




RESEARCH PAPER

# Measuring the Social and Economic Impact of Universities' Entrepreneurial Activity: Introducing the BR-AFC Algorithm to Sort Alumni-Founded Companies

Daniela Uziel<sup>1</sup>, Edison Renato Pereira da Silva<sup>1</sup>, Humberto Henriques de Arruda<sup>1</sup>

<sup>1</sup>Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, RJ, Brazil.

How to cite: Uziel, D., Silva, E.R.P. and Arruda, H.H. (2024), "Measuring the Social and Economic Impact of Universities' Entrepreneurial Activity: Introducing the BR-AFC Algorithm to Sort Alumni-Founded Companies", *Brazilian Journal of Operations and Production Management*, Vol. 21, No. 1, e20241808. <https://doi.org/10.14488/BJOPM.1808.2024>

## ABSTRACT

**Goal:** This study introduces an algorithm to sort alumni-founded companies from the public Brazilian Internal Revenue Service (IRS) database.

**Design/Methodology/Approach:** Departing from IRS data and student data from the university, sequential filters are applied to arrive at a final list of alumni-founded companies.

**Results:** The main result of this study is the establishment of the algorithm itself, which emerged after cycles of iterations of analysis and rewriting. To test its reliability, a sample of 1625 alumni was used. The algorithm successfully identified 140 founders of 159 AFC. Founders were heterogeneously distributed throughout the decades analyzed. Companies belonged to different industry sectors and were classified according to their technological intensity, with predominance of middle-low and low intensity.

**Research limitations/implications:** Although the BR-AFC algorithm is applicable to any Brazilian institution, generalization to other countries depends on access to country-specific databases containing data about companies and its partners. Additionally, the final result depends on the reliability of input data and of user decisions about the rigor of its premises.

**Practical implications:** The BR-AFC algorithm can improve measurements of the socioeconomic impact of educational institutions. It points to the formation of entrepreneurs and, as a consequence, institutions can evaluate courses and educational programs and improve curricula. Policymakers and sponsoring institutions can measure return over investment, outcomes of policies to encourage entrepreneurship and rank universities according to novel criteria.

**Originality/Value:** The main contribution to the literature is exploring novel approaches to measuring university-industry relationship. More specifically, it proposes an algorithm to identify the alumni-founded companies of a given university from large country-based databases.

**Keywords:** Alumni-founded companies; Entrepreneurship; University Industry Relationship; Internal Revenue Service.

## 1 INTRODUCTION

Universities in Brazil are established upon a foundation of teaching, research, and outreach endeavors. In global academic discourse, the latter is commonly termed the "third mission" of a university, which is broadly defined as a "contribution to society." This encompasses a diverse array of activities that entail contributions to the social, economic, and cultural development of not only

**Financial support:** None.

**Conflict of interest:** The authors have no conflict of interest to declare.

**Corresponding author:** [daniela.uziel@ufrj.br](mailto:daniela.uziel@ufrj.br)

**Received:** 14 March 2023.

**Accepted:** 06 September 2023.

**Editor:** Julio Vieira Neto.



This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

the state, region, but also the entire nation. In a recent review, Compagnucci and Spigarelli (2020) cite the work of contemporary authors as evidence showcasing the impact of knowledge and technology transfer on both industry and society, as integral components of the university's third mission. Universities around the world are looking to update their curricula to promote greater interdisciplinarity (Klein, 2015; Marins, Ramos, Ferreira, Costa, & Costa, 2019) and stimulate innovative behaviors (Hoidn & Kärkkäinen, 2014; Täks, Tynjälä, Toding, Kukemelk, & Venesaar, 2014; Ten Caten, Silva, de Aguiar, Filho, & Huerta, 2019).

The scope of university entrepreneurial activity is encompassed within this "third mission" (Li, Yang, & Cai, 2021; Vefago, Trierweiler, & de Paula, 2020), encompassing actions such as patenting, licensing, establishing incubators, science parks, and spinoffs. There is clear empirical evidence indicating an uptick in such activities in developed countries (Mowery, Nelson, Sampat, & Ziedonis, 2004; Phan & Siegel, 2006; Rothaermel, Agung, & Jiang, 2007). This phenomenon has also gained traction in developing nations, evident in cases like Brazil (Freitas, Gonçalves, Cheng, & Muniz, 2011), even preceding the promulgation of the Science, Technology, and Innovation Act (Brazil, 2016).

While universities engage in a multitude of activities, they are primarily assessed based on their scientific output and its influence. Metrics such as the quantity of published papers, citation counts, and journal impact factors are accessible on publishing platforms, fostering a self-reinforcing cycle that bolsters research. However, contemporary literature emphasizes the necessity of quantifying the university's "third mission" (Compagnucci & Spigarelli, 2020) and proposes pertinent indicators (Uziel & Allonso, 2022). One vital aspect of the university-industry relationship is evaluating the number of spinoffs, which serves as a significant dimension in this context.

The literature in the field often refers to companies whose origin of intellectual capital is tied to a university or public research institution as "spinoffs" (Callan, 2001). Shane (2004) defines a spinoff as "a new company founded to exploit a piece of intellectual property created in an academic institution." Similarly, the Organization for Economic Cooperation and Development (OECD) characterizes an intellectual property-based spinoff from a publicly funded research organization as "a new firm whose startup incorporates a significant contribution of knowledge recently developed in a public research organization, with this knowledge being protected by intellectual property rights that are either licensed or transferred to the firm". While there is some variation in the definition of the term, there is a degree of convergence that academic spinoffs are established by individuals originating from universities or research institutions, who then seek to commercially exploit the intellectual property created there.

Quantitative studies adopt diverse measures to assess spinoff creation. North American researchers often utilize Association of University Technology Managers (AUTM) surveys to quantify the number of spinoff companies originating from universities or research institutes (O'shea, Allen, Chevalier, & Roche, 2005). AUTM exclusively tallies companies that have licensed intellectual property from a university or research institution. European scholars, on the other hand, create their own databases, identifying spinoffs through publicly accessible sources (Conceição, Faria, & Fontes, 2017) or official spinoff lists provided by universities and public research centers (Meoli, Pierucci, & Vismara, 2018).

Employing intellectual property-based spinoffs as a measure for a university's new venture creation and its subsequent economic impact can yield results that are ambiguous and incomplete. The ambiguity arises due to the disputable nature of quantifying the extent to which intellectual capital contributes to the inception of a new business. Incompleteness stems from the fact that not all knowledge generated within universities and research institutes conforms to the criteria for formal protection. Consequently, knowledge originating from a university might give rise to new companies that do not meet the classification of spinoffs, leading to the exclusion of numerous economic activities from the tally.

In pursuit of alternatives, institutions like Stanford and MIT have focused on companies founded by alumni to delineate their economic impact, based on factors such as domain of activity, employment rates, revenue, income, and market capital (Lebret, 2017; Roberts, Murray, & Kim, 2019). The University of Campinas (Unicamp) has introduced the term "daughter companies" to encompass those founded by individuals with affiliations to Unicamp, including students, former students, professors, former professors, employees, former employees, and those previously incubated or graduated from the institution's incubator (Unicamp, 2020). This study delves further into the definition of daughter companies and presents a logical algorithm for calculating companies established by alumni of a given institution, exploring their economic impact.

The approach taken in this study quantifies entrepreneurial activity at the Federal University of Rio de Janeiro (UFRJ) and employs its Industrial Engineering course alumni as an initial step in a demographic study of UFRJ's alumni-founded companies. The methodology's resultant data already encompasses companies' economic activities and can be combined with data from other sources (e.g., the technology transfer office) to analyze its societal and economic developmental

impact more comprehensively.

## 2 METHOD

With the aim of formulating an algorithm for categorizing companies founded by alumni, this project utilized two databases. The first database is an anonymized open dataset sourced from the Brazilian Internal Revenue Service (equivalent to the American Internal Revenue Service), updated monthly on a government online open data platform (Receita Federal do Brasil, 2023). The second database is derived from UFRJ's internal student registration system. Since student data is personally identifiable, special permission was secured for this study. Data handling adheres to the principles of the Brazilian General Data Protection Regulation, affording the institution the right to conduct research for public policy development. All outcomes are presented in aggregated form, ensuring individual identification remains unfeasible.

### 2.1 Definition of the term 'alumni-founded company'

In this project, an "alumni-founded company" is defined as any entity listed in the Brazilian Internal Revenue Service database that was established by at least one UFRJ graduate or postgraduate alumnus, regardless of its economic activity. It's important to underscore that the collection of companies identified through present methodology encompasses those utilizing intellectual property from the university (and would, accordingly, meet the classification of spinoffs, as per Shane, 2004), yet is not constrained solely by this criterion. Our pool of companies also includes a subset that might be categorized as "daughter companies" according to Unicamp's standards. However, it's noteworthy that Unicamp designates companies as daughters when they respond to their call and provide data on their platform. In contrast, our chosen terminology differs, as the companies authors identified did not actively enroll as university "daughters," and our focus deliberately excludes companies initiated by staff.

In the Portuguese context, the term "company" can yield a dual interpretation: it could pertain to a legal entity name (Razão Social) or a Corporate Taxpayer Identification Number (CTIN), derived from the Portuguese expression Cadastro Nacional de Pessoa Jurídica - CNPJ. Given that companies can generate additional CTIN entries (as branches) for expansion or other purposes, it's conceivable that a single legal entity name could be linked to multiple CTINs. Thus, this work rests upon the following premises: (1) Each company corresponds to a distinct legal entity name; and (2) A CTIN is singularly associated with a unique legal entity name. Henceforth, when referencing a company, this is a specific reference to a legal entity name.

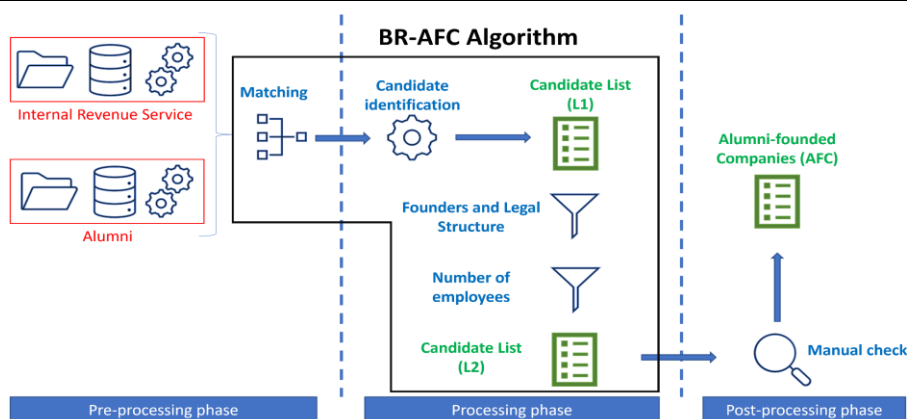
### 2.2 The nature of the method

It is essential to highlight the empiric nature of the method established to identify alumni-founded companies (AFC). The algorithm emerged after cycles of iterations, in which the following steps were repeated:

- Establish criteria for selecting AFC;
- Implement the criteria on the database;
- Analyze the results;
- Identify results that are incompatible (examples below);
- Restart the process.

In the first attempt to sort AFC, our results included multinational corporations and state-owned companies, which was incompatible with reality. This was the first clue to design an algorithm that sorts companies based on alumni as founders and to deal with companies that have a legal entity as a partner as a special subset of data. To be classified as AFC, two main criteria were followed: (1) When the company does not have a legal entity as a partner, an alumni must have joined the company as a partner up to day 30 from the company's starting date; and (2) If the company does have a legal entity as a partner, it must have joined the company after day 30, counting from the company's starting date. Limitations and possible variations of these criteria will be addressed in further topics of the Method and in the Discussion section "Limitations on the algorithm settings".

Our method can be divided in three main steps, as presented in Figure 1: pre-processing, processing and post-processing. The conceived algorithm (box in Figure 1) permeates the Pre-processing and the Processing phases.



**Figure 1** - Overview of the proposed method of sorting alumni-founded companies.

Briefly, the pre-processing phase aims at preparing the data, since the original data from the Internal Revenue Service (IRS) and from UFRJ presents different types of inconsistency (use of special characters, abbreviations, missing zeros etc.). The processing phase focuses on generating the AFC list (L2) closest to the final result, based on the application of two successive filters to the initial list of candidate companies (L1). Filters are based on the algorithm detailed in the next section. Finally, in the post-processing phase inconsistencies that may have passed through the processing algorithm are eliminated and data reliability indicators are calculated. The final result is the list of AFC from the studied institution.

### The Algorithm: A Synthesis

The processing phase of the algorithm is composed of the following steps, that will be detailed in the next subsections.

- A. Identify all companies in which alumni are partners (list L1);
- B. Apply first filter to companies in L1: founding partners and legal structure:
  - a. Calculate the founding date of every company in L1 using two different criteria;
  - b. Identify founding partners;
  - c. Exclude companies in which alumni were not founding partners;
  - d. Exclude companies in which legal entities were founding partners;
  - e. Classify companies according to the presence of legal entity as non-founding partners;
- C. Apply second filter to resulting companies: number of employees over time:
  - a. Calculate the number of employees of each company per year, since its founding date;
  - b. Classify companies as "CTIN as a service provider for another company" (also known in Brazil as "pjotização", when a worker is hired as a legal entity) or as regular company.
- D. Arrive at the L2 list of AFC, to be post-processed.

### 2.3 Pre-processing phase

In this phase, a first step of extraction, transformation and loading of IRS and UFRJ data took place, followed by a step of matching of the databases.

The first step carried out was to upload data using a Relational Database Management System, which allows efficient manipulation of large volumes of the data containing all companies in the country.

For the matching, only the Unique National ID (UNID, derived from the Portuguese term *Cadastro de Pessoa Física - CPF*) and Full Name fields were used from the IRS data. It is important to note that the IRS makes UNID data available truncated to 6 digits. That is, the data is displayed in the form XXX.NNN.NNN-XX, where N stands for available digits and X for unavailable digits. In the case of alumni data, the full UNID and full name were available.

The initial difficulty to be overcome was to make the appropriate matching between the same individual in both databases. For this it was necessary to:

1. Cross the truncated UNID data from the IRS with the full UNID data: Only entries in which all available digits were equal were kept as matching candidates;
2. Cross the Full Name data from both databases: When the UNIDs matched, the names were verified. If the first and last name were the same, the entry was accepted.

When matching and understanding the reliability of the results, the pre-processing stage of the data was completed, making them ready for the beginning of the processing phase.

Limitations and possible variations of these criteria will be addressed in the Discussion section "Limitations on the algorithm settings".

## 2.4 Processing phase

Data processing took place in three stages: identification of candidate companies; filter based on founding partners and legal structure; and filter based on the number of employees over time. Initially, all candidate companies were identified, which were those that are linked to alumni in the IRS database, leading to an initial list of candidates (L1). From there, filters were applied and a more restrictive list of companies (L2) emerged. For all the steps in the processing phase, specific attributes of the IRS database were used, as follows.

### Identification of candidates

This stage is designed to comprehensively identify all companies that hold the potential to qualify as "Alumni-Founded Companies" (AFCs) through a meticulous comparison of IRS and university databases. The initial step involved extracting all companies from the IRS dataset that exhibited any form of connection to alumni. This formed the foundational L1 candidate list. This preliminary roster encompassed companies genuinely established by alumni as well as instances of discrepancies. The discrepancies stemmed from companies where an alumnus was attributed any form of social responsibility, such as holding directorial or presidential roles, or even serving as shareholders within corporations. These discrepancies materialize at this phase due to the IRS database's inclusion of all entries—including both natural persons and legal entities—that maintain partnerships with a given company.

### Filter based on founding partners and legal structure

This step aims to select from L1 companies only those founded by alumni.

Precisely, founding partners of a given company are those that join it since it legally comes into existence (founding date or starting date). In practice, however, authors noticed from data sample that there was a dissociation between the company's starting date (a column in the IRS dataset named "activity starting date") and the date assigned as the earliest entry of a partner in the partnership (calculated by comparing all the entries of partners for a given CTIN). No clear cause was identified: it could be attributable to bureaucratic procedures regarding the CTIN creation or other reason. To account for this, all partners (natural persons and legal entities) who joined within 30 days from the company's starting date were considered founders. Aiming at more accurate results, for each CTIN, there was a verification whether a partner could be considered a founder (within the 30-day lag after the starting date) using two criteria (attributes from the IRS database): the "activity starting date" of a company and the earliest "date of entry into society". For instance, if a company's "activity starting date" was January 1st but the first partner joined the society on February 15th ("date of entry into society" field), everyone entering the society from January 1st to January 30th was considered a founder by criterion 1 and 2. From January 31st to March 17th (D+30) it would be considered a founder only on the basis of criterion 2. Each criteria was tracked individually, to measure the degree of certainty that a given person was indeed a founder.

In terms of the specific procedures to filter L1, as mentioned before, more than one CTIN can be associated with the same company. Thus, the first step was to obtain the date the company legally came into existence, based on the earliest founding date of all CTIN linked to a given legal entity name ("razão social"). This can be done step-by-step, as follows:

- i. Scan the database to identify the earliest start date for each CTIN linked to a given legal entity name. This is the first criterion to obtain the founding date of each company.
- ii. Scan the database to identify the earliest date among all the dates assigned as a partner joining the company for all the CTIN numbers associated with that given legal entity name. This is the second criterion to obtain the founding date of each company.
- iii. List all partners that could be considered founders by criteria 1 and 2 and check whether they are alumni: search for alumni becoming partners within 30 days since the founding date of that company. All companies in which the alumni entered after 30 days from the founding date (calculated by criteria 1 or 2) are excluded as AFC candidates.
- iv. List all other founding entries in the previous step to check for companies that might be AFC candidates after the previous step. List all other founders' entries and check if there are legal entities or only natural persons listed as founders. This is a classificatory

step that allows the method to account for differing uncertainties regarding whether a person was indeed a founder of an entirely new company or of a spinoff or branch of another company that holds a different company name from the parent company.

Based on this procedure, companies can be classified into three groups:

1. Companies that have only natural persons as partners;
2. Companies that have a mix of natural persons and legal entities as partners:
  - a) Those that were founded by a mix of natural persons and legal entities; and
  - b) Those that were founded only by natural persons (and in which legal entities entered after the 30-day lag).

Acceptance Criteria: Based on the outcome from step iv, each company is classified in one of the three categories listed above. Companies classified as 1 or 2b show a higher probability of being an AFC, whereas those classified as 2a show a higher risk of false positives.

Limitations and possible variations of these criteria will be addressed in the Discussion section "Limitations on the nature of the IRS database".

### **Filter based on the number of employees over time**

During this stage, companies that continue to meet the eligibility criteria went through an assessment based on their recorded employee count over a span of time. This stage serves the purpose of distinguishing entities engaged in authentic productive operations from those established primarily to function as "CTIN service providers" for other companies (a phenomenon colloquially referred to as "pjetização" in Brazil, wherein an individual is hired as a legal entity).

To verify the number of employees hired by AFC from their starting date, the Annual List of Social Information (ALSI, derived from the term in Portuguese *Relação Anual de Informações Sociais - RAIS*) of the Ministry of Labor was used. For the analysis described in the Results section, RAIS data from 1985 to 2017 were used. Based on companies' CTIN, the database was scanned, and the number of employees was registered yearly. Companies were then classified according to their hiring pattern. AFC were classified according to their number of employees, number of partners, existence of legal entities as partners, number of branches, existence of an assumed business name ("nome fantasia"). Those companies that matched all the following criteria were classified as an "CTIN as a service provider for another company" (its creation is probably attributed to an event of hiring workers as entities): hired no employees along their lives, have up to two partners that are both natural persons, does not have legal entities as partners, does not have branches and did not register any assumed business name.

At the end of this stage, a list of companies (L2) came out. It contained all the data from IRS - including founders' entries dates and company start date, that were the basis of the sorting - and the number of employees along time (from the Annual List of Social Information). The list was further refined in the post-processing phase.

### **2.5 Post-processing phase**

From all the parameters previously calculated, the algorithm generates an AFC list (L2). This list L2 can still not be considered the final and definitive one, and it is recommended that during the post-processing phase this list of AFC is manually checked to verify inconsistencies that may have passed throughout processing, due to the inherent uncertainties of the imprecise matching of UNIDs between the bases, the quality of founders' entries dates and company starting date. A final sorting of the data based on its analysis is considered very important (see Discussion section "Limitations on the nature of the IRS database").

## **3 RESULTS AND DISCUSSION**

The main result of this study was the establishment of an algorithm to extract a list of alumni-founded companies from a large IRS database that includes all companies and partners available in Brazil.

For the proposal of the method and for its validation, a pilot study was carried out with alumni from the Industrial Engineering course of the Federal University of Rio de Janeiro.

First, initial population consisted of 2849 alumni who graduated between 1970 and February 2021. From 1970 to 2000, 33% of alumni had either missing or invalid UNIDs registered in the database, since only by 2001 the university created a unified institutional data system, gathering data from different and independent internal sources. Then, only 1625 individuals from the original list were used in the search.

Among the 1625 alumni analyzed, there were 140 founders of 159 AFC companies. Founders were heterogeneously distributed throughout the decades analyzed: 5 in the 1970's (out of 17 alumni), 10 in the 1980's (out of 270 alumni), 15 in the 1990's (out of 867 alumni), 40 in the 2000's (out of 774 alumni), 70 in the 2010's (out of 690 alumni) and none in 2020/2021 (out of 32 alumni).

Alumni combined themselves as partners to found companies that had the following temporal distribution: 2 in the 1970's, 1 in the 1980's, 21 in the 1990's, 39 in the 2000's, 88 in the 2010's and 8 in 2020/2021. The range of alumni-founders per company varied from 1 to 4 in 97% of the companies analyzed. Ten percent of the companies, however, showed a high number of partners. Among those, one company presented 116 partners.

From the 159 AFC companies identified, 116 did not have any employees since their founding date. Regarding companies' economic activity, these 116 companies were initially considered small businesses with no employees, which may characterize early stage-startups, IT companies, consultancy partnerships or other business arrangements that, although highly active, are not typical employers. Among the 116 companies, 34 of them fulfilled a profile of "CTIN as a service provider for another company": did not employ any individuals since their date of creation, had up to two partners (and none of them were legal entities), had no branches or had no assumed business name.

Another 25 companies were small businesses that had up to 10 employees. Their economic activity based on ISIC classification (section) were distributed as follows: 47 (Retail trade, except of motor vehicles and motorcycles,  $n = 4$ ), 46 (Wholesale trade, except of motor vehicles and motorcycles,  $n = 3$ ), 82 (Office administrative, office support and other business support activities,  $n = 3$ ) and 70 (Activities of head offices; management consultancy activities,  $n = 3$ ). There was a large variability of economic activities among the other 12 companies.

A second group of AFC ( $n = 18$ ) was characterized by an increasing number of employees (always more than 10 employees) throughout the period analyzed. Among those, three companies had "Retail trade, except motor vehicles and motorcycles" (ISIC section 47) as their primary economic activity. Three companies were in "Wholesale trade, except of motor vehicles and motorcycles activities" (ISIC section 46), two other companies in ISIC section 66 (Activities auxiliary to financial service and insurance activities), two in ISIC section 62 (Computer programming, consultancy and related activities), two in ISIC section 68 (Real estate activities) and two in ISIC section 85 (Education). Four remaining companies were sparsely distributed among other economic activities.

All companies were also classified according to their technological intensity, based on the OECD taxonomy (Galindo-Rueda & Verger, 2016). Most of the companies were classified either as middle-low intensity ( $n=45$ ) or low intensity ( $n=78$ ). Only few companies ( $n=18$ ) were middle-high or high intensity companies, as exemplified by those in ISIC section 72 of economic activity (Scientific research and development).

In respect to their spatial distribution, most of the companies were situated in the state of Rio de Janeiro ( $n=134$ ), mainly in the cities of Rio de Janeiro ( $n = 116$ ) and Niteroi ( $n = 6$ ). Sixteen companies were located in the state of Sao Paulo and the remaining were in other Brazilian states.

Although not described in this paper, it is also possible to correlate the academic performance levels of alumni or their quota status with their entrepreneurial aptitude or to other attributes available in the institution's database.

### 3.1 Methodological framework and limitations of the method

It is important to emphasize that a simple association of an alumni UNID and a company CTIN within the Internal Revenue Service database will not strictly bring a list of companies that were founded by alumni. It would rather bring all cases in which UNID and CTIN are associated, including those cases where alumni are allocated as director or president of a company or where they are shareholders, among other situations which are not of interest here. This methodological framework faced diverse possibilities and obtained a list of companies founded by alumni by subtracting cases that are not of interest and selecting companies where alumni can be identified since their creation. Parameters adopted in the algorithm can lead to more restrictive or to wider results, depending on how they are calibrated (see further discussion).

Limitations of the method can be related to the algorithm itself, to the quality of the data provided by the university or downloaded from the Internal Revenue Service platform or to how strict were the criteria adopted throughout the algorithm. Here, there is a discussion about some issues that may affect the quality of the final output.

#### Limitations on the algorithm settings

In the pursuit of identifying founding partners, the algorithm scans the database in two different ways: (1) to identify the earliest starting date for each CTIN linked to a given legal entity name and

(2) to identify the earliest date among all the dates assigned as a partner joining the company. If the time interval between the founding date and the date of entry into the society of the graduate is null, it is possible to be fully confident that given alumnus is a founder. In the Revenue Service database, however, the date assigned as the starting date of the company frequently precedes the date of any partner joining it. Therefore, one must define an acceptable time interval between these two events, in which a partner is considered a founder.

The creation of new companies involves several steps, from building a business model, getting funded and developing the solution until selling it. No evidence was found in the current literature that having a team of co-founders from the very beginning is a factor related to business success. Literature points to characteristics of the partners (passion, openness, ethic etc) (Forbes, 2021) and of the partnership (mutual accountability, true commitment, shared goal etc) (Forbes, 2018) rather than to the time lag of joining the company. Nonetheless, it was necessary to design a solution to exclude from our dataset companies where the participation of alumni does not involve founding it. Authors then admitted that the closest to 30 days the time lag is, the highest is the probability of a new partner being a founder. Other shorter or larger periods can be used by other authors. The larger this parameter, the wider the results from the algorithm will be. Further research can propose a specific rule or logic that could help determine the best rigorosity for each dataset.

It is important to emphasize the empirical and data-based nature of the method proposed in this work. Analyzing the method in a simplified way, filters are applied to a list of companies to classify each one of them at the end of the sequence of selections. The quality of classifying a given company as an AFC is highly related to the rigorosity of the parameters used. There is always a risk of false positives or false negatives, depending on this setup.

If a company was founded by a legal entity, it does not fit in the definition of an AFC, even if one of its founders is an alumnus. This corporate composition can be attributable for purposes other than the company's productive activity, such as avoiding tax or labor issues, or because of a spinoff or joint venture. For instance, venture capital (VC) firms usually create new companies under a business structure called "society of specific purpose" (SPE in the Portuguese acronym) to carry out investments. Executives from the VC become founding partners of the SPE, but that doesn't imply they are entrepreneurs: this is only part of their attribution as executives of the original company.

On the other hand, if a certain company is composed only of natural persons, it is very unlikely that it was created as a spinoff of another company, for example. A company whose partners are only natural persons and meets the selection criteria described above will have a very high probability of actually being an AFC.

As a result, it is possible to notice that there are two extremes in the paragraphs above. On one hand, companies that have a legal entity founder. On the other hand, companies that only have natural persons as partners. But there are companies lying between these two extremes: companies that have a non-founding legal entity as partner – startups that received investments later, for example. If a 30-day window between foundation and entry of a legal entity is adopted, any legal entity that becomes partner after that period will be considered non-founding, and the company may be classified as AFC. As the matter is never clear-cut, the post-processing procedure is necessary to check for false positives and false negatives.

Therefore, the choice was to divide the analyzed companies into:

- i. Companies that do not have a legal entity as a partner: if they meet the other criteria, will be classified as AFC.
- ii. Companies that have a legal entity partner: it is necessary to analyze whether the legal entity partner is a founder or not.

For case (ii) above, the greater the time interval between the entry of the legal entity partner and the founding of a given company, the greater the chance that this legal entity partner will not be a founder.

### **Limitations on the nature of the Internal Revenue Service database**

The Brazilian IRS makes available large amounts of data partitioned in sequential files in the governmental open data platform. The database contains only public data that is adequate for the analysis proposed here but does not provide information regarding employees or revenues (which are private information) for deeper analysis. In Brazil, Unicamp use alumni-directed surveys to gather data from "daughter companies" and track their financial performance ("Unicamp companies," n.d.).

An Application Programming Interface (API) of IRS data is not yet available: files must rather be downloaded as csv and processed as desired. Such data is ideal for cross-sectional studies, such as the one described here. Although it allows a bulge of results, they are limited in time. For example, a merge or an acquisition of a former AFC will not be detected by a cross-sectional analysis such as the one proposed here, if the alumni UNID is no longer related to the present CTIN. Progress,



improvements, and modifications of CTIN along the years would only be possible based on a temporal analysis of the data. It however implies downloading and storing of files by researchers when made available by the IRS, since monthly updates replace former files.

After preprocessing the pilot industrial engineering database, authors noticed that some companies were named "ACME 1", "ACME 2", "ACME 3". According to our premise, each company has only one legal entity name and, therefore, ACMEs 1, 2 and 3 were displayed as three different companies after running our computational algorithm. However, any human would notice that they are in fact the same company, since they share the same characteristics, but it is a limitation for the algorithm. Future research could develop an improved version of the algorithm that would not adopt the premise and would expand the approach to business groups that have more than one company name. Petrobras Corporation (which is not an AFC for obvious reasons), for instance, has about 40 different company names. An algorithm that identifies business groups given a list of company names could be used before the processing stage and potentially improve the solution our algorithm arrives at.

### Limitation on the UNID and its uses

Data availability from the alumni institution plays an important role in the results obtained from the proposed algorithm. As mentioned before, 33% of alumni from the Federal University of Rio de Janeiro had either missing or invalid UNIDs registered in the database, and most of the missing data belong to students registered before 2001. Reliability of the results can also be affected by other aspects related to data quality. Until the end of the 1990s, UFRJ allowed students to enroll to the university using their parents' UNID, when they were minors and were not legally obliged to have an UNID of their own. This fact plus the lack of identified UNID numbers forced us to do the matching between databases based not only on the UNID, but also on the person's name. Matching of names could only be skipped if the university database has high data quality and access to entire UNID numbers is provided by the IRS.

Another common practice in the Brazilian labor market that brings inaccurate, incomplete, or inconsistent data to the algorithm is what is commonly referred to in Brazil as "PJotização" and is referred throughout the text as a "CTIN as a service provider for another company". Its creation is attributed to an event of hiring workers as entities, thus escaping from protective labor regulation. Therefore, such companies are a special case and might not be considered an AFC. Identifying this practice among other companies is a hard task. Authors proposed a way that can lead both to false positives and false negatives but which, in their evaluation, is better than leaving the data as is and considering all companies approved by the L2 filter as AFC. Since the changing of the Brazilian Labor Law in 2002, there is a rise in the number of "CTIN as a service provider for another company" and the analysis deserves a more specific study to identify further ways and criteria to differentiate between companies per se and this workaround.

In a second group of cases, the alumni UNID is used in founding a new company that is not related to his/her field of knowledge. This happens either because the person herself found an opportunity for a franchising or another business unrelated to her former studies, or a relative or a friend needs a partner for a new business. This latter situation was more frequent until 2011, since it was not possible to establish a single-partner company. The creation of the category "Empresa Individual de Responsabilidade Limitada" (EIRELI), "Empresário Individual" and "Sociedade Limitada Unipessoal" allowed opening of new companies in several economic areas with no need of a partner.

Finally, the last group of cases can be described as a variation of the former one. Authors observed that an alumni UNID takes part in founding a company and eventually becomes its administrator, when the other person involved in the founding team is a public agent. This special case happens because Law 8112/90 allows public agents to be partners in private companies but forbids them from being the manager of the business. Therefore, for the same reason, when an alumnus becomes a public agent, he/she must partner with a non-public agent in founding a new company (a case that was also observed in our sample).

## 4. CONCLUSION

Brazilian universities, especially public institutions, bear the responsibility to demonstrate their contribution to the collective welfare. University's influence can be evaluated both directly - through the quantity and caliber of their alumni, as well as traditional research benchmarks - and indirectly, such as gauging the economic endeavors ignited by their graduates worldwide. This paper introduces an algorithm designed to be adaptable for any Brazilian academic establishment, with potential adaptations in datasets for use in other countries. The algorithm serves to quantify the economic and societal influence of academic entrepreneurs. Its applicability is versatile. Initially, it can bolster undertakings aimed at evaluating the returns on universities' investments.

Furthermore, it can guide internal efforts to reshape curricula by contrasting intended and actual entrepreneurial engagement in different courses. Lastly, through analyzing logically improbable yet frequently occurring scenarios, universities can assess the secondary effects of education and professional practice. For instance, cases wherein professionals who graduated from saturated market segments evolve into small business entrepreneurs or analogous roles.

## REFERENCES

- Brazil (2016), "LEI Nº 13.243, DE 11 DE JANEIRO DE 2016 DA PRESIDÊNCIA DA REPÚBLICA". Retrieved from [https://www.planalto.gov.br/ccivil\\_03/\\_Ato2015-2018/2016/Lei/L13243.htm](https://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2016/Lei/L13243.htm). Last access March 2023.
- Receita Federal do Brasil (2023), "Portal de Dados Abertos - CNPJ". Retrieved February 22, 2023, from <https://dadosabertos.rfb.gov.br/CNPJ/>
- Callan, B. (2001), "Generating spin-offs: evidence from across the OECD". *STI Review: Special Issue on Fostering High-Tech Spin-Offs: A Public Strategy for Innovation*.
- Compagnucci, L. and Spigarelli, F. (2020), "The Third Mission of the university: A systematic literature review on potentials and constraints", *Technological Forecasting and Social Change*, 161, 120284.
- Conceição, O., Faria, A. P. and Fontes, M. (2017), "Regional variation of academic spinoffs formation", *The Journal of Technology Transfer*, 42, 654–675.
- Forbes Business Councils Member (2018). Eight important elements of successful business partnership. *Forbes*. Retrieved from <https://www.forbes.com/sites/forbesbusinessdevelopmentcouncil/2018/12/18/eight-important-elements-of-successful-business-partnerships/?sh=6123a336128d>
- Forbes Councils Member (2021), "Eight Essential Qualities Of A Successful Business Partner". *Forbes*. Retrieved from <https://www.forbes.com/sites/theyec/2021/07/15/eight-essential-qualities-of-a-successful-business-partner/?sh=685365e9374f>
- Freitas, J. S., Gonçalves, C. A., Cheng, L. C. and Muniz, R. M. (2011), "The academic spin-off phenomenon: structuring a new field of research in Brazil", *Innovation & Management Review*, Vol. 8, No. 4, 67.
- Galindo-Rueda, F. and Verger, F. (2016), "OECD taxonomy of economic activities based on R&D intensity", *OECD Science, Technology and Industry Working Papers*, 4, 24.
- Hoidn, S. and Kärkkäinen, K. (2014), "Promoting Skills for Innovation in Higher Education. A literature-review on the effectiveness of problem-based learning and of teaching behaviours". *OECD Education Working Papers*, No. 100 (January 2014). <https://doi.org/10.1787/5k3tsj671226-en>
- Klein, J.D. (2015), "Teaching and learning limits in engineering education", *Engineering Studies*, Vol. 7, No. 2–3, pp. 135–137. <https://doi.org/10.1080/19378629.2015.1062494>
- Lebret, H. (2017), "Startups and Stanford University". *ArXiv Preprint ArXiv:1711.00644*.
- Li, H., Yang, X. and Cai, X. (2021), "Academic spin-off activities and research performance: the mediating role of research collaboration", In *Journal of Technology Transfer*. <https://doi.org/10.1007/s10961-021-09869-y>
- Marins, B. V. de S., Ramos, H. C., Ferreira, G. de S., Costa, S. R. R. and Costa, H. G. (2019), "Interdisciplinarity in Higher Education: a Cross-Sectional Analysis of the Literature in the Period 2014-2018", *Brazilian Journal of Operations and Production Management*, Vol. 16, No. 1, pp. 113–125. <https://doi.org/10.14488/BJOPM.2019.v16.n1.a11>
- Meoli, M., Pierucci, E. and Vismara, S. (2018), "The effects of public policies in fostering university spinoffs in Italy", *Economics of Innovation and New Technology*, Vol. 27, No. 5–6, pp. 479–492.
- Mowery, D. C., Nelson, R. R., Sampat, B. N. and Ziedonis, A. A. (2004), "Ivory Tower and Industrial Innovation: University-Industry Technology Transfer Before and After the Bayh-Dole Act in the United States".
- O'shea, R. P., Allen, T. J., Chevalier, A. and Roche, F. (2005), "Entrepreneurial orientation, technology transfer and spinoff performance of US universities". *Research Policy*, Vol. 34, No. 7, pp. 994–1009.
- Phan, P. H. and Siegel, D. S. (2006), "The effectiveness of university technology transfer". *Foundations and Trends® in Entrepreneurship*, Vol. 2, No. 2, pp. 77–144.
- Roberts, E. B., Murray, F. and Kim, J. D. (2019), "Entrepreneurship and innovation at MIT: Continuing global growth and impact - An updated report", *Foundations and Trends in Entrepreneurship*, Vol.

- 15, No. 1, pp. 1–55. <https://doi.org/10.1561/03000000093>
- Rothaermel, F. T., Agung, S. D. and Jiang, L. (2007), "University entrepreneurship: a taxonomy of the literature", *Industrial and Corporate Change*, Vol. 16, No. 4, pp. 691–791.
- Shane, S. A. (2004), "*Academic entrepreneurship: University spinoffs and wealth creation*". Edward Elgar Publishing.
- Täks, M., Tynjälä, P., Toding, M., Kukemelk, H. and Venesaar, U. (2014), "Engineering students' experiences in studying entrepreneurship. *Journal of Engineering Education*, Vol. 3, No. 4, pp. 573–598. <https://doi.org/10.1002/jee.20056>
- Ten Caten, C. S., Silva, D. S., de Aguiar, R. B., Filho, L. C. P. S. and Huerta, J. M. P. (2019) , "Reshaping Engineering Learning To Promote Innovative Entrepreneurial Behavior. *Brazilian Journal of Operations and Production Management*, Vol. 16, No. 1, pp. 141–148. <https://doi.org/10.14488/BJOPM.2019.v16.n1.a13>
- Unicamp (2020). *Relatório de empresas-filhas*. Retrieved from <https://unicampventures.org.br/wp-content/uploads/2021/03/Relatório-de-empresas-filhas-da-Unicamp-2020-v.2.pdf>
- Unicamp companies. (n.d.). Retrieved from <https://www.inova.unicamp.br/2022/10/empresas-filhas-da-unicamp-faturaram-juntas-r-193-bilhoes-em-2022/#:~:text=Os dados são do mapeamento,Agência de Inovação Inova Unicamp.&text=As empresas-filhas da Universidade,relação ao ano de 2021>
- Uziel, D. and Allonso, D. (2022), "Interação academia-indústria: Métodos de avaliação e caso de estudo na área biomédica. *E-Papers*, Vol. 1, No. 1, 126.
- Vefago, Y. B., Trierweiller, A. C. and de Paula, L. B. (2020), "The third mission of universities: the entrepreneurial university", *Brazilian Journal of Operations and Production Management*, Vol. 17, No. 4, pp. 1–9. <https://doi.org/10.14488/BJOPM.2020.042>

**Author contributions:** DZ: Actively participated in the conception and planning of the project. Contributed to writing and revisions of the manuscript for submission; ERPS: Actively participated in the conception and planning of the project. Contributed to writing and revisions of the manuscript for submission; HHA: Actively participated in the conception and planning of the project. Contributed to writing and revisions of the manuscript for submission.