EconomiA

251

Received 14 August 2024
Revised 11 March 2025
9 August 2025
Accepted 16 August 2025

Academic credentials and career outcomes: evidence from Brazilian economists

Francis Petterini
Department of Economics, Federal University of Santa Catarina, Florianópolis, Brazil, and
Victor Buttignon
Court of Accounts of Santa Catarina, Florianópolis, Brazil

Abstract

Purpose – We analyze how academic credentials relate to career outcomes using linked microdata on Brazilian economists, focusing on two credentials (admission test scores and program prestige) and multiple outcomes spanning the labor market (e.g. wages) and research productivity (e.g. publication incidence).

Design/methodology/approach – We follow 888 master's graduates over nearly a decade to track their career development. To address endogeneity and sample-selection issues, we estimate a set of Heckman selection models with endogenous regressors.

Findings – Wages are positively associated with admission test scores but not with program prestige. Research productivity shows mixed associations with these credentials. Overall, the estimates are suggestive of – but do not establish – a sorting mechanism whereby a larger share of high-scoring students from top programs select into non-academic, higher-paying careers, whereas graduates from lower-ranked programs are relatively more likely to pursue academic careers within Brazil.

Research limitations/implications – Partial observability and the context-specific nature of the data limit the generalizability of our findings; future work should lengthen the observation window and follow additional cohorts.

Practical implications – Findings inform applicants, programs and policymakers about potential trade-offs between prestige and career paths.

Social implications – Insights into sorting between academic and non-academic careers can inform policies to retain research talent.

Originality/value – We provide novel evidence on Brazilian economists' early careers using linked administrative and bibliometric microdata, combining sample-selection correction with endogenous regressors.

Keywords Academic credentials, Career outcomes, Brazilian economists, Heckman model with endogeneity, Graduate education, Signaling mechanism

Paper type Research article

1. Introduction

Academic credentials – such as entrance exam scores and institutional prestige – are widely regarded as proxies for individual ability, shaping both educational trajectories and career outcomes. A seminal contribution to this discussion is the study by Krueger and Wu (2000), which tracked 325 U.S. economics graduates and found a positive correlation between entrance exam performance and later success. The authors argue that higher admission scores may be shaped by unobserved factors such as cognitive skills, and that these latent

JEL Classification — C34, I23, J24, J31

Funding: This work was supported by the National Council for Scientific and Technological Development (CNPq) – grant number 308582/2021-7.

Declaration of competing interest: The authors declare no competing interests.

EconomiA
Vol. 27 No. 2, 2026
pp. 251-272
Emerald Publishing Limited
e-ISSN: 2358-2820
p-ISSN: 1517-7580
DOI 10.1108/ECON-08-2024-0116

characteristics help explain the observed relationship. An elite academic environment, they argue, may amplify the effects of such unobserved characteristics.

Extending this line of research, Dale and Krueger (2002) and Athey, Katz, Krueger, Levitt, and Poterba (2007) used richer U.S. datasets to explore the driving forces behind these associations. Their findings suggest that career outcomes are shaped by two distinct channels related to academic credentials. First, students with higher unobserved ability and/or better preparation – often combined with a favorable family background – are more likely to access elite institutions, partly because they tend to score higher on entrance exams. Second, the mere fact of having attended a prestigious school can serve as a strong signal to employers, regardless of actual education.

From this perspective, admission scores may serve as proxies for individual ability, while enrollment in a top-ranked program functions as a proxy for institutional signals that shape expectations and opportunities in the labor market. Evidence from the United States and Europe supports this view. Studies show that graduates from prestigious programs – especially those placed in elite institutions – tend to achieve higher quality-adjusted research output, benefit from stronger international job placement prospects, and are more likely to secure positions in top academic departments (Cardoso, Guimarães, & Zimmermann, 2010; Chen, Liu, & Billger, 2013; Sullivan, Dubnicki, & Dutkowsky, 2018).

Institutional rank also interacts with advising quality, professional networks, and hiring connections to boost productivity and career outcomes (García-Suaza, Otero, & Winkelmann, 2020; Hadlock & Pierce, 2021), while collaborative networks fostered in such environments are associated with greater publication success, albeit with uneven rewards across genders (Hussey, Murray, & Stock, 2022). Together, these findings indicate that prestige operates not merely as a status marker but as a multidimensional signal shaping graduates' expectations and opportunities in both academic and broader labor markets.

In the case of Brazil, empirical evidence on the relationship between academic credentials and career outcomes remains limited, particularly among professionals in the field of economics. Notable exceptions include Rocha *et al.* (2021) and Pereda *et al.* (2023), who examine gender disparities in academic outcomes, and Estevan and Santos (2022), who documents that attending a more selective economics master's program increases the likelihood of enrolling in doctoral programs abroad, though not in top-tier universities.

This paper helps to fill this gap by examining whether academic credentials predict the career outcomes of professional economists trained in Master's programs in Brazil. We focus on two key indicators of academic credentials: performance on the national standardized entrance exam (ANPEC) and a proxy for institutional prestige based on the official quality assessment published by Brazil's Ministry of Education (CAPES program ranking). The former serves as a proxy for unobserved ability, while the latter captures the signaling effect of institutional affiliation.

In terms of career trajectories, we adopt a novel approach by tracking 888 students who took the ANPEC exam in 2009 (415 students) or 2010 (473 students), following their labor market outcomes through 2018 and their academic output through 2023. To this end, we link microdata from the ANPEC exam and the CAPES program ranking to three national administrative sources.

First, the Comprehensive Registry of Employment in the Formal Sector (RAIS), maintained by Brazil's Ministry of Labor, provides detailed information on individuals' wages, occupations, and employers. Second, the Lattes platform – a nationwide academic CV system administered by a federal science agency and extensively used in research evaluation and funding decisions – offers structured data on researchers' educational trajectories, institutional affiliations, and scholarly output. Third, we incorporate journal quality information from the Qualis classification system, which ranks scientific publications according to their perceived quality and relevance within specific academic fields.

Based on this dataset, our goal is to estimate the relationship between academic credentials and career outcomes, conditional on other observable characteristics. A key challenge is the potential endogeneity of academic credentials – particularly the proxy for institutional signaling – since signaling effects may themselves be correlated with other unobserved attributes, such as access to professional networks, that may directly influence career outcomes. Estimates may also be affected by sample selection bias arising from the non-random observability of outcomes, which is inherent to this type of analysis. These compounding sources of endogeneity and selection bias pose serious threats to identification and underscore the need for a credible empirical strategy to isolate the impact of institutional prestige.

Considering these challenges, we estimate a system of recursive simultaneous equations that extends the standard Heckman correction framework by incorporating endogenous regressors, explicitly modeling correlated error terms, and accommodating mixed outcome types – including both continuous variables (e.g. wages) and categorical outcomes (e.g. journal quality rankings). This approach allows for the estimation and comparison of conditional expectations of outcomes given academic credentials and other covariates, thereby assessing their association while accounting for key sources of bias that may confound causal interpretation.

Our analysis reveals heterogeneous associations between academic credentials and career outcomes. Higher ANPEC exam scores are positively associated with subsequent wages, supporting the interpretation that these scores indicate cognitive ability or motivation valued in the labor market. In contrast, the relationship between ANPEC scores and research productivity is sensitive to the specific definition of productivity adopted. CAPES program rankings show no statistically significant effect on wages and tend to be negatively associated with the volume of publications, while exhibiting a positive association with indicators of publication quality. Moreover, individuals from lower-ranked programs appear more likely to pursue academic careers as university faculty within Brazil.

While this pattern may seem counterintuitive at first, it is consistent with a plausible career-sorting mechanism: graduates from higher-ranked programs often obtain higher-paying, non-academic positions with fewer requirements for publishing in academic journals, although when such publications occur, they tend to be of higher quality. In contrast, those from lower-ranked programs are more likely to follow academic paths in domestic institutions, where publishing is both encouraged and often required for career advancement.

However, this career-sorting interpretation should be treated with caution, as a key limitation of our study is the inability to observe individuals working as independent contractors or sole proprietors (e.g. under Brazil's PJ regime), or those who pursued doctoral studies or employment abroad. These alternative career paths may be systematically associated with both academic credentials and labor market or academic outcomes. Moreover, our analysis focuses on only two graduate cohorts and relies on a research productivity measure that is self-reported and not consistently updated, potentially introducing measurement error. As a result, our findings may not reflect the full range of career trajectories, thus limiting their external validity.

Despite these limitations – and the empirical strategy designed to partially address them – this study contributes to the literature in two main ways. First, to our knowledge, it provides the first empirical assessment of whether academic credentials predict career outcomes for economists in Brazil – a large country with centralized systems for graduate admissions and program evaluation – offering a test of how well-established findings from other contexts apply in this setting. Second, it presents a novel linked administrative microdata set that brings together admission scores, program rankings, wage records, and publication data – a resource that can inform and support future empirical research.

The remainder of the paper is organized as follows. Section 2 describes the data sources. Section 3 presents the descriptive statistics. Section 4 outlines the econometric strategy. Section 5 reports the empirical findings. Section 6 concludes and discusses future research. Additionally, the online supplementary material presents further statistical and econometric results, as well as detailed derivations of expressions arising from the modeling framework adopted in this study.

2. Data sources

This section briefly describes the microdata sources used in the study. The goal is both to support future efforts to build alternative or more comprehensive datasets and to provide an overview for readers unfamiliar with these Brazilian databases.

2.1 ANPEC

We begin with records from ANPEC (National Association of Centers for Postgraduate Studies in Economics), which has administered a standardized national admissions exam since the mid-1970s. Today, all but one of over 50 graduate programs use the ANPEC exam, typically as the sole or primary admissions criterion. For further details on the exam and its administration, see Fernández and Suprinyak (2018), Petterini (2020).

The current exam format was established in the early 1990s and covers Macroeconomics, Microeconomics, Mathematics, Statistics, and the Brazilian Economy at basic to intermediate levels, with some complementary tests. Held annually in September, the exam determines admissions for the following academic year.

A standardized performance score is commonly used to rank candidates. Each subject grade is normalized (mean zero, standard deviation one), and the unweighted average of these scores is multiplied by 1,000 for presentation. We adopt this score without the multiplier, interpreting it as the number of standard deviations above or below the mean.

Admissions are determined through a centralized matching process based on mutual preferences of candidates and institutions. Higher-ranked programs typically attract top scorers, though some high performers opt for lower-ranked programs due to personal factors such as job location or family. In practice, candidates who score more than two standard deviations above the mean almost invariably enroll in elite programs. Those scoring between one and two standard deviations may or may not enter top programs, depending on other factors, while applicants below one standard deviation are, with few exceptions, placed in lower-ranked programs – see supplementary material for details.

While microdata are available since the late 1990s, detailed information on candidates and placements is only available from 2009 onward. We therefore focus on the 2009 and 2010 cohorts, corresponding to admissions in 2010 and 2011. This allows sufficient time to observe post-graduate career outcomes.

Our sample includes students admitted to Master's programs, as the ANPEC exam was not widely used for doctoral admissions at the time. For these students, we compiled test scores, program placements, and individual characteristics. Access to detailed microdata is restricted under LGPD and requires formal request, a signed confidentiality agreement, and a clearly defined research project.

2.2 CAPES

Once the student microdata were compiled, we incorporated a measure of program prestige based on information from CAPES (Coordination for the Improvement of Higher Education Personnel), a federal agency under Brazil's Ministry of Education responsible for evaluating graduate education. Since the mid-1970s, CAPES has coordinated a national framework for accreditation and quality assessment of graduate programs.

A central feature of this framework is a recurring evaluation cycle in which all graduate programs are rated on a 1-to-7 scale, based on standardized criteria such as faculty credentials, student outcomes, research output, internationalization, and infrastructure. Scores of 6 or 7 indicate international-level excellence; scores below 3 result in disqualification. Accordingly, we assigned each student the CAPES score of their program at enrollment, which serves as our proxy for institutional prestige.

2.3 RAIS

The next step was to link data on individuals' professional careers using RAIS – Brazil's Annual Social Information Report, an administrative database maintained by the Ministry of Labor. RAIS is based on mandatory declarations submitted annually by all formal-sector employers and covers every registered employment relationship in the country.

For each employee, RAIS provides information on wages, occupations, employer characteristics, contract type, job tenure, and demographic attributes. Crucially, it includes identifiers that allow individuals to be tracked over time and linked to other administrative records.

As with the ANPEC data, access to detailed RAIS microdata is regulated under Brazil's LGPD and requires a formal request, including a signed confidentiality agreement and a justified research purpose. For this study, we obtained authorization to use RAIS data through 2018, which marks the endpoint of our observation window.

2.4 Lattes

We also required data on students' subsequent academic trajectories. For this, we relied on Lattes – Brazil's national academic CV platform, maintained by the National Council for Scientific and Technological Development (CNPq). The platform aggregates curricula vitae of individuals involved in research, teaching, and academic activities. It is widely used in evaluation processes, including funding decisions, graduate program assessments, and hiring or promotion in higher education institutions. Each CV is publicly accessible and self-reported, with users legally responsible for the information provided. The platform includes detailed data on academic output – such as journal articles, books, conference papers, and other publications. Entries are date-stamped and linked to metadata such as journal titles and co-authors.

In this study, we use Lattes to construct publication profiles. We systematically collected and tabulated publication data in the second half of 2023 using web scraping techniques. This allows us to track research output over time and enrich the dataset with indicators of academic productivity.

2.5 Qualis

To complete our dataset, we assessed the quality of the publications listed on the Lattes platform using the Qualis classification system – an official journal ranking developed and maintained by CAPES. Qualis organizes academic journals into hierarchical strata based on their perceived quality and relevance within each academic field, typically ranging from A1 (highest) and A2, through B1 to B5, and down to C (lowest). As the classification applies only to journal articles – excluding books and other publication types – we restricted our research productivity measure to journal outputs.

To convert these qualitative strata into an individual quantitative publication score, we applied the official weighting rule established by CAPES at the time:

$$\begin{aligned}\text{publication score} &= 100\times\#A1 + 80\times\#A2 + 60\times\#B1 + 40\times\#B2 \\ &\quad + 25\times\#B3 + 15\times\#B4 + 10\times\#B5 + 5\times\#C\end{aligned} \tag{1}$$

where $\#A1, \ldots, \#C$ denote the number of articles in each Qualis stratum.

3. Descriptive statistics

Table 1 presents summary statistics for labor market and academic outcomes. Panel I describes selected variables across graduate programs, grouped by their CAPES rank (column A), which ranges from 3 (lowest) to 7 (highest) – meaning that none of the ANPEC-affiliated programs in

Table 1. Summary statistics of labor market, academic outcomes, and control variables

Panel I

(A) Rank	(B) Observations	(C) Scores	(D) RAIS (%)	(E) Wages	(F) Professor (%)	(G) Lattes (%)	(H) Publication (%)	(I) A1-A2	(J) B1-B2
3	108	−0.4	41.7	14.6 (15.2)	37.8	52.8	56.1	4.7	18.3
4	333	0.1	51.7	15.9 (9.4)	42.4	59.5	54.0	9.1	38.6
5	271	0.3	47.2	19.1 (19.0)	40.6	58.3	59.5	13.1	36.6
6	99	1.4	50.5	24.1 (16.1)	20.0	62.6	45.2	14.6	52.4
7	77	1.9	40.3	25.7 (17.9)	19.4	54.5	26.2	15.0	50.0
Total	888	0.4	48.0	18.4 (15.2)	37.1	58.2	52.6	10.0	38.1

Panel II

Binary variable	Mean (%)	Binary variable	Mean (%)
Male	69.1	Lived in same state as program	65.7
White	73.1	Primary employment as economist(ii)	19.5
Started Master's before age 30	90.9	Primary employment as analyst(iii)	12.7
Undergraduate in Economics	89.5	Primary employment as other occupation	30.7
Undergraduate in STEM(i)	5.1	Primary employment in the public sector	39.4
Public undergraduate institution	76.9	Completed a doctorate	50.0
Single at time of Master's	93.5	Regularly updates academic CV	17.0

Note(s): *Panel I:* (A) Program rank based on CAPES evaluation. (B) Number of observations per rank. (C) Mean ANPEC score (in standard deviations from overall mean). (D) Share of students matched in RAIS. (E) Mean of highest monthly wages in RAIS (R$ 1,000; 2024), with SD in parentheses. (F) Share of RAIS-matched students classified as professors, based on CBO codes: 231205, 231210, 232115, 233115, 234115, 234120, 234310, 234405, 234505, 234510, 234515, 234520, 234725, 234770, 234805, 234810, 331205, 331305. (G) Share of students found on Lattes. (H) Share of Lattes-identified students with at least one journal article. (I) Share of publishing students with articles classified as Qualis A1 or A2. (J) Same as (I), but for Qualis B1 or B2. *Panel II:* (i) STEM stands for Science, Technology, Engineering, and Mathematics fields. (ii) Economists include CBO codes: 251205, 251215, 251220, 251225. (iii) Analysts: 142330, 142335, 203210, 203310, 203505, 203510, 212410, 252405, 252515, 252540, 252545, 253125

Source(s): The authors

our dataset were disqualified. The sample comprises 888 observed individuals, unevenly distributed across ranks, with the largest share enrolled in rank 4 programs (column B).

Column (C) shows the mean ANPEC score among students, standardized relative to the sample mean. That is, a value of 0 indicates an average performance, 1 corresponds to one standard deviation above the average, −1 to one standard deviation below, and so forth. As expected, average scores increase monotonically with program rank, from −0.4 in rank 3 to 1.9 in rank 7, consistent with the use of the ANPEC exam in selective admissions processes. The overall sample mean is 0.4 rather than zero because applicants with very low scores are typically not admitted to any program and therefore do not appear in the data.

Column (D) presents the RAIS match rate, indicating the percentage of students with identifiable labor market records. Match rates range from 40.3% to 51.7%, with no clear trend across program ranks. This may reflect differences in formal employment rates or limitations in correctly identifying individuals in the labor market.

Essentially, RAIS does not cover students who moved abroad to study and/or work, nor those who operate as sole proprietors, either as independent contractors or as individual microentrepreneurs (referred to in Brazil as PJ). These are common legal arrangements through which professionals provide services by registering a personal business entity, rather than holding a standard formal job contract. Such arrangements are frequently adopted by skilled professionals, particularly in technical or consulting roles. Therefore, there is a potential issue of sample selection bias in the analysis, which will be addressed later in the paper.

Column (E) reports the mean of the highest monthly wages in RAIS, with standard deviations in parentheses. This variable serves as our main indicator of labor market compensation. We rely on the maximum observed monthly wage – rather than the mean or median across months – to better capture peak earning potential, which is particularly relevant for individuals with intermittent contracts or short-term academic appointments. Furthermore, all wages were first converted into hourly equivalents expressed in monthly multiples of the minimum wage, and then rescaled to 2024 values using the prevailing minimum wage for that year.

Monthly wages increase substantially with program prestige: individuals from rank 3 programs earned an average of R$ 14,600, while those from rank 7 programs earned R$ 25,700. This pattern supports the notion that higher-ranked programs are associated with stronger labor market returns. Note also that standard deviations are relatively large, reflecting high intra-group variability. This dispersion likely captures heterogeneity in occupational trajectories, regional labor market conditions, and the distribution of individuals across public, private, and academic sectors.

Column (F) shows the proportion of matched individuals working as professors, based on the Brazilian Classification of Occupations (CBO). Interestingly, the share is around 40% for ranks 3, 4, and 5, compared to approximately 20% for ranks 6 and 7. This pattern may reflect differences in professional orientation, with lower-ranked programs placing a greater share of graduates in teaching positions.

It is also worth noting that academic salaries in Brazil tend to be lower than those earned in careers in the financial sector or in high-level government institutions, such as the Central Bank, National Treasury, or state finance secretariats. Microdata inspection confirms that graduates from top-ranked programs are disproportionately represented in these higher-paying positions, which helps explain the inverse relationship between program rank, and consequently wages, and the prevalence of academic employment.

Column (G) reports the proportion of students found on the Lattes platform. The share stays around 55–60% across ranks. It is worth noting that Lattes profiles are self-declared and often outdated. This limitation raises concerns about data quality and coverage. As a result, academic outcomes derived from Lattes may be subject to sample selection bias, particularly if individuals with non-academic trajectories or weaker research profiles are less likely to maintain an updated online CV.

Column (H) captures the share of Lattes-identified individuals with at least one publication in academic journals. This share remains relatively stable across ranks 3 to 5 (around 54–59%) but drops significantly for ranks 6 (45.2%) and 7 (26.2%), indicating that publication activity is not strictly increasing with program prestige. This pattern is consistent with the inverse relationship observed in Column (F), as publishing is a core component of academic careers and tends to be more prevalent among those employed in teaching and research positions.

Columns (I) and (J) detail the distribution of publication quality. Among those who publish, the share of students with at least one article in A1–A2 journals (the highest Qualis classifications) rises monotonically with program rank, reaching 15.0% in rank 7. The B1–B2 share (second-highest classifications) also increases, peaking at 52.4% in rank 6 before slightly declining. These results suggest that while higher-ranked programs do not necessarily lead to higher publication rates, they are associated with a greater proportion of top-tier publications. This pattern is plausible, given that students from top programs tend to have stronger unobserved skills – as reflected in their selection scores – and may also be more comfortable writing in English, which facilitates access to high-quality journals that are often international in scope.

Overall, Panel I in Table 1 indicates a positive association between program rank and entry exam scores, as well as between exam scores and labor market wages. By contrast, the relationship with academic output is more nuanced, reflecting heterogeneity in publication incidence, career orientation, and the quality of journals in which graduates publish. Higher-ranked programs tend to produce graduates with stronger academic publication records, although in relative terms lower-ranked programs appear to generate a higher volume of publications.

Panel II in Table 1 complements the previous information by reporting the means of other control variables included in the analysis. All variables are binary indicators equal to 1 if the described characteristic applies.

A majority of the sample were male (69.1%), and an even larger share identified as white (73.1%). Most individuals began their Master's program before the age of 30 (90.9%), indicating that graduate studies are typically pursued early in professional life. Regarding academic background, the vast majority held an undergraduate degree in Economics (89.5%), while a smaller fraction (5.1%) graduated in STEM fields – a term referring to Science, Technology, Engineering, and Mathematics.

A large share of participants (76.9%) attended a public university for their undergraduate studies, reflecting the prominent role of public higher education institutions in Brazil. Unlike in many other countries, a significant portion of the most prestigious and competitive universities in Brazil are publicly funded and tuition-free, with admission based on competitive entrance exams. Most students were single at the time of the Master's program (93.5%).

In terms of geographic proximity, 65.7% lived in the same state as their graduate institution. Although program rankings are generally correlated with national entrance exam scores, it is not uncommon for some high-performing individuals to choose lower-ranked programs in order to remain in a specific location, due to personal factors such as family ties or local job opportunities.

Regarding labor market outcomes, 19.5% of the sample reported primary employment as economists, 12.7% worked in analyst positions – which often involve consulting and financial activities – and 30.7% were engaged in other professional occupations. Notably, 37.1% held positions as professors, as shown earlier, highlighting the academic sector as a significant destination for graduates. In addition, 39.4% worked in the public sector, underscoring the government's role as a significant employer of graduate-trained economists in Brazil.

Furthermore, 50.0% had completed a doctorate by 2018, highlighting the Master's program as an important pathway to doctoral studies and potential academic careers. Only 17.0% regularly updated their Lattes CV within intervals shorter than one year during the 2018–2023 period, indicating limited engagement with ongoing academic self-reporting.

Taken together, the variables reported in Panel II portray a relatively homogeneous group in terms of demographic and educational background, but one that follows heterogeneous professional paths. The prevalence of early enrollment, degrees from selective public universities, and strong representation in both academic and public sector roles suggests that Master's programs in Economics in Brazil attract candidates who subsequently disperse into a diverse set of career trajectories. This diversity likely reflects a combination of individual preferences, geographic constraints, and the dual role of the Master's degree as preparation for doctoral studies and as a credential valued in specialized professional occupations.

4. Modeling

This section outlines the empirical strategy used to estimate the relationship between academic credentials and subsequent career outcomes. Specifically, we aim to quantify how entrance exam scores and graduate program prestige are associated with labor market performance, career choices (such as becoming a professor or completing a doctorate), and academic advancement (e.g. publishing scholarly work and the quality tier of those publications), conditional on observable characteristics.

Entrance exam scores and graduate program prestige enter as explanatory variables in the conditional expectation of career outcomes. The former serves as a plausibly exogenous proxy for intrinsic ability, as it is determined prior to program enrollment and is therefore unaffected by post-admission factors – a widely accepted assumption in the literature. In contrast, the latter may influence outcomes through its signaling role in both labor and academic markets, independent of any actual gains in knowledge or productivity derived from attending a top-tier institution.

Specifically, we assume that degrees from prestigious programs are often perceived as credible signals of a candidate's potential, which may enhance career opportunities regardless of true ability. As such, program prestige is potentially endogenous, as it may be correlated with unobserved traits such as social capital – including access to influential networks, informal mentoring relationships, or reputational spillovers from affiliation with prestigious institutions – all of which can also affect career success.

Another challenge is sample selection. Labor market outcomes are only observed for individuals successfully matched to RAIS administrative records, while academic outcomes are available only for those with publicly accessible Lattes CVs. Both sources are inherently incomplete and may systematically exclude individuals in non-standard labor arrangements or those working or studying abroad.

4.1 Baseline model

Our empirical strategy extends the Heckman selection framework to accommodate endogeneity and multiple types of outcomes – including continuous (e.g. wages), binary (e.g. doctoral completion or professor status), and ordered categorical variables (e.g. hierarchy of journal quality). The structural system is specified as:

$$y_1^* = \beta_1' X_1 + u_1 \tag{2}$$

$$y_2^* = \gamma_2 y_1 + \beta_2' X_2 + u_2 \tag{3}$$

$$y_3^* = \begin{cases} \gamma_3 y_1 + \beta_3' X_3 + u_3 & \text{if } y_2 = 1 \\ \textit{missing} & \text{if } y_2 = 0 \end{cases} \tag{4}$$

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho_{12} & \rho_{13}\sigma \\ \rho_{12} & 1 & \rho_{23}\sigma \\ \rho_{13}\sigma & \rho_{23}\sigma & \sigma^2 \end{bmatrix} \right) \tag{5}$$

Equation (2) defines the latent variable y_1^* underlying the binary indicator y_1, where $y_1 = 1$ if $y_1^* > 0$ (higher-prestige program) and $y_1 = 0$ otherwise. This variable serves as a central endogenous regressor in both the sample selection and outcome equations. The covariate vector X_1 includes a constant, the admission test score, and a set of instrumental variables for identification. The vector β_1 captures the structural effect of X_1, and u_1 is the associated error.

Equation (3) accounts for sample selection. The latent variable y_2^* generates the indicator y_2, equal to one if the individual is observed (i.e. matched to RAIS or Lattes) and zero otherwise. The coefficient γ_2 captures the impact of program prestige on the likelihood of observation. The vector X_2 includes the admission test score and controls (in part distinct from X_1), and β_2 captures their effects. u_2 is the structural disturbance.

Equation (4) models the outcome y_3^*, which can be:

(1) Continuous, when $y_3^* = y_3$ is directly observed (e.g. log wages).

(2) Ordered categorical, where $y_3 = j \in \{1, \ldots, J\}$ if $\xi_{j-1} < y_3^* \leq \xi_j$, with cutoffs ξ_j to be estimated, except for $\xi_0 = -\infty$ and $\xi_J = +\infty$. The binary outcome is a special case with $J = 2$ and $\xi_1 = 0$.

In all cases, X_3 includes the admission test score and other controls – partly distinct from the covariates in X_1 and X_2, as required for identification. β_3 captures their effects, γ_3 is the coefficient on y_1, and u_3 is the error term.

Equation (5) imposes the joint normality assumption on the structural disturbances. The correlation parameters ρ_{12}, ρ_{13}, and ρ_{23} capture unobserved factors linking selection mechanisms and outcomes. When all three correlations are equal to zero, the system simplifies to a standard model with independent equations and no endogeneity or selection concerns. The variances of u_1 and u_2 are normalized to one, consistent with the binary nature of y_1 and y_2, while the variance σ^2 of u_3 is freely estimated when y_3 is continuous and fixed to one otherwise. This model structure can be implemented using Stata's eregress or eoprobit commands, depending on the nature of the outcome variable.

4.2 Analysis of continuous outcomes

In the case of continuous outcomes, our goal is to estimate the conditional expectation of wages (for individuals observed in RAIS) and publication points (for individuals observed in Lattes), conditional on selection and covariates. This expectation is expressed as:

$$E(y_3|\, y_1, y_2 = 1, X) = \gamma_3 y_1 + \beta_3' X_3 + E(u_3|\, y_1, y_2 = 1, X) \tag{6}$$

where X denotes the full set of covariates, partitioned into X_1, X_2, and X_3 for each equation.

Consistent with Heckman-type models, the final term in equation (6), $E(u_3|X)$, is generally nonzero. Omitting this term results in biased estimates of $E(y_3|X)$, with an indeterminate direction of bias. To clarify, we consider the following expressions:

(1) If $y_1 = 1$, then $E(u_3|\, y_1 = 1, y_2 = 1, X) = E(u_3|\, u_1 > -\beta_1' X_1,\ u_2 > -\gamma_2 - \beta_2' X_2)$.

(2) If $y_1 = 0$, then $E(u_3|\, y_1 = 0, y_2 = 1, X) = E(u_3|\, u_1 \leq -\beta_1' X_1,\ u_2 > -\beta_2' X_2)$.

These expressions formally demonstrate how the correlation between u_3 and the latent selection disturbances u_1 and u_2 influences the conditional expectation of u_3. The direction of the inequalities, together with the inclusion of the parameter γ_2, induces distinct truncation regimes across values of y_1, which in turn modify the conditional density of y_3 and thereby affect the structure of $E(y_3|X)$ itself.

Under the joint normality assumption in equation (5), the conditional expectation $E(u_3|X)$ is given by the following expression:

$$E(u_3|\, y_1, y_2 = 1, X) = \sigma[\rho_{13}\lambda_1(y_1, X) + \rho_{23}\lambda_2(y_1, X)] \tag{7}$$

where λ_1 and λ_2 are correction terms in the form of inverse Mills ratios, constructed from the parameters and controls of the y_1 and y_2 equations – their functional forms are derived and discussed in detail in the supplementary material.

Equation (7) is an expansion of the classic correction, which adjusts for non-random sample selection in models with normally distributed errors. In the standard Heckman two-step framework, the inverse Mills ratio arises from the conditional expectation of a truncated normal distribution, when the outcome is observed only if a selection variable exceeds a threshold. In our setting, λ_1 accounts for selection into the treatment y_1, while λ_2 adjusts for subsequent selection into the observed sample $y_2 = 1$, conditional on y_1. Their inclusion ensures that the conditional expectation of the outcome error term u_3 is properly adjusted for endogenous selection processes, allowing for consistent estimation of the structural parameters in the outcome equation.

In addition, this structure permits estimation via either the 2-Step Heckman procedure or full information maximum likelihood (FIML). The 2-Step method estimates the recursive bivariate probit system first, obtaining consistent estimates of the correction terms λ_1 and λ_2, which are subsequently included in a linear outcome equation estimated by least squares.

When the outcome is not continuous, however, the 2-Step is not applicable due the lack of a closed-form correction. FIML, in turn, jointly estimates all structural parameters, exploiting the full covariance structure among disturbances.

Considering that y_3 is specified in logarithmic form, we focus on two semi-elasticity measures:

(1) Holding program prestige (y_1) constant, we evaluate how much $E(y_3|y_1, y_2 = 1, X)$ changes when the admission test score increases by one unit (i.e. one standard deviation above the reference level). This corresponds to the range where the proxy for ability begins to rise appreciably.

(2) Holding the admission test score constant, we assess the change in $E(y_3|y_1, y_2 = 1, X)$ when a student moves from a lower-prestige to a higher-prestige graduate program. Such transitions are particularly relevant for students whose entrance exam scores lie between one and two standard deviations above the mean.

These semi-elasticities follow directly from equations (6) and (7):

$$\begin{aligned} E(y_3|\, y_1, y_2 = 1, s+1) \quad & -E(y_3|\, y_1, y_2 = 1, s) = \beta_{3s} + \sigma[\rho_{13}(\lambda_1(y_1, s+1) - \lambda_1(y_1, s)) \\ & +\rho_{23}(\lambda_2(y_1, s+1) - \lambda_2(y_1, s))] \end{aligned} \tag{8}$$

$$\begin{aligned} E(y_3|\, y_1 = 1, y_2 = 1, s) \quad & -E(y_3|\, y_1 = 0, y_2 = 1, s) = \gamma_3 + \sigma[\rho_{13}(\lambda_1(1, s) - \lambda_1(0, s)) \\ & + \rho_{23}(\lambda_2(1, s) - \lambda_2(0, s))] \end{aligned} \tag{9}$$

where s represents the selection score, and β_{3s} is the corresponding coefficient in β_3 – all other covariates in X are held constant in both cases.

4.3 Analysis of ordered outcomes

When the outcome variable is ordinal, marginal effects correspond to changes in the conditional probability of each response category. Hence, we examine how the probability of attaining a specific outcome level varies with changes in admission test scores and program prestige.

As previously established, sample selection implies that the distribution of u_3 conditional on y_1 and y_2 is no longer standard normal. Rather, it follows a truncated normal distribution, and the probability of observing outcome level j must be evaluated conditional on the truncation regime induced by the selection mechanism:

$$\begin{aligned} & \Pr(y_3 = j|\, y_1, y_2 = 1, X) \\ & \quad = \frac{\Pr\left(\xi_{j-1} - \gamma_3 y_1 - \beta_3' X_3 < u_3 \leq \xi_j - \gamma_3 y_1 - \beta_3' X_3,\ \text{truncation region}\right)}{\Pr(\text{truncation region})} \end{aligned} \tag{10}$$

Then, consider the marginal effect of increasing the admission test score s by one unit, holding program prestige y_1, selection status y_2, and covariates X constant:

$$\begin{aligned} \Delta_s^{(j)} & = \Pr(y_3 = j|\, y_1, y_2 = 1, s+1, X) - \Pr(y_3 = j|\, y_1, y_2 = 1, s, X) \\ & = \frac{\Pr\left(c_{j-1}^{s+1} < u_3 \leq c_j^{s+1},\ \text{truncation}_{s+1}\right)}{\Pr(\text{truncation}_{s+1})} - \frac{\Pr\left(c_{j-1}^{s} < u_3 \leq c_j^{s},\ \text{truncation}_{s}\right)}{\Pr(\text{truncation}_{s})} \end{aligned} \tag{11}$$

where the upper and lower limits of the interval are given by $c_j^s = \xi_j - \gamma_3 y_1 - \beta_3' X_3$, which varies with s since s is included in the covariates X_3. Additionally, the truncation region changes as s affects the selection equations via X_1 and X_2.

Now consider the marginal effect of moving from a lower-to a higher-prestige program, holding s fixed:

$$\begin{aligned}\Delta_{y_1}^{(j)} &= \Pr(y_3 = j |\, y_1 = 1, y_2 = 1, X) - \Pr(y_3 = j |\, y_1 = 0, y_2 = 1, X)\\ &= \frac{\Pr\left(c_{j-1}^{y_1=1} < u_3 \leq c_j^{y_1=1},\ \text{truncation}_{y_1=1}\right)}{\Pr\left(\text{truncation}_{y_1=1}\right)} - \frac{\Pr\left(c_{j-1}^{y_1=0} < u_3 \leq c_j^{y_1=0},\ \text{truncation}_{y_1=0}\right)}{\Pr\left(\text{truncation}_{y_1=0}\right)},\end{aligned} \tag{12}$$

where $c_j^{y_1} = \xi_j - \gamma_3 y_1 - \beta_3' X_3$ shifts with y_1, and the truncation region adjusts accordingly due to the presence of $\gamma_2 y_1$ in the sample selection equation.

The expressions (11) and (12) make clear that marginal effects in the ordered case depend not only on structural parameters such as γ_3 and β_3, but also on how changes in s or y_1 affect the truncation region – that is, the region of support for (u_1, u_2) under which $y_2 = 1$. Since the distribution of u_3 is conditional on this selection, any variation in s or y_1 shifts the truncation boundaries and consequently alters the conditional probability associated with each outcome level. Additional derivations for $\Delta_s^{(j)}$ and $\Delta_{y_1}^{(j)}$ are provided in the supplementary material accompanying this article.

5. Results

Table 2 presents estimation results for equations (2)–(5), using both 2-Step and FIML procedures. The outcome equation is specified in terms of log wages and log publication points. Program prestige is operationalized as a binary indicator for graduate programs ranked 6 or 7 in the national evaluation system. We focus here on the core results derived from the main outcome specification, although alternative definitions of prestige and additional econometric specifications can be explored in the supplementary material – which includes the anonymized dataset and replication code.

In both estimation strategies and across both outcomes, only two covariates are consistently statistically significant in the first equation (y_1, program prestige): the standardized admission test score and an undergraduate degree in Economics – see columns (A), (B), (G), and (H). The strong association with test scores is unsurprising, as the entrance exam serves as the main selection mechanism across graduate programs. The positive association with an Economics degree likely reflects the alignment between the content of the admission test and the typical undergraduate curriculum in Economics, which emphasizes economic theory, mathematics, and statistics.

In contrast, the selection equation (y_2, presence in RAIS or Lattes) reveals a markedly different pattern – see columns (C), (D), (I), and (J). Regardless of the estimation method employed, the only significant predictor of being observed in RAIS is having started the Master's program before the age of 30. Age at enrollment thus serves as a proxy for career timing and labor market attachment, supporting a lifecycle interpretation: younger individuals are more likely to be active in the formal labor market and, therefore, more likely to appear in administrative employment records.

In the case of Lattes, by contrast, the most salient predictor of observability is current employment in the public sector. This is consistent with the logic that academic environments not only encourage but often require faculty and researchers to maintain updated online CVs and publication records. The incentive structure within these institutions reinforces participation in platforms such as Lattes.

Table 2. Estimates from equations (2)–(5) obtained via 2-Step and FIML methods, with the outcome equation specified in terms of log wages and log publication points. Standard errors are reported in parentheses below the estimates and are clustered by graduate program rank

	When y_3 denotes log wages						When y_3 is log publication points					
	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)	(L)
Variable	y_1 2S★	y_1 FIML	y_2 2S	y_2 FIML	y_3 2S	y_3 FIML	y_1 2S	y_1 FIML	y_2 2S	y_2 FIML	y_3 2S	y_3 FIML
Graduated from a top program (yes = 1)	–	–	−0.293	−0.253	−15.832	−0.033	–	–	−0.511	−0.338	8.944	−1.062
			(0.311)	(0.433)	(10.770)	(0.666)			(0.328)	(0.433)	(15.151)	(0.668)
Admission test score (continuous)	2.271***	2.282***	0.063	−0.052	1.035**	0.327***	2.273***	2.176***	0.022	−0.052	−0.902	0.627***
	(0.167)	(0.284)	(0.125)	(0.136)	(0.380)	(0.080)	(0.169)	(0.288)	(0.135)	(0.139)	(0.566)	(0.221)
Same state as Master's program (yes = 1)	0.040	0.022	–	–	–	–	0.046	0.128	–	–	–	–
	(0.172)	(0.899)					(0.171)	(0.129)				
Single during Master's (yes = 1)	0.637	0.661	–	–	–	–	0.648	0.640	–	–	–	–
	(0.427)	(0.431)					(0.420)	(0.421)				
Started Master's before age 30 (yes = 1)	−0.087	−0.115	0.573***	0.563***	3.399	−0.422	−0.069	−0.115	0.359	0.310	−0.536	0.458
	(0.430)	(0.398)	(0.163)	(0.152)	(2.390)	(1.006)	(0.429)	(0.398)	(0.374)	(0.209)	(0.452)	(0.503)
Male (yes = 1)	−0.216	−0.250	0.039	0.043	0.421	0.316	−0.235	−0.250	0.036	0.043	0.341	0.316
	(0.185)	(0.217)	(0.094)	(0.130)	(0.198)	(0.222)	(0.184)	(0.217)	(0.098)	(0.130)	(0.431)	(0.221)
White (yes = 1)	−0.055	−0.062	−0.090	−0.131	−0.579	−0.031	−0.020	−0.025	−0.074	−0.131	0.003	−0.031
	(0.196)	(0.102)	(0.096)	(0.125)	(0.388)	(0.228)	(0.195)	(0.101)	(0.100)	(0.124)	(0.332)	(0.226)
Undergraduate degree in Economics	0.773	0.774*	0.366*	0.079	2.220	0.193	0.771*	0.774***	0.131	0.079	−0.543	0.193
(yes = 1)	(0.490)	(0.405)	(0.196)	(0.262)	(1.489)	(0.501)	(0.485)	(0.205)	(0.205)	(0.262)	(1.156)	(0.501)
Undergraduate degree in STEM (yes = 1)	0.667	0.696	0.279	−0.019	1.852	0.160	0.714	0.696	−0.016	−0.019	−0.240	0.166
	(0.576)	(0.436)	(0.274)	(0.547)	(1.139)	(1.035)	(0.576)	(0.434)	(0.290)	(0.547)	(0.541)	(1.032)
Public undergraduate (yes = 1)	−0.301	−0.420	0.168	0.437*	1.002	−0.191	−0.315	−0.422	0.377***	0.437***	−0.916	−0.195
	(0.202)	(0.349)	(0.106)	(0.225)	(0.665)	(0.219)	(0.201)	(0.348)	(0.117)	(0.125)	(2.264)	(0.217)
Completed a doctorate by 2018 (yes = 1)	–	–	–	–	0.127	0.217	–	–	–	–	1.266***	1.314***
					(0.105)	(0.062)					(0.089)	(0.090)
Worked in the public sector (yes = 1)	–	–	–	–	0.289**	0.217***	–	–	–	–	0.420**	0.324***
					(0.076)	(0.062)					(0.140)	(0.072)
Primary employment as professor (yes = 1)	–	–	–	–	−0.093	−0.012	–	–	–	–	0.409***	0.418***
					(0.157)	(0.102)					(0.097)	(0.100)
Primary employment as economist (yes = 1)	–	–	–	–	0.163	−0.028	–	–	–	–	−0.127	−0.028
					(0.139)	(0.267)					(0.256)	(0.263)
Primary employment as analyst (yes = 1)	–	–	–	–	0.373**	0.288**	–	–	–	–	0.049	−0.010
					(0.109)	(0.107)					(0.083)	(0.110)

Note(s): ★ denote 2-Step. *Error structure (FIML):* Error variance is 0.701 for wages and 2.239 for points; error correlations are $\rho_{12} = 0.180$ [Wages], 0.141 [Points], $\rho_{13} = 0.013$ [Wages], 0.445 [Points], and $\rho_{23} = -0.812$ [Wages], -0.808 [Points]. *Model fit:* R^2 from 2-Step is 0.227 [Wages] and 0.401 [Points]. Pseudo-R^2 is not reported under FIML, as a constant-only baseline is not identified. *Omissions:* Constant terms and first-stage estimates of λ_1 and λ_2 are omitted. *Significance:* $*p < 0.10$, $**p < 0.05$, $***p < 0.01$

Source(s): The authors

Turning to the outcome equation (y_3), we find divergence in relevant covariates depending on the nature of the outcome – see columns (E), (F), (K), and (L). In the case of log wages, the strongest predictors are the standardized admission test score – serving as a proxy for intrinsic ability – along with indicators for public sector employment and analyst positions.

These associations reflect well-documented features of the Brazilian labor market. For example, public sector careers at institutions such as the Central Bank, the National Treasury Secretariat, the Federal Revenue Service, state finance departments, and audit courts offer highly competitive salaries. Similarly, the “analyst” category frequently includes lucrative positions in private financial institutions, which tend to reward technical skills and quantitative training.

In contrast, log publication points are significantly associated with holding a doctoral degree, public sector employment, working primarily as a professor, and, occasionally, with the proxy for academic ability. These factors identify individuals embedded in academic or research-oriented careers, particularly those situated in public universities. In such environments, research productivity is not only incentivized but institutionalized through formal requirements for grant eligibility, promotion, and tenure evaluations. The admission test score also emerges as statistically significant in many specifications, underscoring the long-run importance of cognitive ability and preparation in determining research output.

Overall, the results suggest distinct sorting patterns across sectors: high wages appear to be more closely associated with technical and administrative roles in the public and financial sectors, while research output may be predominantly driven by engagement with academia and related institutions. These findings are consistent across estimation strategies, highlighting the potential role of both individual and institutional factors in shaping heterogeneous career outcomes.

Interestingly, the prestige of the graduate program does not emerge as statistically significant in any of the selection or outcome equations. This null result is robust across estimation methods (2-Step and FIML) and for both dependent variables (log wages and log publication points). While this finding runs counter to the commonly held view that prestigious programs provide long-term career benefits, it should be interpreted with caution, as other unobserved mechanisms may still play a role.

One possible interpretation is that, once unobserved individual ability (presumably closely related to the admission test score) and other observable characteristics are taken into account, program prestige provides limited additional explanatory power. This is consistent with the hypothesis that institutional prestige primarily acts as a signaling mechanism early in a graduate’s career. Over time, as employers and academic institutions gain access to more accurate measures of performance – including publications, job history, and network reputation – the influence of the initial signal diminishes. In this sense, prestige may facilitate early access to desirable positions but appears less decisive for sustained success, which is ultimately more dependent on personal attributes and demonstrated outcomes.

An alternative, non-exclusive explanation for the absence of statistically significant effects of program prestige relates to data limitations. Although the dataset contains rich, standardized indicators of academic and labor market performance, it may overlook informal or relational channels through which prestige operates. Elite networks, subfield-specific reputations, and unequal access to early-career mentorship could shape career trajectories in ways not captured by administrative records.

The estimated error correlations from the FIML specification offer additional insights into the underlying data structure. In the wage models, the correlation between the unobserved determinants of program prestige and selection into the RAIS dataset is modest and positive ($\rho_{12} = 0.180$, see notes on Table 2), while the correlation between program prestige and the wage outcome is effectively zero ($\rho_{13} = 0.013$). Empirically, this indicates that selection into top programs, driven by unobserved characteristics, does not appear to confound the relationship between observed credentials and wages.

In contrast, the strong negative correlation between selection and wages ($\rho_{23} = -0.812$) underscores the relevance of addressing selection bias. The negative sign indicates that unobserved traits that are associated with a lower likelihood of being observed in RAIS correspond to higher predicted wages.

One plausible explanation for this pattern is that individuals who are not captured in the administrative records may be working under alternative contractual arrangements, such as self-employment or service provision through legal entities (PJ contracts). These forms of employment are common among high-skilled professionals in Brazil, especially in finance and consulting sectors, where they can offer greater flexibility and tax advantages. Importantly, such arrangements are not recorded in RAIS, which primarily captures formal wage employment under standard labor contracts. As a result, high-ability individuals with superior labor market outcomes may be systematically excluded from the observed sample, generating a form of negative selection on unobservables that biases naive estimates. The strong negative correlation between selection and wages observed in the model is consistent with this interpretation.

A similar pattern emerges in the publication models. The correlation between unobserved factors affecting program prestige and selection into the Lattes dataset remains low ($\rho_{12} = 0.141$), but the correlation between prestige and publication points is moderately positive ($\rho_{13} = 0.445$), potentially reflecting some alignment between institutional reputation and research output that is not fully captured by observable variables. Once again, the strong negative correlation between selection and the outcome ($\rho_{23} = -0.808$) highlights the necessity of accounting for selection in the estimation of marginal effects – specifically, the semi-elasticity of academic productivity.

Regarding model fit, the R^2 values from the 2-Step approach indicate moderate explanatory power: 22.7% of the variation in log wages and 40.1% in publication points are explained by the models. Under FIML, Pseudo-R^2 is not reported due to the absence of a well-defined constant-only baseline.

In what follows, Table 3 presents the estimates from models in which the outcome equation is specified with a binary dependent variable, defined in four alternative ways to capture distinct dimensions of academic engagement. While the main analysis focuses on these binary outcomes, we also explore a complementary approach based on an ordered categorical variable that differentiates levels of research productivity. Specifically, we consider an ordinal measure coded as: (1) no publications; (2) best publication in C–B3 journals; (3) in B1–B2; and (4) in A1–A2. This and other modeling variants can be replicated using the accompanying code.

We report estimation results for employment as a professor (=1; columns (A)–(C), corresponding to equations y_1–y_3); publication of at least one article (=1; columns (D)–(F)); publication in an A1 or A2 journal (=1; columns (G)–(I)); and completion of a doctorate (=1; columns (J)–(L)). Each block of three columns presents estimates from the recursive system, with results for y_1, y_2, and y_3 shown side by side.

As for the selection equation for program prestige (y_1), the two covariates most strongly and consistently associated with entry into top-ranked programs are the standardized admission test score and holding an undergraduate degree in Economics. These results are consistent with those presented previously and reinforce the role of prior academic performance and field-specific background in shaping program allocation.

Additionally, being single during the Master's program emerges as a positive and statistically significant predictor of attending a top program. This may reflect greater geographic and temporal flexibility among unmarried students, which could facilitate relocation or full-time commitment to selective academic tracks.

Regarding the sample selection equation (y_2), we find that being relatively younger at the time of enrollment significantly increases the likelihood of being observed in Lattes. This effect likely reflects a lifecycle pattern: younger individuals are more likely to remain engaged in research or academic employment, as they are still in the process of establishing

Table 3. Estimates from equations (2)–(5) obtained via the FIML method, with the outcome equation (y_3) specified in terms of: *primary employment as professor*, *published at least one journal article*, *published at least one article in A1 or A2 journals*, and *completed a doctorate by 2018* (yes = 1 in all cases). Standard errors are reported in parentheses below the estimates and are clustered by graduate program rank

	Professor			Published			A1-A2			Doctorate		
	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)	(L)
Variable	y_1	y_2	y_3	y_1	y_2	y_3	y_1	y_2	y_3	y_1	y_2	y_3
Graduated from a top	–	−0.215	0.409	–	−0.241***	−0.544	–	−0.249**	−0.873	–	−0.237	0.077
program (yes = 1)		(0.470)	(0.438)		(0.082)	(0.730)		(0.107)	(0.772)		(0.222)	(0.412)
Admission test score	2.267***	0.034	−0.408**	2.270***	0.133***	0.007	2.263***	0.137***	0.401	2.287***	0.094	0.095
(continuous)	(0.282)	(0.184)	(0.185)	(0.278)	(0.029)	(0.267)	(0.286)	(0.031)	(0.313)	(0.279)	(0.095)	(0.118)
Same state as Master's	0.043	–	–	0.044	–	–	0.002	–	–	0.046	–	–
program (yes = 1)	(0.071)			(0.077)			(0.085)			(0.098)		
Single during Master's	0.673**	–	–	0.660**	–	–	0.617*	–	–	0.637**	–	–
(yes = 1)	(0.342)			(0.310)			(0.366)			(0.319)		
Started Master's before	−0.128	0.573***	−0.577**	−0.074	0.545***	−0.193	−0.020	0.545***	0.237	−0.068	0.622***	−0.702
age 30 (yes = 1)	(0.403)	(0.135)	(0.254)	(0.393)	(0.192)	(0.130)	(0.418)	(0.199	(0.234)	(0.373)	(0.170)	(0.449)
Male (yes = 1)	−0.243	0.042	0.039	−0.233	−0.038	0.004	−0.241	−0.038	0.033	−0.147	−0.005	−0.168
	(0.219)	(0.105)	(0.160)	(0.213)	(0.116)	(0.106)	(0.213)	(0.115)	(0.161)	(0.228)	(0.128)	(0.112)
White (yes = 1)	−0.043	−0.088	−0.146	−0.007	−0.066	−0.036	−0.108	−0.066	−0.104	−0.065	−0.089	−0.165
	(0.134)	(0.159)	(0.198)	(0.072)	(0.141)	(0.162)	(0.167)	(0.142)	(0.090)	(0.140)	(0.162)	(0.136)
Undergraduate degree	0.653***	0.364	0.602	0.751***	0.269	0.003	0.816**	0.270	−0.192	0.304	0.358	−0.137
in Economics (yes = 1)	(0.247)	(0.240)	(0.633)	(0.284)	(0.188)	(0.201)	(0.406)	(0.189)	(0.422)	(0.224)	(0.224)	(0.201)
Undergraduate degree	0.536	0.277	0.992	0.684*	0.185	−0.254	0.730	0.185	−0.140	0.146	0.283	−0.307
in STEM (yes = 1)	(0.411)	(0.363)	(0.641)	(0.417)	(0.428)	(0.405)	(0.495)	(0.429)	(0.356)	(0.254)	(0.437)	(0.337)
Public undergraduate	−0.320	0.174***	0.085	−0.302	0.179***	0.403*	−0.284	0.178***	0.359*	−0.468	0.197***	0.475
(yes = 1)	(0.328)	(0.023)	(0.153)	(0.327)	(0.028)	(0.213)	(0.319)	(0.027)	(0.198)	(0.352)	(0.017)	(0.405)
Regularly updates	–	–	1.379***	–	–	1.158***	–	–	0.834***	–	–	1.566***
academic CV (yes = 1)			(0.173)			(0.100)			(0.151)			(0.291)

Note(s): Correlations among errors: ρ_{12} = 0.112 [Professor], 0.067 [Published], 0.072 [A1-A2], 0.094 [Doctorate]; ρ_{13} = 0.504 [Professor], 0.324 [Published], 0.419 [A1-A2], 0.034 [Doctorate]; ρ_{23} = −0.102 [Professor], −0.100 [Published], −0.251 [A1-A2], −0.482 [Doctorate]. *Omissions:* Constant terms are omitted. *Significance:* $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Source(s): The authors

their careers, and are thus more inclined to maintain updated academic records. In addition, holding an undergraduate degree from a public university is positively associated with Lattes observability. This may stem from institutional norms and incentive structures within Brazil's public higher education system, which tend to emphasize research activities and promote participation in formal academic documentation systems such as Lattes.

Interestingly, the estimated coefficient on program prestige (y_1) is statistically significant only in the selection equations – specifically, for having published at least one journal article and for having published in A1 or A2 journals. Notably, these coefficients are negative, which may reflect self-selection into academia by graduates from less prestigious programs or strategic responses to differing opportunities across sectors.

An alternative explanation for this negative sign is a sorting mechanism whereby academically oriented individuals within top programs self-select into research careers abroad, which are not documented in the Lattes platform. Additionally, this finding could be partly driven by sample limitations that obscure the broader influence of program prestige on long-term academic trajectories. As such, the interpretation offered here should be viewed with caution, as it may lack external validity.

Regarding the outcome equation (y_3), maintaining an updated academic record appears to be the predominant factor associated with all outcomes, given its positive sign and an absolute value roughly three times larger than that of other statistically significant parameters. This pattern suggests that individuals who remain engaged with the academic system – as indicated by the upkeep of their Lattes curriculum – are also more likely to accumulate academic achievements over time. In this sense, record maintenance may serve both as a proxy for academic integration and as an indicator of continued investment in research-oriented career paths.

When y_3 represents the indicator for having become a professor, both the admission test score and the age-at-enrollment indicator have negative and statistically significant effects. The first finding may reinforce the previously discussed sorting mechanism, whereby individuals with higher ability tend to pursue non-academic career paths.

The negative coefficient on age suggests that older graduates are more likely to become professors. This pattern may reflect institutional dynamics specific to the early 2010s, when many new public universities were established in the interior of the country and predominantly hired Master's graduates. In this context, older individuals may have been more willing to relocate and secure a stable academic position – even in less urbanized regions. The negative sign on age thus captures, at least in part, this potential geographic and career mobility pattern.

As in previous analyses, the estimated correlations among the error terms offer insights into the role of unobserved heterogeneity. The correlation between the selection equation and the program participation equation, ρ_{12}, is positive but modest across all model specifications. This suggests that unobserved factors influencing the likelihood of being observed in the data (e.g. being traceable in RAIS or Lattes) are only weakly correlated with those that determine program status (y_1). In other words, conditional on observed characteristics, the latent factors driving sample inclusion appear to be distinct from those associated with access to more prestigious programs.

The remaining error correlations reveal more substantial patterns. The estimates for ρ_{13} are consistently positive and substantively large – particularly for the outcomes professor (0.504), published (0.324), and A1-A2 publications (0.419). These results imply that the same unobserved factors that facilitate access to elite programs – such as academic orientation or self-selection based on prior preparation – also contribute to enhanced academic performance.

Conversely, the correlation ρ_{23} is negative across all outcomes. This suggests that unobserved factors promoting inclusion in the observed sample may actually be inversely related to academic achievement, possibly reflecting that those with stronger academic profiles are more likely to pursue paths (e.g. international careers) that reduce their visibility in the linked administrative datasets.

Although the regression outputs reported in Tables 2 and 3 are highly informative, as previously discussed in the methodological section – equations (8), (9), (11), and (12) –, the marginal effects must be computed using a more refined procedure to adequately address potential biases stemming from both endogeneity and sample selection. In particular, it is necessary to account for changes in the truncation of the error term distributions, fundamentally by controlling for residual correlations across equations. This adjustment is especially important when interpreting the estimated effects of program type (y_1), whose assignment is likely correlated with unobserved characteristics that also influence the outcomes of interest – including, for instance, latent factors related to signaling mechanisms in the academic and labor markets.

We focus on marginal effects derived from the FIML estimation in order to ensure comparability between linear and non-linear outcome models. Moreover, these effects are evaluated at the mean covariate profile of individuals with entrance exam scores between 1 and 2 standard deviations above the mean. This focus is motivated by the observation that students scoring below one standard deviation are rarely admitted into top programs, whereas those above two standard deviations are almost always admitted.

In contrast, individuals within the [1, 2] interval represent a subset for whom both academic pathways – enrollment in a prestigious program ($y_1 = 1$) or not ($y_1 = 0$) – are realistically attainable. Therefore, it is only within this range that comparisons of potential outcomes under alternative program statuses are meaningful, and the marginal effects of program prestige can be appropriately interpreted. Moreover, all computed effects use a baseline score of 1 in order to simulate a one-standard-deviation increase from 1 to 2 within the interval of interest.

Confidence intervals for the estimated marginal effects are constructed via nonparametric bootstrap procedures. Bootstrapping is particularly well-suited in this context, as it provides reliable inference while avoiding the potential misspecification associated with applying the delta method to multi-equation, non-linear models whose probability distributions – such as the trivariate normal used here – are complex and not easily handled through standard analytical methods, inevitably requiring the use of advanced numerical techniques.

The results discussed in the following paragraphs are reported in detail in the supplementary material. While the latter presents point estimates and confidence intervals, the focus here is on summarizing the main patterns of association through an interpretive lens, with values rounded for clarity.

Starting with labor market outcomes, we observe that a one-standard-deviation increase in exam score is associated with a sizable rise in predicted log wages, regardless of program type. Interpreted as semi-elasticities, these effects imply that a shift from a score of 1 to 2 increases expected wages by approximately 35% for individuals who enrolled in prestigious programs ($y_1 = 1$), and by about 20% for those in lower-prestige programs ($y_1 = 0$), holding other characteristics constant. In contrast, the direct effect of program prestige, conditional on exam score, corresponds to a wage premium that is statistically indistinguishable from zero. These results are consistent with those reported in Table 2 and suggest that program rank, by itself, does not substantially affect wages once individual ability is taken into account.

In contrast, results for publication performance, measured in log Qualis points, reveal a more complex pattern. A one-standard-deviation increase in exam score is associated with a decline of approximately 50% in publication points among students in prestigious programs, whereas the same increase leads to a rise of about 10% for those in lower-prestige programs, although this latter effect is only marginally significant. The estimated effect of program status, conditional on exam score, is not statistically significant.

It is important to note that this inverse effect observed in log Qualis points does not necessarily contradict the results presented in Table 2, as the computation of marginal effects is more nuanced than the direct inspection of estimated outcome coefficients. In particular, the calculation accounts for the estimated correlations between unobserved factors captured by the error terms. Moreover, a plausible explanation for the reversed signs lies in a selection mechanism whereby higher-scoring students in top programs may be more likely to pursue non-academic career paths, which offer fewer incentives to publish but higher financial rewards.

The academic career outcomes further support this interpretation. One-standard-deviation increase in exam score is associated with an increase of approximately 30% in the probability of holding a faculty position among students from lower-prestigious programs, compared to a smaller increase of 10% for those from prestigious programs. Additionally, the estimated effect of program status, holding exam score fixed, implies a 40% lower probability of becoming a professor for those who attended top-ranked programs. This pattern reinforces the idea that prestigious programs may redirect talented individuals toward better-paid careers outside academia – at least among those who do not pursue academic careers abroad.

Regarding research productivity, measured by the likelihood of having published at least one journal article, a one-standard-deviation increase in exam score is associated with a decrease in this probability across both groups. The estimated decline is approximately 10% for students in prestigious programs and around 15% for those in lower-prestige programs.

Although this may appear counterintuitive, the negative effects may reflect threshold dynamics in publication behavior, whereby the decision to publish increases sharply only after certain academic or professional milestones are reached – such as entering a doctoral program, obtaining a research fellowship, or becoming formally affiliated with a higher education institution. In such cases, publishing is no longer optional or incidental, but rather becomes expected, encouraged, or even mandatory. By contrast, students who do not reach these thresholds – even if academically capable – tend to have weaker incentives to publish. This could help explain the observed negative effects, especially if those who meet the thresholds represent a numerical minority compared to those who do not.

For high-impact research, proxied by having published in A1 or A2 journals, no meaningful effects are observed. Regardless of program prestige, the marginal effect of exam score is negligible in both groups, suggesting that academic ability alone does not meaningfully predict top-tier publication. The estimated effect of program status is also not statistically significant, indicating that attending a prestigious program, by itself, does not substantially increase the likelihood of producing high-impact research. This pattern may reflect the fact that publishing in top journals depends less on entry-level academic credentials and more on factors that emerge later in the academic trajectory – such as access to research funding, advisor networks, doctoral training, or institutional publishing culture.

Finally, for the probability of completing a doctorate, the score effect is positive and remarkably similar across groups: a one-standard-deviation increase in exam score raises the likelihood of earning a doctoral degree by approximately 30%, regardless of whether $y_1 = 1$ or $y_1 = 0$. This suggests that academic ability is a strong and consistent predictor of doctoral attainment. The effect of enrolling in a prestigious program, controlling for exam score, is statistically insignificant, suggesting that doctoral completion is determined more by individual characteristics than by program rank.

The estimated parameters and the computed semi-elasticities highlight a nuanced interplay between individual ability, institutional prestige, and career outcomes. Exam score consistently predicts higher wages and an increased likelihood of completing a doctorate, confirming its role as a proxy for individual ability. In contrast, its effects on research productivity are more heterogeneous and less straightforward to interpret, possibly due to selection effects or threshold dynamics whereby incentives to publish become salient only after specific academic milestones are reached.

Program prestige, by comparison, does not exhibit strong direct effects on wages or academic productivity once individual ability is accounted for. Instead, it appears to function primarily as a signaling mechanism that shapes early career opportunities and initial placement. These signaling effects may be particularly relevant at the entry point of professional trajectories, where institutional affiliation can convey valuable information about unobserved traits to employers and academic gatekeepers. Over time, however, individual ability – rather than institutional prestige – seems to play a more central role in shaping long-term outcomes.

6. Conclusion

This study examined how academic credentials – proxied by indicators of individual ability and institutional prestige – relate to the career prospects of Master's graduates in Economics in Brazil. Using a novel set of linked microdata and an extended Heckman model that simultaneously accounts for endogeneity and the effects of sample selection, we addressed methodological challenges that can hinder the interpretation of estimated relationships in observational settings.

When focusing on future wages, our results show that performance on the entrance exam – an indicator of innate ability – is positively associated with higher earnings over time. Institutional prestige appears to operate mainly through early-career signaling, facilitating access to initial labor market opportunities, consistent with most studies in the literature. Nevertheless, in our study it has no strong independent effect on longer-term outcomes once ability and sample selection are considered.

Turning to academic outcomes, the relationship is more complex. Doctoral attainment is positively associated with entrance exam scores, as higher-scoring students are more likely to meet admission requirements and possess the research aptitude for doctoral programs. Yet, among students from top-ranked programs, higher scores are often negatively correlated with the relative volume of reported research output, based on publication records in public CVs. Even so, when these students publish, they tend to do so in higher-quality journals.

This pattern can be plausibly explained by a sorting mechanism. Higher-scoring individuals may pursue better-paying jobs in the private or public sector, where incentives to publish are limited. Graduates from lower-ranked programs appear more likely to remain in academic positions – particularly in the domestic public sector – where publishing is closely tied to career advancement despite lower salaries.

This interpretation is consistent with the context of the early 2010s, when many observed students graduated and numerous faculty openings arose in newly created federal university campuses in Brazil's interior. With few holders of doctoral degrees willing to relocate to smaller cities with limited infrastructure, many positions required only a Master's degree, which may partly explain the relatively high academic placement of graduates from lower-ranked programs in the cohorts analyzed (de Faveri, Petterini, & Barbosa, 2018; Barbosa, Petterini, & Ferreira, 2019; Lazaretti & França, 2020).

In sum, the findings from this study highlight a multifaceted interaction between academic ability, institutional context, and career outcomes, suggesting that graduate trajectories are shaped primarily by individual characteristics, with institutional factors playing a more limited, though still relevant, role. While institutional characteristics may exert greater influence on initial career placement, individual ability appears to have a more enduring impact on long-term outcomes. Nonetheless, several limitations must be acknowledged when interpreting these results.

First, our measure of institutional prestige relies on standardized program rankings commonly used in national evaluations. However, this may overlook important dimensions of prestige, such as discipline-specific reputation and informal academic networks, which may also have an impact on career trajectories. Future research could examine complementary indicators of prestige by identifying where alumni have excelled professionally and developing measures based on these external career successes.

Second, although our dataset integrates rich administrative records on employment and academic outcomes within Brazil, it does not capture research production abroad, potentially underestimating the output of graduates who pursue international academic careers. Moreover, even within Brazil, it fails to track individuals engaged in formal work through legal entities (e.g. independent contractors or business owners, known as PJs), which may lead to the underrepresentation of certain employment arrangements, particularly in the private sector. Addressing these gaps would require access to additional data sources – such as direct surveys of graduate programs or microdata from state-level commercial registries and the federal tax authority – whose use is currently restricted under Brazil's General Data Protection Law (LGPD).

Third, this analysis focuses on only two cohorts of graduates, which limits our capacity to examine long-term trends or generational shifts in labor market and academic contexts. These extensions could not be carried out due to restrictions on personal data access and linkage procedures imposed by the LGPD.

Fourth, our empirical strategy is essentially based on parametric assumptions, in particular the multivariate normality of latent error terms. While we employ full-information maximum likelihood estimation and bootstrap inference to mitigate potential biases, concerns about model misspecification remain. Future research could address this limitation by applying more flexible modeling approaches – such as copula-based econometrics, which retains the parametric structure of the present model and allows the normality assumption to be relaxed. Another possibility is to use semiparametric methods, which allow for greater heterogeneity and impose fewer distributional assumptions.

Overall, we believe that future research should track career trajectories over longer periods, distinguish between domestic and international placements, account for informal or non-salaried work arrangements, and incorporate qualitative insights from graduate surveys. It could also expand the scope of outcomes to include non-academic contributions – such as public policy, entrepreneurship, or applied consulting – to deepen understanding of how academic training translates into real-world impact.

Ethics statements

This manuscript adheres to the guidelines for Ethics in Publishing.

Acknowledgments

The authors would like to thank the three anonymous reviewers and the two editors who oversaw the revisions of this work (Prof. Mauro Rodrigues and Prof. Matheus Albergaria) for their thoughtful and constructive feedback, which significantly contributed to improving the quality and clarity of the manuscript.

Supplementary material

The supplementary material for this article can be found online.

References

Athey, S., Katz, L. F., Krueger, A. B., Levitt, S., & Poterba, J. (2007). What does performance in graduate school predict? Graduate economics education and student outcomes. *American Economic Review*, *97*(2), 512–518. doi: 10.1257/aer.97.2.512.

Barbosa, M., Petterini, F., & Ferreira, R. (2019). Expansion of Brazilian federal universities: Is it possible to raise economic impacts?. *Journal of Contemporary Administration*, *24*(1), 3–24. doi: 10.1590/1982-7849rac2020190230.

Cardoso, A., Guimarães, P., & Zimmermann, K. (2010). Comparing the early research performance of PhD graduates in labor economics in Europe and the USA. *Scientometrics*, *84*(3), 621–637. doi: 10.1007/s11192-009-0136-5.

Chen, J., Liu, Q., & Billger, S. (2013). Where do new Ph. D. Economists go? Recent evidence from initial labor market. *Journal of Labor Research*, *34*(3), 312–338. doi: 10.1007/s12122-013-9162-4.

Dale, S. B., & Krueger, A. B. (2002). Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables. *The Quarterly Journal of Economics*, *117*(4), 1491–1527. doi: 10.1162/003355302320935089.

de Faveri, D., Petterini, F., & Barbosa, M. (2018). Uma avaliação do impacto da política de expansão dos Institutos Federais nas economias dos municípios brasileiros. *Planejamento e Politicas Publicas*. Available from: https://www.ipea.gov.br/ppp/index.php/PPP/article/view/742

Estevan, F., & Santos, K. (2022). Does it matter which top institution you choose? A case study of Brazilian graduate admissions. A Case Study of Brazilian Graduate Admissions. doi: 10.2139/ssrn.4225787.

Fernández, R. G., & Suprinyak, C. E. (2018). Creating academic economics in Brazil: The Ford Foundation and the beginnings of ANPEC. *EconomiA*, *19*(3), 314–329. doi: 10.1016/j.econ.2018.03.004.

García-Suaza, A., Otero, J., & Winkelmann, R. (2020). Predicting early career productivity of PhD economists: Does advisor-match matter?. *Scientometrics*, *122*(1), 429–449. doi: 10.1007/s11192-019-03277-8.

Hadlock, C. J., & Pierce, J. R. (2021). Hiring your friends: Evidence from the market for financial economists. *ILR Review*, *74*(4), 977–1007. doi: 10.1177/0019793919896755.

Hussey, A., Murray, S., & Stock, W. (2022). Gender, coauthorship, and academic outcomes in economics. *Economic Inquiry*, *60*(2), 465–484. doi: 10.1111/ecin.13047.

Krueger, A. B., & Wu, S. (2000). Forecasting job placements of economics graduate students. *The Journal of Economic Education*, *31*(1), 81–94. doi: 10.1080/00220480009596765.

Lazaretti, L., & França, M. T. (2020). School competition and performance indicators: Evidence from the creation of federal education institutions in Brazil. *International Journal of Educational Development*, *77*, 102211. doi: 10.1016/j.ijedudev.2020.102211.

Pereda, P., Montoya Diaz, M. D., Rocha, F., Matsunaga, L., Borges, P. B., Mena-Chalco, J., Narita, R., & Brenck, C. (2023). Are women less persistent? Evidence from submissions to a nationwide meeting of economics. *Applied Economics*, *55*(16), 1757–1768. doi:10.1080/00036846.2022.2099525.

Petterini, F. C. (2020). Brazilian academic economics: A picture from the ANPEC exam microdata. *EconomiA*, *21*(3), 325–339. doi: 10.1016/j.econ.2020.04.004.

Rocha, F., Pereda, P., Matsunaga, L., Montoya Diaz, M. D., Narita, R., & Borges, B. (2021). Gender differences in the academic career of economics in Brazil. *Cuadernos de Economía, 40*(SPE84), 815–852. doi: 10.15446/cuad.econ.v40n84.86778.

Sullivan, R. S., Dubnicki, A., & Dutkowsky, D. H. (2018). Research, teaching, and 'other': What determines job placement of economics Ph. Ds?. *Applied Economics*, *50*(32), 3477–3492. doi: 10.1080/00036846.2018.1430331.

Corresponding author
Francis Petterini can be contacted at: f.petterini@ufsc.br