Deterministic bibliometric disambiguation challenges in company names

Andrea Belz
Viterbi School of Engineering
University of Southern California
Los Angeles, California USA
abelz@usc.edu

Alexandra Graddy-Reed
Sol Price School of Public Policy
University of Southern California
Los Angeles, California USA
graddyre@usc.edu

FNU Shweta
Viterbi School of Engineering
University of Southern California
Los Angeles, California USA
s779682@usc.edu

Aleksandar Giga
Faculty of Technology, Policy and Management
Delft University of Technology
Delft, the Netherlands
A.Giga@tudelft.nl

Shivesh Meenakshi Murali
Viterbi School of Engineering
University of Southern California
Los Angeles, California USA
smeenaks@usc.edu

***Abstract*—Peer-reviewed publications and patents serve as important signatures of knowledge generation, and therefore the authors and their organizations can represent agents of intellectual transformation. Accurate tracking of these players enables scholars to follow knowledge evolution. However, while author name disambiguation has been discussed extensively, less is known about the impact of organization name on bibliometric studies. We expand here on the recently defined phenomenon of "onomastic profusion," high-frequency words used in organization names for semantic reasons, and thus contributing a non-random source of error to bibliographic studies. We use the Small Business Innovation Research (SBIR) Phase I awardees of the National Aeronautics and Space Administration (NASA) as a use case in the field of engineering innovation. We find that firms in California or Massachusetts experience a six percent decrease in the likelihood of using the word "Technologies" in their names. Furthermore, use of the words "Research" and "Science" is linked to doubling the number of awards. We illustrate that, in aggregate, firms executing rational strategic naming decisions can create deterministic bibliometric challenges.**

***Index Terms*—disambiguation, names, patents, NLP, bibliometric, NASA, SBIR**

Introduction

The creation and evolution of technology plays an important role in economic growth [1], and at the earliest stages it can be traced through dissemination processes, such as publication and patenting [2]–[5]. This tracking process requires identification and labeling of each named entity (NE) participating as an agent. To date, significant attention has been paid to challenges in author identification [6]–[8], such as gender prediction and author nationality [9].

Two key considerations are the appearance of *synonyms*, a single name in multiple forms, and *homonyms*, a name shared by different entities [10]. The latter issue is the principal concern of this study. Bibliometric onomastics, the study of proper names, has not yet focused on entity names and the associated impact on accurate disambiguation.

Unlike individual names, organizational entity names are not randomly assigned but instead are deterministic, as they result from important branding decisions [11]–[13]. Competitive equilibrium theory predicts that firms using existing words will be attracted to the same terms [14]. This strategy, pursued in parallel by multiple companies, leads to certain words appearing in firm names at high frequency. In this paper, we expand on the first reports of this phenomenon, termed "onomastic profusion" [15]. We use a publicly available list of recipients of the National Aeronautics and Space Administration (NASA) Small Business Innovation Research (SBIR) Phase I awards as a relevant sample to demonstrate how naming strategies impact bibliometrics.

Names and marketing strategy

Fundamentally, naming a new venture is an exercise in addressing the "liability of newness" [16] in order to generate legitimacy with potential stakeholders. Moreover, names impart information on the quality of the firm [17]. Therefore, name choice is an important decision [11]–[13]. Kohli and Suri describe the set of possible strategies for brand names in general [18]:

- Descriptive ("General Motors" cars)
- Suggestive ("Mr. Clean" cleaners)
- Arbitrary ("Apple" computers)
- Coined ("Microsoft" software)

These four alternatives can be categorized simply as *meaningful* (descriptive or suggestive) or *non-meaningful* (arbitrary or coined) [19]. How should firms select from this list of alternatives?

Marketing scholars have identified factors that may inform this decision. Meaningful names offer two benefits in that they are both easier to recall and they create better responses than

This research was funded in part by the National Science Foundation (NSF) I-Corps awards 1444080 and 1740721. Any opinions, findings, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the aforementioned organizations.

their non-meaningful counterparts [18]. Furthermore, brand names that suggest a benefit lead to better recall of the indicated benefit [20]. The impact of names goes beyond product lines to corporate finance. Early-stage investors prefer ventures with unique names, but investors in more mature firms - where the risk is lower - prefer simpler names [21].

Consequently, leaders of firms in a crowded marketplace should select meaningful, simple names that convey benefits and lower risk. If a market participant identifies a strategic option that confers an advantage, equilibrium theory predicts that all competitors will make the same selection [14]. In the technology innovation context, this effect manifests as companies converging on a small set of words implying applied technical research. Indeed, as more firms select from a limited word palette, other market players may perceive lower risk simply through the ubiquity of those words. To some extent the words lose their meaning *in extremis*; for instance, if everyone selects the word "unique" then it becomes a poor descriptor, as noted by Klink [22].

Despite this semantic dilution, as a market reaches equilibrium, companies naturally move toward selecting the same words in their names. This equilibration is the origin of the previously observed onomastic profusion. It generates a new set of homonyms that can confound clustering or manual (supervised) disambiguation processes [15], similarly to the prevalence of certain names among Chinese authors [10].

Importantly, unlike author name disambiguation problems that occur randomly, such as misspelling or variable use of middle initials, onomastic profusion causes deterministic clustering challenges because firm name selection is a deliberate strategic choice. Therefore, errors in identifying intellectual assets could be correlated with other observable firm characteristics.

DATA ANALYSIS

Corpus generation

To demonstrate the challenge, we use the publicly available set of awardees of the NASA SBIR program, a flagship initiative of the United States government to support small businesses as performers of federal research. Outcomes have been discussed extensively in the literature as contributions to the national ecosystem [23]–[25]. Awards are made in two Phases, with a Phase I award required as eligibility for Phase II. Because Phase I selection is associated with evaluations of legitimacy [26], we hypothesized that the effects of naming decisions would appear in the Phase I award pool.

We downloaded[1] the complete list of Phase I awardees and pre-processed them to remove entity identifiers such as "Inc.", "LLC", and "Incorporated." Because records included the firm address, we also created a binary flag set to one (and zero otherwise) if the firm was located in California or Massachusetts, as these states report high levels of success in innovation and small business growth [24], [27], [28]. Firms from these two states represented 30.4% of our sample. Furthermore, records included the number of Phase I awards[2] associated with the firm throughout the program's history. A corpus of 3,463 firm names was generated.

Word frequencies

We generated a word cloud of the corpus (Fig. 1) and discovered that, as predicted, certain words associated with technological innovation appeared at high frequency. We studied this distribution via binary flags set to one if a given word appeared in a firm's name; these were not exclusive, as discussed further below. The frequencies indicated that the word "Technologies" appeared in almost nine percent of the firm names, and "Technology" appeared in more than five percent (Table I) - in other words, as a stemmed token this term was used in roughly one of seven firm names. Similarly, "Systems" appeared in more than seven percent of the names. These words generally connoted technical innovation.

Fig. 1: Word cloud of corpus of names of firms receiving NASA SBIR Phase I awards.

TABLE I: Frequencies of words appearing in NASA SBIR Phase I awardee names

Word	Percent
Technologies	8.75
Systems	7.59
Research	6.18
Technology	5.54
Engineering	3.87
Science	2.19
Advanced	2.17
Solutions	2.11
Applied	1.53
Scientific	1.18

Notes: The total number of observations was 3,463. The data were downloaded from the SBIR web site as NASA SBIR Phase I awardees. Firm names were pre-processed to remove entity identifiers such as "Inc." and "LLC."

Many of the firm names used more than one of the words, presumably to hedge their bets in selecting strategic options. Examples of co-occurrence include: Advanced Cooling Technologies, Inc.; Advanced Materials Technology, Inc.;

[1] sbir.gov

[2] Phase II awards were not included in this variable.

Advanced Systems and Technologies, Inc.; and Advanced Science and Novel Technology. We explored this co-occurrence through a correlation matrix (Table II) and found that that the words “Applied” and “Science” had a statistically significant correlation coefficient of 13 percent, consistent with the goals of a program designed to bring engineering discovery to the federal government. “Applied” was also associated with “Research” ($\rho = 9\%$). This adjective-noun correlation has important implications in bibliometric scholarship [29]. On the other hand, “Technologies” was generally anticorrelated with most of the other words considered here.

TABLE II: Correlations between words

	Technologies	Systems	Research	Technology	Science	Advanced
Technologies						
Systems	-0.08***					
Research	-0.07***	-0.06**				
Technology	-0.08***	-0.05**	-0.03			
Engineering	-0.05**	0.01	-0.02	-0.05**		
Science	-0.04*	-0.03	-0.02	0.04*		
Advanced	0.04*	-0.01	-0.01	0.02	-0.01	
Solutions	-0.05**	-0.04*	-0.02	0.02	-0.02	0.05**
Applied	0.02	-0.01	0.09***	0.03	0.13***	-0.02
Scientific	-0.01	-0.02	0.02	-0.03	-0.02	0.00

Notes: The total number of observations was 3,463. The data were downloaded from the SBIR web site as NASA SBIR Phase I awardees. Firm names were pre-processed to remove entity identifiers such as “Inc.” and ”LLC.” Binary variables were set to one if the word appears in the firm’s name. The words “Engineering,” “Solutions,” and “Applied” are excluded from this table because all correlation coefficients not yet noted in the rows were statistically insignificant. *p<0.1; **p<0.05; ***p<0.01

Impacts

To explore the impacts more thoroughly, we conducted regressions on two different outcomes (Table III). We first examined the association of specific words for firms located in California or Massachusetts, reasoning that those firms would have less need to demonstrate legitimacy in their names. We conducted a logistic regression to estimate the probability of the firm’s location in California or Massachusetts (model 1) based on its firm name and found that the word “Technologies,” previously shown to be the most frequently used word, was negatively associated with the likelihood of location in one of these two key states with an average marginal effect (AME) estimated at -6.51% ($p < 0.028$); the word “Technology” does not show this effect. A weaker signal was seen for firms using the word “Solutions,” with a decreased probability of 9.97% ($p < 0.099$). On the other hand, the word “Applied” was associated with an increase of 10.5% ($p < 0.082$) in the probability of being located in California or Massachusetts. In other words, the most common word (“Technologies”) connoting innovation was selected by firms in states reporting lower innovation measures. This is consistent with naming as a strategy to confer legitimacy.

In model 2 of Table III, we estimated the firm’s number of Phase I awards with linear regressions. The covariates are the binary variables indicating if the word appears in the firm’s name. We found that the word “Research” in the firm’s name is positively associated with a doubling of the number of Phase I awards, as are both “Science” and “Scientific”. In robustness checks, we confirmed that this was not mediated by a firm’s presence or absence in California or Massachusetts nor by other words.

TABLE III: Estimation of word use in firm name

	Outcome	
	CA or MA	Number of awards
	(1)	(2)
Technologies	-0.308^{**} (0.141)	-0.166 (0.495)
Systems	0.180 (0.137)	0.169 (0.524)
Research	0.111 (0.152)	2.366^{***} (0.578)
Technology	0.024 (0.162)	-0.322 (0.607)
Engineering	-0.006 (0.192)	0.276 (0.717)
Science	-0.083 (0.257)	2.977^{***} (0.951)
Advanced	0.059 (0.255)	0.152 (0.949)
Solutions	-0.472^{*} (0.287)	-1.380 (0.963)
Applied	0.498^{*} (0.287)	-0.953 (1.137)
Scientific	0.279 (0.328)	2.363^{*} (1.276)
Constant	-0.823^{***} (0.046)	3.497^{***} (0.170)
Observations	3,463	3,463
Adjusted R^2		0.007
Note:		*p<0.1; **p<0.05; ***p<0.01

Notes: The total number of observations was 3,463. The data were downloaded from the SBIR web site as NASA SBIR Phase I awardees. Firm names were pre-processed to remove entity identifiers such as “Inc.” and ”LLC.” Binary variables were set to one if the word appears in the firm’s name. A logistic functional form was used for model 1 and a linear regression was conducted for model 2. Standard errors are reported in parentheses.

Discussion and future research

In this work we have described why and how onomastic profusion occurs. Consistent with marketing theory, words connoting technical innovation appear with high frequency in a list of recipient names for grants dedicated to this purpose. Moreover, statistically significant correlations are observed between select pairs of these words, implying that the words may not be selected independently. The most commonly used word, “Technologies,” appears in almost one-tenth of the corpus, and it is associated with a 6% decrease in the probability of the firm’s location in the two states with highest innovation performance. Interestingly, the words “Research,” “Science,” and “Scientific” are associated with a two-fold increase in the firm’s number of awards. Naturally the model is a poor fit (Adjusted $R^2 = 0.007$) in that award decisions are based on much more than the firm’s name; however, this study indicates that there may be links between name, strategy, and quality [17]. For instance, these words could be linked to firms pursuing early-stage technologies with low Technology Readiness Levels [30] aligned with the SBIR program’s goals.

These findings are important for NE disambiguation in systems such as PatentsView, the United States Patent and Trademark Office platform that uses a clustering process to aggregate patents for both inventors and assignees [31], [32]. To date, scant attention has been paid to the assignee as a confounding source. This study points to reasons that similar names appear in the database. Indeed, Hotelling's model states that as product lines overlap more, so will the names if they are perceived to provide advantages [14]. As a result, the disaggregation problem should grow as the market becomes more saturated and moves toward equilibrium; in other words, onomastic profusion should scale with the maturity of the market.

Several key areas emerge for further examination in natural language processing scholarship. We did not consider bigrams, or words that co-occur systematically. These co-occurrences are important in business text identifying industries dynamically [33]–[35]. It is not clear if bigrams are an important naming strategy. Certainly the correlation matrix of Table II suggests that the actors are using words in combination.

Our results add to the repository of practices in data processing. We considered all parts of speech. Many business text analyses limit their attention to nouns [35], but the adjectives here ("Advanced," "Applied") contribute to the naming decisions, suggesting the role of broader linguistics considerations in onomastic studies [29]. In addition, we show that stemming would have diluted the geographic link of the word "Technologies" relative to "Technology." Therefore, stemming must be used carefully in pre-processing analysis workflows in onomastics.

In addition, this study contributes to management research. First, we did not explore eponymous firm names - namely, those named after a founder. Belenzon and colleagues report 19 percent of a sample of European firms are eponymous [36]. We see eponymy in only one percent of our sample (estimated thirty observations). However, a significant clustering problem tied to A&P Technology was identified [15]; it was incorrectly associated with several firms, including G&H Technology, M&A Technology, and R&H Technology. It is not clear if these letters are chosen eponymously. Research on the impact of eponymy on firm outcome has shown mixed results [36]–[38]. However, these data suggest that strategic semantic naming decisions outweigh eponymy in this particular population.

Another key reason to understand naming strategy is that firms may change their names, generating synonym challenges to disambiguation. The renaming decision is also typically not random; companies with financial, reputational, or performance problems are more likely to change their names [39], [40]. Therefore, the semantic factors underlying naming strategy can point to solutions in following low-performing firms that change their names.

If the equilibrium model is correct, then in aggregate, firm naming decisions should evolve with the maturity of the field, as suggested previously [13]. For instance, firms conducting semiconductor research may use these strategies more commonly than companies engaged in synthetic biology, a newer industrial sector. This question could be explored directly and linked to federal innovation grants.

A final area of exploration would be to expand this study to Phase II selection, gendered selection effects, and related questions around innovation financing. For instance, do women use words like "Technologies" in their firm names to improve their legitimacy in peer-reviewed processes relevant to innovation, such as grant selection and patents [26], [41], [42]? Organizational names must be considered as another instrument in the strategic tool box [11]–[13], and thus they impact management scholarship and bibliometric studies.

Conclusion

Intellectual asset production can be effectively tracked in publications and patents, and this process can illuminate how value accrues to the host organization. To date disambiguation scholarship has focused on the author rather than the entity. Company naming is a calculated process. Onomastic profusion takes place when many firms select the same words in their names. Here we demonstrate that recipients of technical innovation grants choose words such as "Technology" and "Systems" at frequencies on the order of several percent. Correlation coefficients indicate that the words are selected to suggest innovation, and the most common word, "Technologies," is typically associated with firms from states reporting lower innovation metrics. Use of the words "Research," "Science," and "Scientific" is associated with the firm receiving twice as many awards. Taken together, these findings suggest that firms follow the predictions of competitive equilibrium theory to confer legitimacy and select names for semantic reasons. As a result, this high-frequency word selection underpins previously observed challenges in disambiguation of patent data in a deterministic, non-random way. These results have important implications, particularly for scholars exploring innovation in small companies.

Acknowledgment

A.B. served as co-PI and Research PI on the awards acknowledged herein prior to service at the National Science Foundation. To manage potential conflicts of interest she resigned from all roles associated with these awards and was recused from associated NSF matters. The authors thank Matt Marx, Florenta Teodoris, participants at the 2021 Academy of Management "New Data And Methods In Strategic Management Research" workshop, the 2022 IEEE International Conference on Big Data, and the Management of INnovation, Entrepreneurial Research, and Venture Analysis (MINERVA) lab.

References

[1] P. M. Romer, "Endogenous Technological Change," *Journal of Political Economy*, vol. 98, no. 5, pp. S71–S102, 1990.

[2] M. Marx and A. Fuegi, "Reliance on science: Worldwide front-page patent citations to scientific articles," *Strategic Management Journal*, vol. 41, no. 9, pp. 1572–1594, 2020.

[3] L. Fleming, H. Greene, G. Li, M. Marx, and D. Yao, "Government-funded research increasingly fuels innovation," *Science*, vol. 364, no. 6446, pp. 1139–1141, 2019.

[4] A. Conti, J. Thursby, and M. Thursby, "Patents as Signals for Startup Financing," *Journal of Industrial Economics*, vol. 61, no. 3, pp. 592–622, 2013.

[5] F. Teodoridis, J. Lu, and J. L. Furman, "Measuring the Direction of Innovation: Frontier Tools in Unassisted Machine Learning," 2021.

[6] V. I. Torvik and N. R. Smalheiser, "Author Name Disambiguation in MEDLINE," *ACM Transactions on Knowledge Discovery from Data*, vol. 3, no. 3, pp. 1–7, 2009.

[7] H. Mihaljević and L. Santamaría, “Disambiguation of author entities in ADS using supervised learning and graph theory methods,” *Scientometrics*, vol. 126, no. 5, pp. 3893–3917, 2021. [Online]. Available: https://doi.org/10.1007/s11192-021-03951-w

[8] C. A. D'Angelo and N. J. van Eck, “Collecting large-scale publication data at the level of individual researchers: a practical proposal for author name disambiguation,” *Scientometrics*, vol. 123, no. 2, pp. 883–907, 2020. [Online]. Available: https://doi.org/10.1007/s11192-020-03410-y

[9] P. Roe, G. Lewison, and R. Webber, “The sex and ethnicity or national origins of researchers in astronomy and oncology in four countries, 2006-2007 and 2011-2012,” *Scientometrics*, vol. 100, no. 1, pp. 287–296, 2014.

[10] D. Yin, K. Motohashi, and J. Dang, “Large-scale name disambiguation of Chinese patent inventors (1985–2016),” *Scientometrics*, vol. 122, no. 2, pp. 765–790, 2020. [Online]. Available: https://doi.org/10.1007/s11192-019-03310-w

[11] Y. Engel, R. van Werven, and A. Keizer, “How novice and experienced entrepreneurs name new ventures,” *Journal of Small Business Management*, vol. 60, no. 4, pp. 828–858, 2022. [Online]. Available: https://doi.org/10.1080/00472778.2020.1738820

[12] M. A. Glynn and R. Abzug, “Institutionalizing Identity: Symbolic Isomorphism and Organizational Names,” *Academy of Management Journal*, vol. 45, no. 1, pp. 267–280, 2002.

[13] I. Stoyneva and V. Vracheva, “Demystifying entrepreneurial name choice: insights from the US biotech industry,” *New England Journal of Entrepreneurship*, vol. 25, no. 2, pp. 121–143, 2022.

[14] H. Hotelling, “Stability in Competition,” *The Economic Journal*, vol. 39, no. 153, pp. 41–57, 1929.

[15] A. Belz, A. Graddy-Reed, F. Shweta, A. Giga, and S. M. Murali, “Patentopia: A multi-stage patent extraction platform with disambiguation for certain semantic challenges,” in *The 6th IEEE International Workshop on Big Data for Financial News and Data*, 2022, p. In press.

[16] A. L. Stinchcombe, “Social structure and organizations,” in *Handbook of Organizations*, J. March, Ed. Chicago, IL: Rand McNally, 1965, pp. 142–193.

[17] R. C. McDevitt, “”’A’” business by any other name: Firm name choice as a signal of firm quality,” *Journal of Political Economy*, vol. 122, no. 4, pp. 909–944, 2014.

[18] C. Kohli and R. Suri, “Brand names that work: A study of the effectiveness of different types of brand names,” *Marketing Management Journal*, vol. 10, pp. 112–120, 2000.

[19] L. Todea, “Technology brands inspired by nature,” in *Proceedings of ICONN 3*, 2015, pp. 856–867.

[20] K. L. Keller, S. E. Heckler, and M. J. Houston, “The Effects of Brand Name Suggestiveness on Advertising Recall,” *Journal of Marketing*, vol. 62, no. 1, pp. 48–57, 1998.

[21] C. S. R. Chan, H. D. Park, and P. Patel, “The effect of company name fluency on venture investment decisions and IPO underpricing,” *Venture Capital*, vol. 20, no. 1, pp. 1–26, 2018. [Online]. Available: http://doi.org/10.1080/13691066.2017.1334369

[22] R. R. Klink, “Creating Meaningful New Brand Names: A Study of Semantics and Sound Symbolism,” *Journal of Marketing Theory and Practice*, vol. 9, no. 2, pp. 27–34, 2001.

[23] S. T. Howell, “Financing Innovation: Evidence from R & D Grants,” *American Economic Review*, vol. 107, no. 4, pp. 1136–1164, 2017.

[24] J. Lerner, “The Government as Venture Capitalist: The Long-Run Impact of the SBIR Program,” *The Journal of Business*, vol. 72, no. 3, pp. 285–318, 1999. [Online]. Available: http://www.jstor.org/stable/10.1086/209616

[25] A. Giga, A. P. Belz, R. J. Terrile, and F. Zapatero, “Helping the Little Guy: The Impact of Government Awards on Small Technology Firms,” *Journal of Technology Transfer*, 2021.

[26] A. P. Belz, A. Graddy-Reed, I. Hanewicz, and R. J. Terrile, “Gender Differences in Peer Review of Innovation,” *Strategic Entrepreneurship Journal*, vol. 16, no. 2, pp. 255–280, 2022.

[27] M. H. Best, “Greater Boston's industrial ecosystem: A manufactory of sectors,” *Technovation*, vol. 39-40, no. 1, pp. 4–13, 2015. [Online]. Available: http://dx.doi.org/10.1016/j.technovation.2014.04.004

[28] J. Guzman and S. Stern, “Where is Silicon Valley?” *Science*, vol. 347, no. 6222, pp. 606–609, 2015.

[29] H. Motschenbacher, “Corpus Linguistic Onomastics: A Plea for a Corpus-Based Investigation of Names,” *Names*, vol. 68, no. 2, pp. 88–103, 2020. [Online]. Available: https://doi.org/10.1080/00277738.2020.1731240

[30] A. P. Belz, R. J. Terrile, F. Zapatero, M. Kawas, and A. Giga, “Mapping the ”Valley of Death”: Managing Selection and Technology Advancement in NASA's Small Business Innovation Research Program,” *IEEE Transactions on Engineering Management*, vol. 68, no. 5, pp. 1476 – 1485, 2019.

[31] N. Monath, C. Jones, and S. Madhavan, “PatentsView: Disambiguating Inventors, Assignees, and Locations,” 2021.

[32] N. Monath, S. Madhavan, C. DiPietro, A. McCallum, and C. Jones, “Disambiguating Patent Inventors, Assignees, and their Locations in PatentsView,” Tech. Rep.

[33] F. Shweta and A. Belz, “Computational Linguistic Analysis of Submitted SEC Information (CLASSI),” in *21 IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI)*, 2021, pp. 1132–1137.

[34] G. Hoberg and G. Phillips, “Product market synergies and competition in mergers and acquisitions: A text-based analysis,” *Review of Financial Studies*, vol. 23, no. 10, pp. 3773–3811, 2010.

[35] G. Hoberg and G. M. Phillips, “Text-Based Network Industries and Endogenous Product Differentiation,” *Journal of Political Economy*, vol. 124, no. 5, pp. 1423–1465, 2016.

[36] S. Belenzon, A. K. Chatterji, and B. Daley, “Eponymous entrepreneurs,” *American Economic Review*, vol. 107, no. 6, pp. 1638–1655, 2017.

[37] R. J. Andrews, C. Fazio, J. Guzman, Y. Liu, and S. Stern, “The Startup Cartography Project: Measuring and mapping entrepreneurial ecosystems,” *Research Policy*, vol. 51, no. 2, p. 104437, 2022. [Online]. Available: https://doi.org/10.1016/j.respol.2021.104437

[38] A. Minichilli, A. Prencipe, S. Radhakrishnan, and G. Siciliano, “What's in a Name? Eponymous Private Firms and Financial Reporting Quality,” *Management Science*, vol. 68, no. 3, pp. 2330–2348, 2022.

[39] Q. Feng, Q. Tao, Y. Sun, and M. Susai, “Fresh look or false advertising: Modeling of investor attention based on corporate name changes,” *Finance Research Letters*, vol. 47, no. PA, p. 102526, 2022. [Online]. Available: https://doi.org/10.1016/j.frl.2021.102526

[40] R. C. McDevitt, “Names and reputations: An empirical analysis,” *American Economic Journal: Microeconomics*, vol. 3, no. 3, pp. 193–209, 2011.

[41] T. Bol, M. de Vaan, and A. van de Rijt, “Gender-equal funding rates conceal unequal evaluations,” *Research Policy*, vol. 51, no. 1, p. 104399, 2022.

[42] K. Jensen, B. Kovács, and O. Sorenson, “Gender differences in obtaining and maintaining patent rights,” *Nature Biotechnology*, vol. 36, no. 4, pp. 307–309, 2018. [Online]. Available: http://dx.doi.org/10.1038/nbt.4120