTRO DE INVESTIGACION EN ALIMENTACION Y DESARROLLO AC
19	BUAP	BENEMERITA UNIVERSIDAD AUTONOMA DE PUEBLA
20	UASLP	UNIVERSIDAD AUTONOMA DE SAN LUIS POTOSI
21	CIMAV	CENTRO DE INVESTIGACION EN MATERIALES AVANZADOS SC
22	IPN	INSTITUTO POLITECNICO NACIONAL
23	UV	UNIVERSIDAD VERACRUZANA
24	UADY	UNIVERSIDAD AUTONOMA DE YUCATAN
25	UAEM	UNIVERSIDAD AUTONOMA DEL ESTADO DE MEXICO
26	UCOL	UNIVERSIDAD DE COLIMA
27	UAQ	UNIVERSIDAD AUTONOMA DE QUERETARO
28	TEC	INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE MONTERREY
29	UDG	UNIVERSIDAD DE GUADALAJARA
30	UGTO	UNIVERSIDAD DE GUANAJUATO
31	IMSS	INSTITUTO MEXICANO DEL SEGURO SOCIAL

	ABBREVIATION	NAMES
32	UNISON	UNIVERSIDAD DE SONORA
33	UMICH	UNIVERSIDAD MICHOACANA DE SAN NICOLAS DE HIDALGO
34	UAEH	UNIVERSIDAD AUTONOMA DEL ESTADO DE HIDALGO
35	UANL	UNIVERSIDAD AUTONOMA DE NUEVO LEON
36	TECNM	TECNOLOGICO NACIONAL DE MEXICO

```

/* Input: A:= weighted directed Adjacency Matrix with zero entries in the main
   diagonal. */
sk:= A top-score defined with an arbitrary small initial value;
S:= An empty vector of max scores;
O:= An empty matrix of orders with u columns;
for i = 0 to n do
    o0:= Generate a random n-tuple of length u;
    M:= Sort the matrix A by o0;
    s0:= Compute current score by adding the upper triangular elements of M by row;
    /* Local Search */
    for i = 0 to n do
        o1:= Swap randomly two elements of o0;
        M:= Sort the matrix M by o1;
        s1:= Compute swap score by adding the upper triangular elements of M by row;
        if s1 ≥ s0 then
            M:= Sort the matrix A by o1;
            sk:= Update the top-score by s1;
        end if
    end for
    dif:= Compute difference between s1 and sk;
    if dif > 0 then
        sk:= Update the top-score by s1;
        S:= Add the current score sk as an element of S ;
        O:= Add a row of orders given by s1 ;
    end if
end for
/* Output1: S:= A set of highest scores */
/* Output2: O:= A matrix of orders for each element in S */

```

**Algorithm 1:** Prestige Ranking Algorithm by Cowan and Rossello (2018).



## NSR Disciplines and Evaluation Procedure

### Disciplines

#### Levels of Rewards linked to the NSR ratings

- SNI Candidates<sup>20</sup>: Granted for 3 years, with the possibility of 2 years of extension.
- Level I: Granted for 3 years the first time, and every 4 years in the following periods.
- Level II: Granted for 4 years the first time, and every 5 years in the following periods.
- Level III: Granted for 5 years the first and second time, and every 10 years in the following periods.
- Emeritus Professors: Candidates must have 65 years of more, and have accumulated at least three periods of level III distinction (15 years) without interruption.

#### NSR Rating and Evaluation

We detail here the NSR rating process. The NSR rating is a peer review evaluation of research performance. Scholars submit their CVs with their scientific production. The evaluation of the CVs takes into consideration primarily the research output and linkages with industry and the public sector and to a less extent human capital formation of research groups<sup>21</sup>. The research output include: Scientific Articles, Books, Book Chapters, Patents, Technological Developments, Innovations and Transfers of Technology. These research products contain not only publications indexed in Web Of Science or Scopus but also local research in Spanish which is relevant for the country.

The “Evaluation Commission” reviews the quantity and the quality of the aforementioned research products in each CV (application). All the members of the commission participate in the revision of applications, but the evaluation is the direct responsibility of two members of the commission. The Evaluation Commission is comprised by a heterogeneous group of 14 prominent scholars from different institutions, changed every 3 years<sup>22</sup>. The rotation of the members and evaluation process is overlooked by the “Council of Approval” and several CONACYT authorities. Their job is to eradicate any personal bias and discrimination to ensure a meritocratic evaluation.<sup>23</sup>

**Table 7** Scientific Disciplines of the NSR System

Discipline no.	Scientific discipline description
Discipline I	Physics-Mathematics and Earth Sciences
Discipline II	Biology, Chemistry, Life Sciences
Discipline III	Medicine and Health Sciences
Discipline IV	Humanities and Behavioral Sciences
Discipline V	Social Sciences
Discipline VI	Biotechnology and Agricultural Sciences
Discipline VII	Engineering

<sup>20</sup> Applicants can only receive this distinction one time.

<sup>21</sup> See “Glosario de Términos” for more details. Available at [https://www.conacyt.gob.mx/images/SNI/GLOSARIO\\_DE\\_TERMINOS\\_BASICOS\\_Y\\_RECOMENDACIONES\\_SNI.pdf](https://www.conacyt.gob.mx/images/SNI/GLOSARIO_DE_TERMINOS_BASICOS_Y_RECOMENDACIONES_SNI.pdf). (Last accessed July 2021).

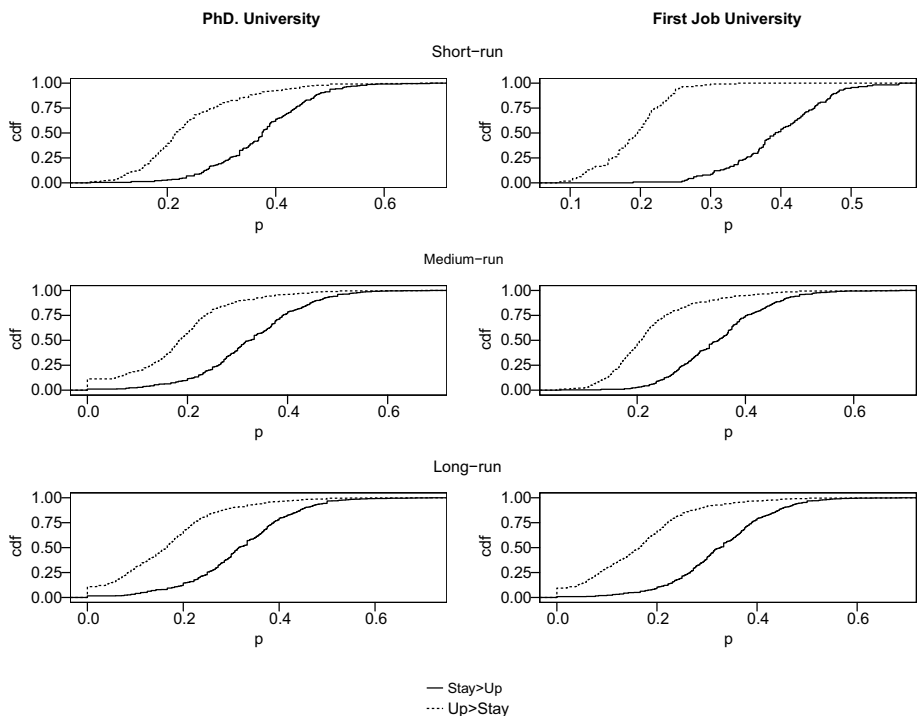
<sup>22</sup> See “Miembros de Comisiones” for more details. Available at <https://www.conacyt.gob.mx/Miembros-de-comisiones.htm>. (Last accessed July 2021).

<sup>23</sup> See “Lineamientos para el funcionamiento de las Comisiones Dictaminadoras” for more details. Available at <https://www.conacyt.gob.mx/PDF/sni/lineamientos-comisiones-dictaminadoras-y-transversales-sept-2019.pdf>.

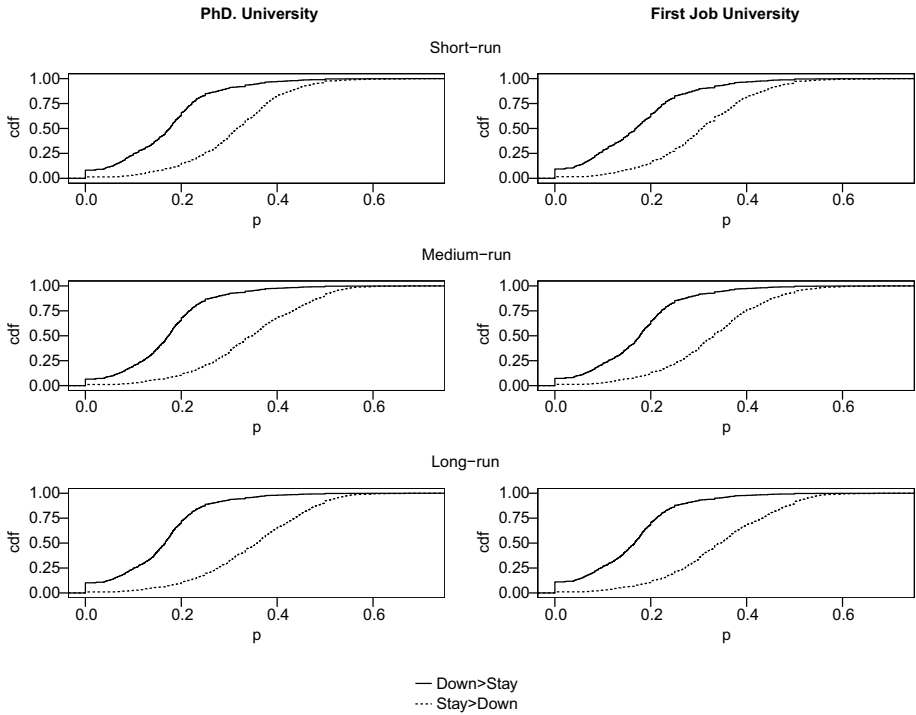
There are multiple commissions for each of the corresponding research areas described in Table 7. The commission evaluates and assigns researchers to one level of research performance. The levels of research performance (NSR ratings) are ‘SNI Candidate’, ‘Level I’, ‘Level II’, ‘Level III’ (ordered from low to high performance). There is a special category called ‘Emeritus’ excluded from our analysis because this recognition is uncommon. An Emeritus recognition is an honorary recognition to professors at the end of a career. Every NSR rate has an associated economic reward that increases linearly. Belonging to the NSR, implies high recognition of the quality and academic prestige of the researcher, the result of a scientific production of considerable importance at the national level and, in some cases, also at the international level (Reyes and Suriñach, 2013). Thus, our dependent variable, NSR ratings, is not a productivity measure only, but it measures research performance in a broader sense. In contrast with bibliometric measures, our measure of research performance accounts for ‘standard’ of quality from the national context within each discipline.

## Results Static Rank

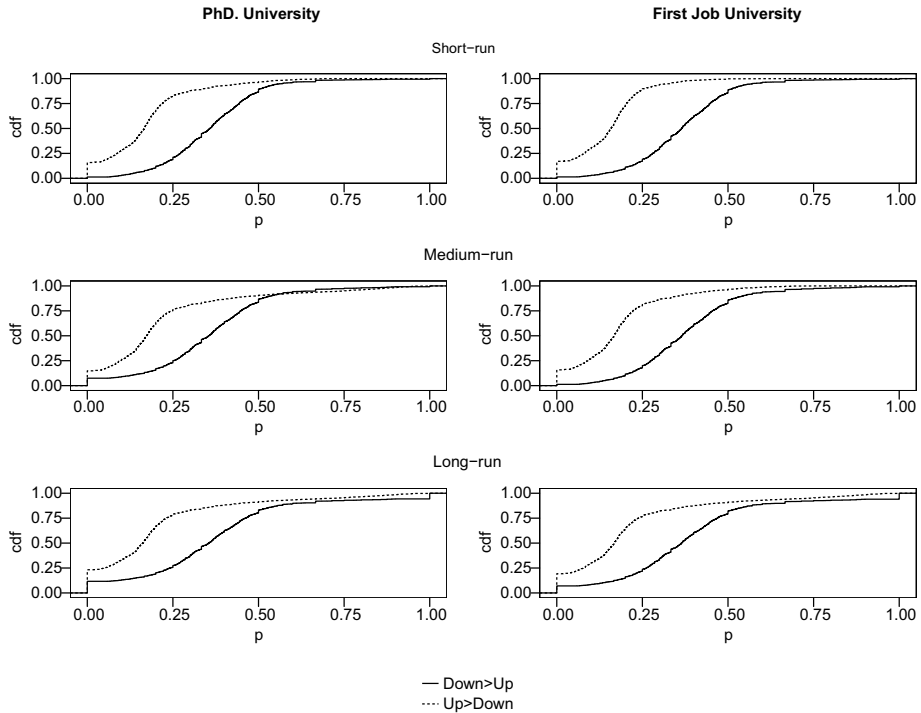
Figures (7, 8 and 9).



**Fig. 7** Stochastic Analysis of **Up vs Stay**. The solid curves are CDFs of the proportion of pairs in which  $R_{Stay} > R_{Up}$ . Dotted curves are CDFs for  $R_{Up} > R_{Stay}$ . Pairs matched by gender, age, discipline, graduation years, and same PhD university (left), or same first job university (right). From top to bottom: short-run (Up to 2 years), medium-run (3–5 years), and long-run (6–25 years) after PhD graduation



**Fig. 8** Stochastic analysis of **Stay vs Down**. The solid curves are CEDFs of the proportion of pairs in which  $R_{Down} > R_{Stay}$ . Dotted curves are CEDFs for  $R_{Stay} > R_{Down}$ . Pairs matched by gender, age, discipline, graduation years, and same PhD university (left), or same first job university (right). From top to bottom: short-run (Up to 2 years), medium-run (3–5 years), and long-run (6–25 years) after PhD graduation



**Fig. 9** Stochastic analysis of **Up vs Down**. The solid curves are CEDFs of the proportion of pairs in which  $R_{Down} > R_{Stay}$ . Dotted curves are CEDFs for  $R_{Stay} > R_{Down}$ . Pairs matched by gender, age, discipline, graduation years, and same PhD university (left), or same first job university (right). From top to bottom: short-run (Up to 2 years), medium-run (3–5 years), and long-run (6–25 years) after PhD graduation

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Abbott, W. F., & Barlow, H. M. (1972). Stratification theory and organizational rank: Resources, functions, and organizational prestige in American Universities. *Sociological Perspectives*, 15(4), 401–424. <https://doi.org/10.2307/1388261>
- Allison, P. D., & Long, J. S. (1987). Interuniversity mobility of academic scientists. *American Sociological Review*, 52(5), 643. <https://doi.org/10.2307/2095600>
- Allison, P. D., & Long, J. S. (1990). Departmental effects on scientific productivity. *American Sociological Review*, 55(4), 469. <https://doi.org/10.2307/2095801>
- Allison, P. D., & Stewart, J. A. (1974). Productivity differences among scientists: Evidence for accumulative advantage. *American Sociological Review*, 39(4), 596–909.
- Appelt, S., de Pinho, R., Fernández-Zubieta, A., Franzoni, C., Freeman, R. B., Galindo-Rueda, F., Ganguli, I., Geuna, A., Huang, W., Kataishi, R., Lawson, C., Scellato, G., Shibayama, S., Stephan, P., Toselli,

- M., van Beuzekom, B., Van Bouwel, L., & Veugelers, R. (2015). Introduction. In A. Geuna (Ed.), *Global mobility of research scientists* (pp. 15–20). Academic Press. <https://doi.org/10.1016/B978-0-12-801396-0.02003-2>
- Astin, A. W. (1969). A student's dilemma: Big fish-little pond or little fish-big pond. *Journal of Counseling Psychology*, *16*(1), 20–22. <https://doi.org/10.1037/h0026684>
- Azoulay, P., Stuart, T., & Wang, Y. (2014). Matthew: Effect or fable? *Management Science*, *60*(1), 92–109. <https://doi.org/10.1287/mnsc.2013.1755>
- Bair, J. H. (2003). Hiring practices in finance education: Linkages among top-ranked graduate programs. *American Journal of Economics and Sociology*, *62*(2), 429–433. <https://doi.org/10.1111/1536-7150.t01-2-00007>
- Baldi, S. (1995). Prestige determinants of first Academic job for New Sociology Ph.D.s 1985–1992. *Sociological Quarterly*, *36*(4), 777–789. <https://doi.org/10.1111/j.1533-8525.1995.tb00464.x>
- Barnett, G. A., Danowski, J. A., Feeley, T. H., & Stalker, J. (2010). Measuring quality in communication doctoral education using network analysis of faculty-hiring patterns. *Journal of Communication*, *60*(2), 388–411. <https://doi.org/10.1111/j.1460-2466.2010.01487.x>
- Bazeley, P. (2003). Defining early career in research. *Higher Education*, *45*(3), 257–279.
- Bedeian, A. G., & Feild, H. S. (1980). Academic stratification in graduate management programs: Departmental prestige and faculty hiring patterns. *Journal of Management*, *6*(2), 99–115. <https://doi.org/10.1177/014920638000600201>
- Bedeian, A. G., Cavazos, D. E., Hunt, J. G., & Jauch, L. R. (2010). Doctoral degree prestige and the academic marketplace: A study of career mobility within the management discipline. *Academy of Management Learning & Education*, *9*(1), 11–25. <https://doi.org/10.5465/amle.9.1.zqr11>
- Bol, T., De Vaan, M., & Van De Rijt, A. (2018). The Matthew effect in science funding. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(19), 4887–4890. <https://doi.org/10.1073/pnas.1719557115>
- Buela-Casal, G., Paz Bermúdez, M., Sierra, J. C., Quevedo-Blasco, R., Castro, Á., & Guillén-Riquelme, A. (2012). Ranking de 2011 en producción y productividad en investigación de las universidades públicas Españolas. *Psicothema*, *24*(4), 505–515.
- Burris, V. (2004). The academic caste system: Prestige hierarchies in PhD exchange networks. *American Sociological Review*, *69*(2), 239–264. <https://doi.org/10.1177/000312240406900205>
- Capponi, G., & Frenken, K. (2021). Does academic inbreeding affect scientific success? Paper to be presented at DRUID21 Copenhagen Business School, Copenhagen [https://conference.druid.dk/acc\\_papers/q6gp2qtessqddcj7uirgdbwemf3faj.pdf](https://conference.druid.dk/acc_papers/q6gp2qtessqddcj7uirgdbwemf3faj.pdf).
- Chan, K. C., Chen, C. R., & Steiner, T. L. (2002). Production in the finance literature, institutional reputation, and labor mobility in academia: A global perspective. *Financial Management*, *31*(4), 131. <https://doi.org/10.2307/3666177>
- Clauset, A., Arbesman, S., & Larremore, D. B. (2015). Systematic inequality and hierarchy in faculty hiring networks. *Science Advances*, *1*(1), e1400005. <https://doi.org/10.1126/sciadv.1400005>
- Cowan, R., & Rossello, G. (2018). Emergent structures in faculty hiring networks, and the effects of mobility on academic performance. *Scientometrics*, *117*(1), 527–562. <https://doi.org/10.1007/s11192-018-2858-8>
- Crane, D. (1970). The academic marketplace revisited: A study of faculty mobility using the Cartter ratings. *American Journal of Sociology*, *75*(6), 953–964. <https://doi.org/10.1086/224848>
- Cruz-Castro, L., & Sanz-Menéndez, L. (2010). Mobility versus job stability: Assessing tenure and productivity outcomes. *Research policy*, *39*(1), 27–38.
- Cyrenne, P., & Grant, H. (2009). University decision making and prestige: An empirical study. *Economics of Education Review*, *28*(2), 237–248. <https://doi.org/10.1016/j.econedurev.2008.06.001>
- de Solla Price, D. J. (1965). Networks of scientific papers. *Science*, *149*(3683), 510–515.
- Debackere, K., & Rappa, M. A. (1995). Scientists at major and minor universities: Mobility along the prestige continuum. *Research Policy*, *24*(1), 137–150. [https://doi.org/10.1016/0048-7333\(95\)98447-3](https://doi.org/10.1016/0048-7333(95)98447-3)
- Demeter, M., & Toth, T. (2020). The world-systemic network of global elite sociology: The western male monoculture at faculties of the top one-hundred sociology departments of the world. *Scientometrics*, *124*(3), 2469–2495. <https://doi.org/10.1007/s11192-020-03563-w>
- Farys, R., & Wolbring, T. (2021). Matthew effects in science and the serial diffusion of ideas testing old ideas with new methods. *Quantitative Science Studies*, *2*, 1–36.
- Finn, MG. (2010). Stay rates of foreign doctorate recipients from us universities, 2007. Tech. rep., Oak Ridge Inst. for Science and Education (ORISE), Oak Ridge.
- Fogarty, T. J., & Saftner, D. V. (1993). Academic department prestige: A new measure based on the doctoral student labor market. *Research in Higher Education*, *34*(4), 427–449. <https://doi.org/10.1007/BF00919183>

- Fox, M. F. (1983). Publication productivity among scientists: A critical review. *Social Studies of Science*, 13(2), 285–305. <https://doi.org/10.1177/030631283013002005>
- Gerhards, J., Hans, S., & Drewski, D. (2018). Global inequality in the academic system: Effects of national and university symbolic capital on international academic mobility. *Higher Education*, 76(4), 669–685. <https://doi.org/10.1007/s10734-018-0231-8>
- Giorgi, G. (2012). Workplace bullying in academia creates a negative work environment. An Italian study. *Employee Responsibilities and Rights Journal*, 24(14), 261–275. <https://doi.org/10.1007/s10672-012-9193-7>
- Gorelova, O., & Yudkevich, M. (2015). *Academic inbreeding: State of the literature* (pp. 17–44). Palgrave Macmillan.
- Gras, N. (2018). La evaluación de la investigación y su relación con la producción de conocimiento: un análisis estructural para el sistema nacional de investigadores de México. Phd dissertation UAM-Xochimilco, Mexico City.
- Headworth, S., & Freese, J. (2016). Credential privilege or cumulative advantage? Prestige, productivity, and placement in the academic sociology job market. *Social Forces*, 94(3), 1257–1282. <https://doi.org/10.1093/sf/sov102>
- Henning, M. A., Zhou, C., Adams, P., Moir, F., Hobson, J., Hallett, C., & Webster, C. S. (2017). Workplace harassment among staff in higher education: A systematic review. *Asia Pacific Education Review*, 18(4), 521–539. <https://doi.org/10.1007/s12564-017-9499-0>
- Horta, H., Veloso, F. M., & Grediaga, R. (2010). Navel gazing: Academic inbreeding and scientific productivity. *Management Science*, 56(3), 414–429.
- Horta, H., Cattaneo, M., & Meoli, M. (2018). PhD funding as a determinant of PhD and career research performance. *Studies in Higher Education*, 43(3), 542–570. <https://doi.org/10.1080/03075079.2016.1185406>
- Kalaitzidakis, P., Stengos, T., & Mamuneas, T. P. (2003). Rankings of academic journals and institutions in economics. *Journal of the European Economic Association*, 1(6), 1346–1366. <https://doi.org/10.1162/154247603322752566>
- Keashly, L., & Neuman, J. H. (2010). Faculty experiences with bullying in higher education. *Administrative Theory & Praxis*, 32(1), 48–70. <https://doi.org/10.2753/atp1084-1806320103>
- Keyserlingk, L. V., Becker, M., & Jansen, M. (2020). Do social comparisons matter for university major choices? a longitudinal study from a gender perspective do social comparisons matter for university major choices? A longitudinal study from a gender perspective. *International Journal of Gender, Science and Technology*, 12(1), 46–64.
- Long, J. S. (1978). Productivity and academic position in the scientific career author (s). *American Sociological Review*, 43(6), 889–908.
- Lang, L., Wang, Y., Chen, Q., & Zheng, T. (2019). University ranking based on faculty hiring network with minimum weighted violation rankings. *International Journal of Modern Physics C*, 30(7), 1940017. <https://doi.org/10.1142/S0129183119400175>
- Laudel, G., & Gläser, J. (2008). From apprentice to colleague: The metamorphosis of Early Career Researchers. *Higher Education*, 55(3), 387–406. <https://doi.org/10.1007/s10734-007-9063-7>
- Laurance, W. F., Useche, D. C., Laurance, S. G., & Bradshaw, C. J. (2013). Predicting publication success for biologists. *BioScience*, 63(10), 817–823. <https://doi.org/10.1525/bio.2013.63.10.9>
- Lee, D. H. (2019). Predicting the research performance of early career scientists. *Scientometrics*, 121(3), 1481–1504. <https://doi.org/10.1007/s11192-019-03232-7>
- Levy, H. (1992). Stochastic dominance and expected utility: Survey and analysis. *Management Science*, 38(4), 555–593. <https://doi.org/10.1287/mnsc.38.4.555>
- Li, R., Zhang, H., Zhu, X., & Li, N. (2020). Does employee humility Foster performance and promotability? Exploring the mechanisms of LMX and peer network centrality in China. *Human Resource Management*, 60(3), 399–413. <https://doi.org/10.1002/hrm.22025>
- Long, J. S., Allison, P. D., & McGinnis, R. (1979). Entrance into the academic career. *American Sociological Review*, 44(5), 816. <https://doi.org/10.2307/2094529>
- Lopez-Olmedo, R., Marmolejo-Leyva, R., Perez-Angon, M. A., Villa-Vazquez, L. L., & Zayago-Lau, E. (2017). The role of public policies in the decentralization process of Mexican science and the formation of new researchers in institutions outside the Mexico City area. *Scientometrics*, 112(3), 1343–1366. <https://doi.org/10.1007/s11192-017-2423-x>
- Mai, B., Liu, J., & González-Bailón, S. (2015). Network effects in the academic market: Mechanisms for hiring and placing PhDs in communication (2007–2014). *Journal of Communication*, 65(3), 558–583. <https://doi.org/10.1111/jcom.12158>

- Marsh, H. W. (1991). Failure of high-ability high schools to deliver academic benefits commensurate with their students' ability levels. *American Educational Research Journal*, 28(2), 445–480. <https://doi.org/10.3102/00028312028002445>
- Marsh, H. W., & Hau, K. T. (2003). Big-fish-little-pond effect on academic self-concept: A cross-cultural (26-country) test of the negative effects of academically selective schools. *American Psychologist*, 58(5), 364–376. <https://doi.org/10.1037/0003-066X.58.5.364>
- Marsh, H. W., Seaton, M., Trautwein, U., Lüdtke, O., Hau, K. T., O'Mara, A. J., & Craven, R. G. (2008). The big-fish-little-pond-effect stands up to critical scrutiny: Implications for theory, methodology, and future research. *Educational Psychology Review*, 20(3), 319–350. <https://doi.org/10.1007/s10648-008-9075-6>
- McKay, R., Arnold, D. H., Fratzl, J., & Thomas, R. (2008). Workplace bullying in academia: A Canadian study. *Employee Responsibilities and Rights Journal*, 20(2), 77–100. <https://doi.org/10.1007/s10672-008-9073-3>
- Merton, R. K. (1968). The Matthew effect in science. *Science*, 159(3810), 56–63. <https://doi.org/10.1126/science.159.3810.56>
- Merton, R. K. (1988). The Matthew effect in science, ii: Cumulative advantage and the symbolism of intellectual property. *Isis*, 79(4), 606–623. <https://doi.org/10.1086/354848>
- Miller, C. C., Click, W. H., & Cardinal, L. B. (2005). The allocation of prestigious positions in organizational science: Accumulative advantage, sponsored mobility, and contest mobility. *Journal of Organizational Behavior*, 26(5), 489–516. <https://doi.org/10.1002/job.325>
- Moodie, G. (2009). Four tiers. *Higher Education*, 58(3), 307–320. <https://doi.org/10.1007/s10734-008-9195-4>
- Nevin, A. D. (2019). Academic hiring networks and institutional prestige: A case study of Canadian sociology. *Canadian Review of Sociology*, 56(3), 389–420. <https://doi.org/10.1111/cars.12252>
- Olcay, G. A., & Bulu, M. (2017). Is measuring the knowledge creation of universities possible?: A review of university rankings. *Technological Forecasting and Social Change*, 123, 153–160. <https://doi.org/10.1016/j.techfore.2016.03.029>
- Oyer, P. (2006). Initial labor market conditions and long-term outcomes for economists. *Journal of Economic Perspectives*, 20(3), 143–160.
- Oyer, P. (2008). Ability and employer learning: Evidence from the economist labor market. *Journal of the Japanese and International Economies*, 22(2), 268–289. <https://doi.org/10.1016/j.jjie.2007.01.001>
- Perc, M. (2014). The Matthew effect in empirical data. *Journal of The Royal Society Interface*, 11(98), 20140378.
- Pinheiro, D. L., Melkers, J., & Newton, S. (2017). Take me where I want to go: Institutional prestige, advisor sponsorship, and academic career placement preferences. *PLoS ONE*, 12(5), e0176977. <https://doi.org/10.1371/journal.pone.0176977>
- Reyes, R. G., & Suriñach, J. (2013). Las evaluaciones internas del SNI: Coherencias o coincidencias. *Secuencia*, 1(83), 181. <https://doi.org/10.18234/secuencia.v0i83.1159>
- Reyes Ruiz, G., & Suriñach, J. (2015). Análisis sobre la Evolución del Sistema Nacional de Investigadores (SNI) de México. *Investigación Administrativa*, 44(115), 55–69.
- Rivero, P. J., & Peña, A. P. T. (2020). Los programas de repatriación de investigadores de argentina y México. *Controversias y Concurrencias Latinoamericanas*, 12(21), 79–105.
- Rosman, T., Mayer, A. K., Lechner, N., & Krampen, G. (2020). Putting big fish into a bigger pond: Self-concept changes in psychology undergraduate entrants. *Journal of Further and Higher Education*, 44(1), 14–28. <https://doi.org/10.1080/0309877X.2018.1493095>
- Salchegger, S. (2016). Selective school systems and academic self-concept: How explicit and implicit school-level tracking relate to the big-fish-little-pond effect across cultures. *Journal of Educational Psychology*, 108(3), 405–423. <https://doi.org/10.1037/edu0000063>
- Sherf, E. N., & Morrison, E. W. (2019). I do not need feedback! Or do i? Self-efficacy, perspective taking, and feedback seeking. *Journal of Applied Psychology*, 105(2), 146. <https://doi.org/10.1037/apl0000432>
- Slavova, K., Fosfuri, A., & De Castro, J. O. (2016). Learning by hiring: The effects of scientists' inbound mobility on research performance in academia. *Organization Science*, 27(1), 72–89. <https://doi.org/10.1287/orsc.2015.1026>
- Stapel, D. A., & Koomen, W. (2005). Competition, cooperation, and the effects of others on me. *Journal of Personality and Social Psychology*, 88(6), 1029–1038. <https://doi.org/10.1037/0022-3514.88.6.1029>
- Su, X. (2011). Postdoctoral training, departmental prestige and scientists' research productivity. *Journal of Technology Transfer*, 36(3), 275–291. <https://doi.org/10.1007/s10961-009-9133-3>
- Teplitskiy, M., Duede, E., Menietti, M., & Lakhani, KR. (2020). Citations systematically misrepresent the quality and impact of research articles: Survey and experimental evidence from thousands of citers. [arXiv:2002.10033](https://arxiv.org/abs/2002.10033).

- Wang, J. (2014). Unpacking the Matthew effect in citations. *Journal of Informetrics*, 8(2), 329–339. <https://doi.org/10.1016/j.joi.2014.01.006>
- Wang, Y., Jones, B. F., & Wang, D. (2019). Early-career setback and future career impact. *arXiv*, 10(1), 1–10. <https://doi.org/10.2139/ssrn.3353841>
- Waters, J. L., & Leung, M. W. (2017). Domesticating transnational education: Discourses of social value, self-worth and the institutionalisation of failure in ‘meritocratic’ hong kong. *Transactions of the Institute of British Geographers*, 42(2), 233–245.
- Way, S. F., Morgan, A. C., Larremore, D. B., & Clauset, A. (2019). Productivity, prominence, and the effects of academic environment. *Proceedings of the National Academy of Sciences*, 116(22), 10729–10733.
- Weiss, H. M., & Knight, P. A. (1980). The utility of humility: Self-esteem, information search, and problem-solving efficiency. *Organizational Behavior and Human Performance*, 25(2), 216–223.
- Whelpley, C. E., & McDaniel, M. A. (2016). Self-esteem and counterproductive work behaviors: A systematic review. *Journal of Managerial Psychology*, 31(4), 850–863. <https://doi.org/10.1108/JMP-01-2014-0008>
- Williamson, I. O., & Cable, D. M. (2003). Predicting early career research productivity: The case of management faculty. *Journal of Organizational Behavior*, 24(1), 25–44. <https://doi.org/10.1002/job.178>
- Zhu, Y., & Yan, E. (2017). Examining academic ranking and inequality in library and information science through faculty hiring networks. *Journal of Informetrics*, 11(2), 641–654. <https://doi.org/10.1016/j.joi.2017.04.007>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.