STANDING OUT OR FITTING IN, BUSINESS PITCH AND PIVOT: INSTITUTIONAL ISOMORPHISM AND DIFFERENTIATION STRATEGY

BY

CHUQING ZHAO

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS IN COMPUTATIONAL SOCIAL SCIENCE

Faculty Advisor: James Evans

Preceptor: Shilin Jia

THE UNIVERSITY OF CHICAGO

JULY 2022

Standing out or Fitting in, Business Pitch and Pivot: Institutional Isomorphism and Differentiation Strategy

Chuqing Zhao

MACSS 20-22’, the University of Chicago

JULY 2022

1. RESEARCH QUESTIONS

The game of “standing out” or “fitting in” is pervasive in the social world. In the term of “standing out” strategy, individuals and organizations become dissimilar and differentiated from their peers; in contrast, “fitting in” means that individuals and organizations become homogeneous or resemble the common form. The game of “fitting in ” or “standing out” faced by individuals and organizations alike, navigating between these two ends, and finding a balance become important to the survival and prosperity of the social entity. For individuals, this game is played within the organization that the individual takes part in, such as a company or social group. For organizations, this game is reified in the structure that they engage with, such as an industry. This reification is especially true for business organizations. While innovation emerges from “standing out” from the rest of the industry, such novelty could also alienate the company from the rest of the pack. On the other hand, fitting into the industry also prevents the company from capturing additional market share and becoming the industry leader. How do companies find this balance?

Given the height of U.S. venture capital funding at the moment, at all time high in 20 years mirroring the funding level of the dot-com boom in the 2000s, the question of standing out or fitting in becomes increasingly important to early-stage companies. To partake and benefit from the venture capital carnival, companies often pivot their products and business models to generate more value to attract investors. Though the startup and venture capital community disagree on the key value drivers of a company, there is consensus on the importance of pivots. What are patterns of pivots and companies that pivot, and do they stand out or fit in overtime with these pivots?

In this paper, we investigate this central question through analyzing company description text from Crunchbase (early-stage companies) and annual 10-K's (mature, public companies). There are three dimensions to this analysis. First, how do companies change their pitches overtime? Second, what drives and results from pivoting? Third, how do companies fit in with the rest of the industry they find themselves in – and how is this different between early-stage v.s. mature, public companies? We hypothesize that pitches become more professional overtime, and that early-stage companies are more likely to pivot and disrupt the industry because pivoting is beneficial.

2. THEORETICAL FRAMEWORKS

This tradeoff between fitting in and standing out draws on literatures in organizational theory. Many studies observe a trade-off between fitting in and standing out from the culture perspective (Goldberg, 2016; Askin, 2017) [1][2]. The mechanism between

[1] Goldberg, A., Srivastava, S. B., Manian, V. G., Monroe, W., & Potts, C. (2016). Fitting in or standing out? The tradeoffs of structural and cultural embeddedness. *American Sociological Review*, *81*(6), 1190-1222.

[2] Askin, N., & Mauskapf, M. (2017). What makes popular culture popular? Product features and optimal differentiation in music. *American Sociological Review*, *82*(5), 910-944.

differentiation and integration strategy is linked with entrepreneur identity construction, network structure, and culture fitness (Navis & Glynn, 2010[3]; Navis & Glynn, 2011[4]; Granqvist et al., 2013[5]). First, the choice of fitting in is a problem of organizational isomorphism. In terms of homogeneity, sociological theories have long sought to explain why organizations are very similar to one other, and how fitting into the market could benefit firms' profitability. Earlier researches focus on the technical and rational forces that shape organizations and therefore similarities among organizations were thought to be a consequence (Weber, 1952[6]). Accordingly, rationality can control human behavior leading to homogeneity in structure, culture, and outputs (Giddens, 1979[7]). Sociological research has found various empirical evidence: Coser, Kadushin and Powell (1982)[8] describe the evolution of American college textbook publishing from a period of initial diversity to the current hegemony largely due to the large bureaucratic generalist and a small number of specialists. Starr (1980)[9] provides evidence of mimicry in the development of the hospital field. More modern theories

[3] Navis, C., & Glynn, M. A. (2010). How new market categories emerge: Temporal dynamics of legitimacy, identity, and entrepreneurship in satellite radio, 1990–2005. Administrative Science Quarterly, 55(3), 439-471.

[4] Navis, C., & Glynn, M. A. (2010). How new market categories emerge: Temporal dynamics of legitimacy, identity, and entrepreneurship in satellite radio, 1990–2005. Administrative Science Quarterly, 55(3), 439-471.

[5] Granqvist, N., Grodal, S., & Woolley, J. L. (2013). Hedging your bets: Explaining executives' market labeling strategies in nanotechnology. *Organization science*, *24*(2), 395-413.

[6] Weber, Max 1952 The Protestant Ethic and the Spirit of Capitalism. New York: Scribner. 1968 Economy and Society: An Outline of In- terpretive Sociology. Three volumes. New York: Bedminster.

[7] Giddens, A. (1979). Agency, structure. In *Central problems in social theory* (pp. 49-95). Palgrave, London.

[8] Coser, L. A., Kadushin, C., & Powell, W. W. (1982). *Books: The culture and commerce of publishing* (p. 22). New York: Basic Books.

[9] Starr, P. (1980). Medical care and the boundaries of capitalist organization. *Unpublished manuscript. Program on Non-Profit Organizations, Yale University, New Haven, CT*.

incorporate the interdependence among organizations. DiMaggio and Powell (1983)[10] observe that organizations become increasingly homogeneous over time. The authors theorize three mechanisms through which organizations demonstrate isomorphism: coercive, mimetic, and normative.

More specifically for businesses, the question of fitting in or standing out is often directly tied to generating returns on the stock market and tangible value for shareholders. Zuckerman (1999) believes that organizations being perceived as belonging to certain product categories was important for the stock market performance of companies. Companies outside the "zone of recognition" of a particular industry classification were forced by organizational isomorphism to fit in within an industry. However, if a differentiation occurs that alters the entire industry, this would bring significant financial returns to the company.[11]

In contrast, many studies also emphasize the importance of standing out from competitors. According to the resource-based view of the firm, competitive advantage is critical to firms' future performance: the theory assumes that firms can be conceptualized as bundles of resources and that resources are heterogeneously distributed across firms. Combining resources that cannot be easily replicated or mimicked by competing firms can help the firm gain competitive advantages. It has emphasized the importance of competition, and identity variables of "the differences

---

[10] DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American sociological review*, 147-160.

[11] Zuckerman, Ezra W. 1999. "The Categorical Imperative: Securities Analysts and the Illegitimacy Discount." *American Journal of Sociology* 104 (5): 1398–1438. https://doi.org/10.1086/210178

that make a difference on organizations' performance", that is the strategies of firms are dissimilar in "meaningful" ways. As a result, the differentiation could establish comparative advantages in market competition and therefore firms possess a distinctive competence to have skills, expertise, knowledge, or technology which is superior or different from competitors (Selznick, 1957[12]; Porter, 1985[13]). As firms have different capabilities to access scarce resources, firms have different performances on profitability. This view is largely constrained by the assumption that these resources are unique, scarce and finite.

However, the resource-based theory fails to explain the firms' ecology between standing-out and fitting-in strategy in a dynamic market, and also lacks empirical grounding (Williamson, 1999[14]; Priem and Butler, 2000[15]). Extending the understanding of comparative advantage theory, Teece et al. (1997)[16] and Eisenhardt & Martin (2000)[17] argues that the resource-based theory would break down in high-velocity markets, where the duration of comparative advantage is unpredictable. In the context of a stable industry structure, organizations rely extensively on existing knowledge and resources to produce predictable outcomes. The firm's processes of using resources to match and even create the market change are defined as "dynamic

---

[12] Selznick, P. (1957). Law and the Structures of Social Action.

[13] Porter, M. E. (2011). *Competitive advantage of nations: creating and sustaining superior performance*.

[14] Williamson, O. E. (1999). Strategy research: governance and competence perspectives. *Strategic management journal*, *20*(12), 1087-1108.

[15] Priem, R. L., & Butler, J. E. (2001). Is the resource-based "view" a useful perspective for strategic management research?. *Academy of management review*, *26*(1), 22-40.

[16] Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic management journal*, *18*(7), 509-533.

[17] Eisenhardt, Kathleen M., and Jeffrey A. Martin. "Dynamic capabilities: what are they?." *Strategic management journal* 21, no. 10-11 (2000): 1105-1121.

capabilities".

Adaptability—being able to recognize and deal with changes can be an important component of successful entrepreneurial strategies, especially for early-stage firms that normally lack fewer resources, mature products, or a well-defined corporate vision. Nascent firms face uncertainty in technology, product definition and ambiguous characteristics. Consequently, to establish organizations' identities, nascent firms would experience more experiential and unstable processes than later-stage companies. Acquiring resources is critical to firms' success to establish legitimacy by adjusting firms' vision according to investors' expectations. The financial investment could be a driver for firms to choose between "fitting in" and "standing out" strategies. For instance, institutionalists and ecologists believe that external forces from isomorphic pressure would shape organizations to become similar to the rest units while other threads of theory in entrepreneurship claim that unique features of organizations serve as touchstones for legitimacy and innovation (Navis & Glynn, 2010). Learning mechanisms in nascent firms would occur during the process of dealing with the expectation of investors and markets.

In the social process of sensegiving and sensemaking, organizations' linguistic frames of their activities are critical to establishing organizations' identity and legitimizing innovation. Pitch, the business self-description could be an indicator that reflects a firm's business strategy, product, or self-identity. Understanding different version of linguistic changes in pitches allows us to trace the changes in business strategy across different business life cycles.

Hypothesis 1: Early-stage firms would receive more financial capital if they update their pitches, especially recognizing and dealing with the uncertainty and investors' expectation.

Considering the process of revising a research paper, individuals make evaluations of whether and how they should update after consulting with others in their networks. Similarly, social interaction in a given network would influence the revision of pitches and business proposals (Couzin-Frankel, 2013[18]; Lee et al., 2013; Huang and Pearce, 2015[19]; Wu, 2016[20]; Greenberg2019a, Bian et al.2021[21]). In reality, not all firms have the same opportunities for social interaction and receiving feedback or the ability to make use of the opportunities to engage others. Different network positions lead to different outcomes—if occupying a position that allows for information flows or places at the center of the social network, then social actors can access more information and opportunities and cognitive freedom to resist peer effects (Burt, 2010). In the context of early-stage companies, venture capital would provide firms with more information and guidance on how to deal with the uncertainty, if they are at a central position in the social network. Receiving information from disconnected networks or having diverse

[18] Couzin-Frankel, J. (2013). Secretive and subjective, peer review proves resistant to study.
[19] Huang, L., & Pearce, J. L. (2015). Managing the unknowable: The effectiveness of early-stage investor gut feel in entrepreneurial investment decisions. *Administrative Science Quarterly*, *60*(4), 634-670.
[20] Wu, A. (2016). Organizational decision-making and information: Angel investments by venture capital partners. In *Academy of Management Proceedings* (Vol. 2016, No. 1, p. 11043). Briarcliff Manor, NY 10510: Academy of Management.
[21] Bian, J., Greenberg, J., Li, J., & Wang, Y. (2022). Good to go first? Position effects in expert evaluation of early-stage ventures. *Management Science*, *68*(1), 300-315.

investment portfolios could help invested firms to integrate various resources.

Homogeneity can be viewed as a touchstone for legitimacy and conformity (DiMaggio & Powell, 1983 [22]) while heterogeneity can indicate innovations or differentiation from the prototype. Innovation stems from the process of legitimization process of new categories (Khaire, 2015[23]). The optimal differentiation is described as a combination of traditional features and some degree of novelty distinguishing them from their peers (Askin & Mauskapf, 2017[24]). New categories that are often regarded as a combination of legitimacy and distinctiveness do not only align with the expectation of consumers and producers but also have been institutionalized by other actors and intermediaries (Rosa et al., 1999[25]; Kleinbaum, 2012[26]). For example, Khaire (2015) argues that "modern Indian art", a new art category that emerged in the last decades, have been established with the support of art historians, museum, action institutions, etc. When new categories have emerged, the entrepreneurial identity in a nascent market would be definitional for the new category, and it emphasizes the differences from already established market categories but aligns with institutionalized understanding (Navis & Glynn, 2010[27]). The collective identity serves as a foundation

---

[22] DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American sociological review*, 147-160.

[23] Khaire, M., & Wadhwani, R. D. (2010). Changing landscapes: The construction of meaning and value in a new market category—Modern Indian art. *Academy of Management Journal*, *53*(6), 1281-1304.

[24] Askin, N., & Mauskapf, M. (2017). What makes popular culture popular? Product features and optimal differentiation in music. *American Sociological Review*, *82*(5), 910-944.

[25] Rosa, J. A., Porac, J. F., Runser-Spanjol, J., & Saxon, M. S. (1999). Sociocognitive dynamics in a product market. Journal of marketing, 63(4_suppl1), 64-77.

[26] Kleinbaum, A. M. (2012). Organizational misfits and the origins of brokerage in intrafirm networks. *Administrative Science Quarterly*, *57*(3), 407-452.

[27] Navis, C., & Glynn, M. A. (2010). How new market categories emerge: Temporal dynamics of legitimacy, identity, and entrepreneurship in satellite radio, 1990–2005. Administrative Science Quarterly, 55(3), 439-471.

for firms to tailor their unique identities within the already-known meaning systems. It is critical for nascent firms to navigate between shared identities with their peers while preserving their individual distinctiveness.

Hypothesis 2: There is a trade-off between fitting in and standing out strategy. When a new market category achieves legitimacy, entrepreneurial organizations will shift from collective identities to unique identities for individual broadcasters.

Market dynamics could be a potential mechanism to explain the game of fitting in and standing out. Achieving legitimacy or not will affect the organization's strategy. In a nascent market, a firm would like to imitate others because they are not certain whether the innovation would succeed or not. Without sufficient legitimacy, deviating from the others is less likely to be judged as plausible from the perspective of investors. Second, in high-velocity markets where industry structure is blurring, firms would rely on quickly obtaining new information and meeting the expectation of markets, and therefore standing-out strategy would be more likely to occur and become risky.

3. DATA AND METHODS

Text and Other Relevant Data

To test our hypothesis, we collect data from two major resources that cross-validate each other: Crunchbase, and VentureXpert from Thomson Reuters, which is a commercial database serving entrepreneurs and investors (N=298,934 by 2020). While both datasets provide startup-related information (company description, financial performance, etc.), Crunchbase allows for crowdsourcing, where founders and companies can update their profiles. This possibly results in more complete and reliable records.[28] We constructed our research data in three steps.

First, from the Crunchbase dataset, we obtained all the private firms founded from 1950 to 2020 that have a website and have received some form of financing. We were able to obtain a total of 298,934 companies. For each company, we downloaded all data columns available on Crunchbase, including company name, headquarter location, regions, diversity, revenue range, founding date, number of investments, total equity funding amount, money raised by IPO, and so on. We collect firms' funding transactions (N = 178,881) and acquisition transactions (N = 51,217).

Second, we used Wayback Machine, a digital library of persevering archived copies of web pages overtime, to recover the initial website of each startup from the time of founding up to December, 2020. Even though the data is automatically updated, we have recovered Web archives that identify prior business descriptions and their timestamps, which we assembled into a database of signaled pivots (N=397,726). Wayback Machine provides access to a digital archive of over 330 billion web-page snapshots, and for our project dataset, it takes snapshots for the same webpage at least one time from 2014 to 2020. We developed a web-scraping pipeline to automatically obtain all archived webpages, tracing back to the earliest version of the webpage in Crunchbase websites. We first downloaded the homepage and the first level in the

---

[28] Dalle, J., M. den Besten and C. Menon. 2017. "Using Crunchbase for economic and managerial research", OECD Science, Technology and Industry Working Papers, No. 2017/08, OECD Publishing, Paris. https://doi.org/10.1787/6c418d60-en

webpage (fixed URLs suffixes), and we analyzed the web structure changes over time. Then we parsed raw HTML, divided into different parsing versions with the consideration of different time periods and web-page structures. We excluded all pages with HTTP error. After this pre-processing, our final dataset consists of total webpages of 39,668 unique companies.

Third, we identified organization webpages using Crunchbase data with company names. To ensure the best data quality, we narrowed down our time window from 2016 to 2020. We were able to establish record linkages for 22,037 unique companies. Then, we derived a panel dataset to characterize the longitudinal observations of startups, where we separately created a yearly and a quarterly panel from each company's first month/year, up to the last time period of our panel dataset. We also gathered 8,189 10-k companies business descriptions from 10-k files scraped from Edgar API using R.

Additionally, we imputed gender information based on the founder's first name. To minimize missing data, we used the Social Security Administration (SSA) baby name dataset[29]. As foreign names such as Asian names are not included in the SSA dataset, we use Gender API[30] that supports a total of 6,084,389 names across 189 countries. We calculated the gender assignment likelihood for each name — if the probability of "male" is greater than "female", then we assign "male" to the name according to the given source.

---

[29] Office of Chief Actuary. 2022. "Top 10 Baby Names of 2020." Social Security Administration. https://www.ssa.gov/oact/babynames/

[30] Gender API. 2022. https://gender-api.com

Methods

Measuring Funding

The key variable in this study is the *total amount of money raised* in the funding transaction of a startup, since market-based valuation measures are not applicable to privately held startups.[31] We use the total amount of funds received by a startup to represent a startup's value.

Measuring Pivots

Our measurement approach builds on the idea that startup pitch reflects their strategies in business communication. The text similarity is a good indicator of how distant from one pitch to another. We measure distance among company statements by a standard text-analysis algorithm that allows us to quantify the relatedness of those statements to effectively create a measure of cosine similarity between nearby pitches for each startup. We project the pitch on high-dimensional vector space and calculate the cosine similarity between two pitches.

Measuring Linguistic Features

Pitch change is then simply the inverse of similarity. Social innovations are embedded and diffused in human language. To measure the direction of linguistic changes, we use the Brysbaert Concreteness Index (BCI). The BCI relies on abstraction norms that were created for 40,000 commonly used word lemmas in contemporary

---

[31] Lin, Yu-Kai, Likoebe M. Maruping. 2021. "Open Source Collaboration in Digital Entrepreneurship." *Organizational Science* 33(1): 212-230.

English.[32] This concreteness data is collected from crowdsourcing on 4,000 participants who rated the concreteness of different words on a scale ranging from 1 (abstract or language based) to 5 (concrete or experience based). BCI score for each pitch is computed based on words in the Brysbaert dictionary. We computed both the Brysbaert Concreteness Index and Brysbaert Variance Index for a particular pitch.

To measure the professionalization of pitch writing, we apply the Flesch-Kincaid readability tests,[33] which indicates how difficult a passage in English is to understand. The readability tests counts the number of difficult words, total words and syllables in a sentence:

$$206.835 - 1.015\left(\frac{\text{total words}}{\text{total sentences}}\right) - 84.6\left(\frac{\text{total syllables}}{\text{total words}}\right)$$

The score of Flesch reading ease ranges between 0 to 100. Higher scores indicate the reading material that is easier to read while lower numbers indicate passages are more difficult to read. We have applied the Gunning Fog Index[34] and the SMOG reliability[35] formula to validate our hypothesis.

We use business buzzwords to proxy for social innovation and institutional change. Specifically, we measure the adoption of business buzzwords in pitches through

---

[32] Brysbaert, Marc, Amy Beth Warriner, and Victor Kuperman. 2014. "Concreteness ratings for 40 thousand generally known English word lemmas." *Behavior Research Methods* 46: 904-911.

[33] Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., & Chissom, B. S. 1975. "Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel." Naval Technical Training Command Millington TN Research Branch.

[34] Londoner, Carroll A. 1967. "A readability analysis of randomly selected basic education and vocational education curriculum materials used at the Atterbury Job Corps Center as measured by the Gunning Fox Index." School of Education, Indiana University.

[35] Mclaughlin, G. Harry. 1969. "SMOG Grading - a New Readability Formula." *Journal of Reading* 12(8): 639-646.

word embedding method and cosine distance calculation. We project each pitch onto a pre-trained business model and calculate the distance of each document to the centroid that represents the buzzwords. We also validate our hypothesis by using discourse atoms by looking into specific words and its vector location on word embedding space.

Measuring Industry Boundaries: Startup Companies

Defining industry boundaries and competitiveness is critical for analyzing industrial organizations. Although traditional SIC or NAICS industry classification is convenient to set boundaries for firms and their products, it lacks time variation to reclassify firms over time as the product and strategy evolves. In terms of startup companies, misclassification of firms' industry categories is more severe. As most firms present potential concepts or ideas, the industrial classification might vary with the development of business plans. At the company's early stage, firms would be classified into multiple industries. Further, NAICS and SIC code fail to accommodate innovations that create entirely new product markets. For example, in the late 1990s, web-based firms and hundreds of technology companies were simply grouped into the classification "business service industry," which lacked meaningful description.[36]

Therefore, we propose a text-based method to re-classify industry categorization for startup companies and to measure its competitiveness across firms and within industries. The major method used is K-Means, the clustering algorithm that partition n observations into k clusters to minimize within-cluster variances based on

[36] Hoberg, Gerard and Gordon Phillips. 2016. "Text-Based Network Industries and Endogenous Product Differentiation." *Journal of Political Economy* 124 (5): 1423-1465. https://doi.org/10.1086/688176

word usage. What makes our method possible is that publicly traded firms would publish a 10-k file each year with SIC code. It enables us to build classification across firms and time, but also validate our results with existing SIC code.

$$f(\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{W}) = \sum_{i=1}^{N} \sum_{k=1}^{K} \sum_{j=1}^{J} \underbrace{\left(W_{ij} - \mu_{jk}\right)^2}_{\text{dissimilarity measure}} \overbrace{\pi_{ik}}^{\text{Cluster indicator}} .$$

Based on K-means clustering, our objective is to assign each pitch into one industry category. We then apply KNN algorithm to find the nearest industry category for each startup company. Consider a startup [i] with [k] competitors in a certain market location, and the similarity between their business statements could be a good indicator of the differentiation between their products. This allows us to investigate how unique a new startup's products or business models are in comparison with those of their competitors. The below figures show two different market structures, where a startup's differentiation from the closest competitors can be obtained by calculating the cosine distance between the focal node and its nearest nodes.

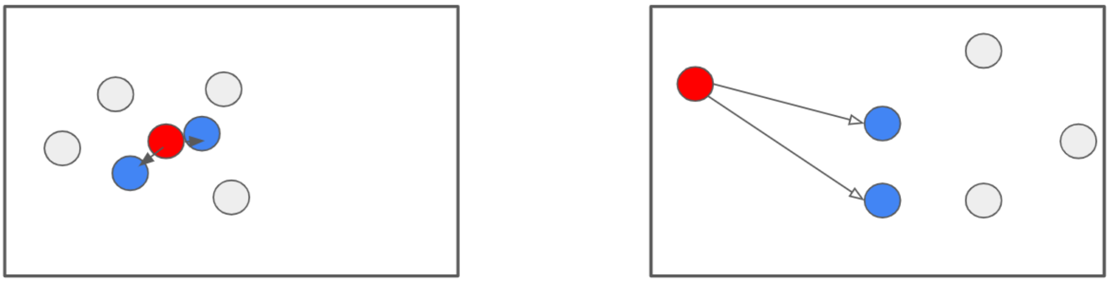

Figure 1. Market structures measured by the cosine distance between the company and its 2 nearest neighbors. (a) On the left, the focal company is in a more crowded and competitive market, so it is "fitting in." (b) On the right, the focal company is in a less crowded market and is "standing out."

After fine-tuning the model (the standard of selecting K is based on the Elbow method; that is, we select the optimal k for which the within-cluster-sum of squared errors starts to diminish), our analysis finds the optimal K equal to 2. We aggregate all distance across all competitors to get an empirical measure of the similarity between firms, then we take the average distance. The similarity score are estimated by the following formula:

$$\delta_{ti,tj} = cos(\theta) = \frac{A_{ti}A_{tj}}{||A_{ti}|| \times ||A_{tj}||} \in [0,1]$$

$$S_i = \frac{1}{k}\sum \delta_{ij}, k = \{\text{2 closest competitors}\}$$

4. RESULTS AND DISCUSSION

Linguistic Features of Pitches Overtime

To investigate the drivers and effects of pivots, we first explore the text corpus by seeking patterns in business pitches and how they change overtime. We found that pitches become more concrete, sophisticated, and specific across the board.

First, using the Brysbaert concreteness index, we compare the concreteness of the first and last pitch with a t-test and found significant difference in concreteness between the two. The Brysbaert concreteness index relies on abstraction norms that were rated by MTurkers for 40,000 commonly used English lemma words.[37]. Here, we interpret the increase in Brysbaert concreteness index from the first to last pitch as the company becoming notably clearer about their product, customers, and market segmentation.

[37] Center for Reading Research. 2022. "Concreteness ratings for 40 thousand English lemmas." Department of Experimental Psychology of Ghent University. http://crr.ugent.be/archives/1330

Second, employing the Flesch reading ease score, we similarly compare the sophistication of the syntax in the first and last pitch by the company. With a t-test, we uncovered a significant difference between the sophistication of the first and the last pitch. The increase in sophistication implies that companies pitch their business more professionally, catering to the professional linguistic register favored by investors.

Lastly, computing the distance by Euclidean norm (n=50) of the first and last pitches from the centroid of commonly used words, we observe that the last pitch tends to include buzzwords in the centroid rather than the first pitch. However, the pitches are still spread out around the centroid because the companies use syntax that is specific to their own products.

Test	Mean (first pitch)	Mean (last pitch)	Std.Dev (first pitch)	Std.Dev (last pitch)	Mean_diff	t
Brysbaert Concreteness Index	2.421113	2.434159	.1766513	.2566383	-.0130458	-4.7473***
Flesch Reading Ease Score	28.08923	23.73196	20.91201	23.48264	4.357273	24.5027***
Convergence	4.941477	4.804689	1.95869	1.971871	.136788	4.5052 ***

Figure 2. Changes in the linguistic features, from comparison between the first pitch and the most recent pitch on Crunchbase. The table shows results of t-test when we measure (a) language concreteness using Brysbaert concreteness index; (b) language sophistication and professionalism using Flesch reading ease score; and (c) convergence of the language used by computing the Euclidean distance (n=50) from both pitches to the centroid (0, 0) of commonly used words, which are validated by

discourse atoms.

Effects of Linguistic Features Change

Comparing the last pitch with the first, we find that companies' pitches become increasingly concrete, professional, and specific overtime. We further analyzed the effect of such linguistic feature changes on companies' funding success. We found that increase in concreteness, decrease in readability, and increase in convergence do not have significant effects on funding raised. It implies a general evolution of firms' pitches. This could be because investors see through the syntactic changes of company pitches and focus more on the actual business strategy of the company. However, when implementing the regression model, we found the linguistic features are not a significant, powerful predictor of future returns or market success of a company, compared to changing the words used in a pitch. By taking the interaction terms of linguistic features and pitch changes, the results show that the pitch changes and linguistic features do not have a significant impact on future money raising (from . One possible explanation could be that becoming professional, concreteness and adoption of innovative languages is a general trend across business lifespan, for all companies. In this next section, we would like to explore how the content or strategic changes would impact the startup companies' market success.

VARIABLES	(1) Log(Money Raised)	(2) Log(Money Raised)	(3) Log(Money Raised)
Pitch change	5.694***	6.056***	6.041***
	(1.621)	(1.587)	(1.580)
Concreteness	-0.839		
	(0.644)		
Convergence		0.140	
		(0.103)	
Readability			0.0244
			(0.0225)
Constant	16.05***	13.72***	13.64***
	(1.564)	(0.247)	(0.359)
Observations	2,959	2,931	2,959
R-squared	0.006	0.006	0.006
Number of company			

Figure 4 a. Correlation between pitch changes and linguistic features. We conducted a linear regression of linguistic features and log scale of money raised – with n>2900. We do not observe significant correlation between increase in linguistic concreteness (-0.839, sd=0.644), sophistication (0.0244, sd=0.0225), and convergence (0.140, sd=0.103) and the log scale of funding raised.

VARIABLES	(1) Log(Money Raised)	(2) Log(Money Raised)	(3) Log(Money Raised)
Pitch Change	9.293	10.42*	10.59*
	(5.976)	(6.093)	(6.111)
Pitch Change#Readability	-0.0307**		-0.0112
	(0.0147)		(0.0242)
Pitch Change#Concreteness		-0.591**	-0.435
		(0.260)	(0.428)
Constant	14.59***	14.74***	14.72***
	(0.207)	(0.245)	(0.247)
Observations	2,009	2,009	2,009
R-squared	0.013	0.015	0.016
Number of company	1,622	1,622	1,622

Figure 4 b. Correlation between pitch changes and linguistic features. We conducted a linear regression of linguistic features with the interaction of pitch change and log scale of money raised. We do not observe significant correlations between the interaction of pitch change and readability and the interaction of pitch change and concreteness in Model 3.

Structural Drivers of Pivots

To measure the drivers of pitch changes across firms from 2016 to 2020, we use the fixed-effected linear regression. The dependent variable is a dummy variable of pitch changes; and the independent variables are those hypothesized factors that influence a firm changes its self-business description, such as amount of funds received in the past (given the skewed distribution of funds, we take the log transformation of this variable), number of employees (team size), geographic locations (the categorical variable is based on Crunchbase's geographic location classifications of Western US, Midwestern US and East Coast), firm age (how many years from a firm's established year to 2021). Since the nature of panel data in linear regression, we assume the individual-specific effect are correlated with the independent variables.

From the linear regression of a few hypothesized variables that drive business pivots for early-stage companies, we find that companies that have (1) more past funding, (2) larger team size, (3) headquarter located in the West Coast, and (4) younger company age are more likely to pivot.

While past funding success and team size positively correlated with likelihood of pivots, firm age was negatively correlated with the log odds of pivots. Specifically, companies with more past funds are more likely to pivot in the future. Moreover, companies with bigger teams are also more likely to pivot, possibly because there are more pivot ideas, since larger teams facilitate more external learning from experts and alliances.[38] This could indicate a matthew-effect on the winner-takes-all market, where

[38] Almeida, Paul, Gina Dohko, and Lori Rosenkopf. 2003. "Startup size and the mechanisms of external

firms with more aggregate resources or more brokerage (a more central position in social networks) could receive more funds and support from venture capitalists.

In contrast, companies that have been in existence longer are less likely to pivot, since the company will have already found their niche within the industry. The results remained consistent when we conducted the log regression across four different models.

Surprisingly, location seems to be a reliable predictor of likelihood of future pivots. Here, companies located on the West Coast and East Coast had a higher location MLE estimate than companies in the Midwest. This could be because the early-stage entrepreneurship networks are stronger on the two coasts than Middle America.

VARIABLES	(1) Pivot_dummy	(2) Pivot_dummy	(3) Pivot_dummy	(4) Pivot_dummy
Log(Past funds)	0.0383*** (0.00343)	0.0383*** (0.00343)	0.0343*** (0.00357)	0.0426*** (0.00374)
Teamsize			8.57e-05*** (1.49e-05)	0.000129*** (1.59e-05)
1. Western US			0.472*** (0.0521)	0.436*** (0.0523)
2.Midwestern US			0.0513 (0.0798)	0.0304 (0.0801)
3.East Coast			0.284*** (0.0582)	0.253*** (0.0585)
Firm age				-0.0315*** (0.00310)
Constant	-1.033*** (0.0495)	-1.033*** (0.0495)	-1.336*** (0.0661)	-0.935*** (0.0769)
Observations	14,445	14,445	13,463	13,463

Figure 5. Logistic regression model of the MLE estimates of coefficient on (a) log scale of past funding, (b) teamsize, (c) geographic location of company headquarters, West Coast / Midwest / East Coast, (d) firm age, and (e) estimated constant term on the log odds of a company having a pivot in its business description. We ran this logit model across four different models of the pivot dummy, and we received results that seem to be generally consistent with each other.

learning: increasing opportunity and decreasing ability?" *Research Policy* 32(2): 301-315. https://www.sciencedirect.com/science/article/pii/S0048733302001014

Another analysis using fixed-effect regression validates the finding that companies that have more prior funding and younger company age are more likely to pivot. Moreover, this analysis reveals that gender can only loosely predict likelihood of pivot. Though there is no significant gender difference in company pivots, the gender difference exists in many dimensions of entrepreneurship, such as funding amount and fundraising success,[39] as well as rate of participation.[40]

Moreover, retrocasting the company funding data and pivot reveals that business pivots has a positive correlation with future funding success. When we run the text model over money raised in the upcoming years, we find that the log of money raised in each of the next four years is positively correlated with pivots. Therefore, companies that have pivoted are more likely to acquire a next round of financing. According to Y Combinator – the leading early-stage accelerator that early-invested successful companies like Airbnb, Reddit, and Coinbase – pivoting is recommended when the opportunity cost of pursuing the current business plan is higher than the benefit from the current venture.[41] In this way, pivoting allows the startup team to capture more value, and investors respond to this value increase by providing more

---

[39] Kanze, Dana, Laura Huang, Mark A. conley, and E. Tory Higgins. 2018. "We Ask Men to Win and Women Not to Lose: Closing the Gender Gap in Startup Funding." *Academy of Management Journal* 61(2). https://journals.aom.org/doi/abs/10.5465/amj.2016.1215

[40] Nigam, Nirijhar, Cristiane Benetti, and Hareesh Mavoori. 2022. "Access to venture capital: Does gender (still) really matter?" *Strategic Change* 31(2): 239-248. https://onlinelibrary.wiley.com/doi/full/10.1002/jsc.2493

[41] Caldwell, Dalton. 2022. "All About Pivoting." Y Combinator Startup School. https://www.ycombinator.com/library/6p-all-about-pivoting

capital.

VARIABLES	(1) Log(Money Raised Next Year)	(2) Log(Money Raised Next Year_2)	(3) Log(Money Raised Next Year_3)	(4) Log(Money Raised Next Year_4)
Log(Money Raised)	0.234***			
	(0.0197)			
Company age	0.000134*	0.000305**	1.17e-05	0.000206
	(7.03e-05)	(0.000124)	(0.000155)	(0.000194)
1.Founder gender	-0.700*	-0.464	-0.445	-4.144
	(0.395)	(0.703)	(0.813)	(2.905)
2.Founder gender	-0.213	-0.599	-0.0358	-0.0243
	(0.246)	(0.449)	(0.595)	(0.974)
Pitch change	194.6***	179.2***	152.4**	156.2*
	(31.67)	(51.74)	(71.43)	(79.84)
Log(Money Raised Next Year)		0.332***		
		(0.0373)		
Log(Money Raised Next Year_2)			0.244***	
			(0.0492)	
Log(Money Raised Next Year_3)				0.240***
				(0.0782)
Constant	10.28***	8.600***	10.93***	11.29***
	(0.335)	(0.665)	(0.890)	(1.388)
Observations	3,541	1,212	661	222
R-squared	0.074	0.127	0.060	0.093

Figure 6. Regression Model of funding success and our hypothesized predictors of pivots.
We conducted a multiple linear regression model of the OLS estimates of coefficient on (a) log scale of past funding, (b) company age, (c) binary gender of the founder, female or male, (d) pitch changes, and (e) estimated constant term on the log odds of a company acquiring additional funding in the future years.

Standing Out v.s. Fitting In

Measuring Industry

Innovation is embedded in social games, and understanding the social structure of startup firms is critical for our analysis on funding strategies. Upon first try, I encountered a challenge with classifying the industry of early-stage companies based on their self-descriptions. We first used the self-classification on Crunchbase of each company, where companies tend to put multiple industry labels instead of providing a primary industry focus, likely for visibility purposes. The Crunchbase classifications totaled over 500 different industries, and this resulted in a messy graph with low legibility.

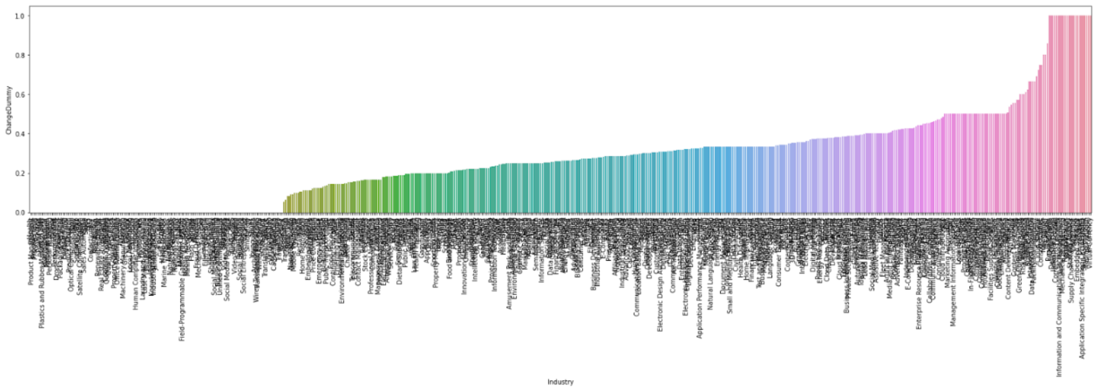

Figure 4. Messy Distribution of Startups Changed Self-description by Industry (classification based on Crunchbase.com)

I then tried to use the official SIC/NAIC codes as the pre-trained model, which is the standard measure of industry. However, these codes turned out to be restrictive and failed to capture innovative products. Shown in the PCA visualization in the figure below, the description of early-stage companies lies at the intersections of multiple industries. The many industry overlaps suggest that companies do not fit well with the SIC/NAIC code industry classification.

Industry Classification of Public Companies (K-means)

Continuing the search for a viable industry classification, I tested a word-embedding model that is pre-trained on 2010-2020 business news. Given that business news tends to report on more mature companies, we tested this model using the business description on 10-K's of public companies. The PCA visualization of this K-means clustering in the below figure shows that the business news model worked well to classify different industries of companies, where companies in the same industry clustered together with minimal overlap.

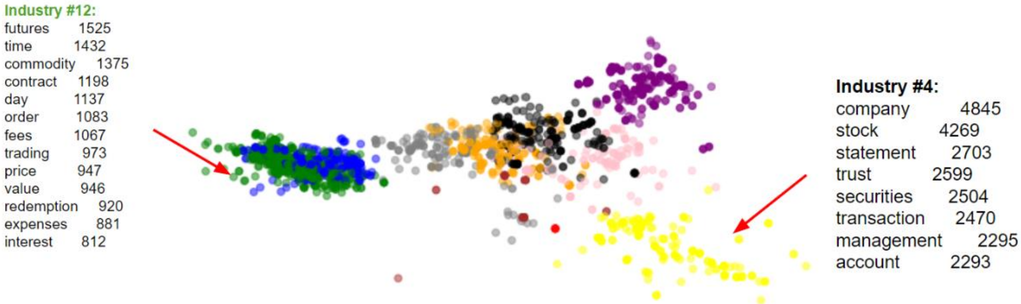

Figure 5. PCA Visualization of K-Means Industry classification.

Industry Classification of Startups (KNN)

Based on the K-Means clustering for mature 10-k companies, the major objective is to find industry classification for startup companies by KNN classification. The Figure below shows the distribution of companies in each industry categories. The classification is imbalance: total number of firms in industry #8 are 6,3476 while in industry #36 is only 36 companies. It is not surprising because industry #8 represents the software industry with key words –"services", "business", "sales", "solutions", "technology", "information", "product", "software", "systems", and industry #36 includes keywords "bank", "income", "services", "management", "operations", because most of startups are software based.

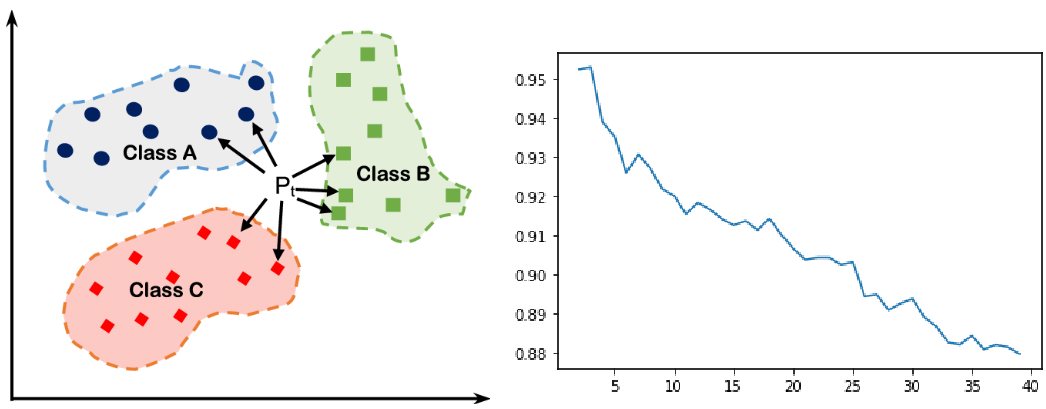

Figure 6. (a) On the left, the graph shows KNN classification for each startup company, classifying each pitch to one industry category. (b) On the right, the

plot shows accuracy score of KNN classification with different k from 2 to 40 (the optimal K = 3).

How does standing out or fitting in strategy relate to firms' performance and opportunities? My measure of strategy is similarity score – defined as the average distance from focal firms to three nearest firms –is highly skewed, indicating that most focal startups are similar to their competitors. The competitors consist of both public listed firms and startups clustered within a same industry. The lower similarity score means the distance between focal firms and competitors is larger, representing the standing out strategy. The number of nearest firms was decided by the highest accuracy score.

The regression results and Figure below shows a positive correlation between similarity score and financing outcomes, and similarity the correlation between pitch changes and fund raising is positive. A higher similarity score predicts a higher amount of financing raised with a coefficient of 5.264. Therefore, integration strategy could be effective in early-stage startups.

VARIABLES	(1) Funds	(2) Funds	(3) Funds
Pitch change	8.531***		8.606***
	(0.681)		(0.681)
Similarity score	5.264***	4.181***	5.458***
	(0.431)	(0.414)	(0.439)
Number of competitors			-3.90e-06**
			(1.76e-06)
Constant	15.12***	15.15***	15.25***
	(0.0585)	(0.0589)	(0.0798)
Observations	35,330	35,330	35,330
R-squared	0.012	0.005	0.012

Figure 7. Correlation between funding success and our hypothesized predictors of pivots and similarity score.

Next, I extend the study to which the differentiation strategy would be changed over business's life cycle and equity exit events such as IPO. I find that overtime the firms fluctuate between "standing-out" and "fitting-in" strategies. Yet, overtime firms become similar to each other. After IPO firms become more conservative and less likely to differentiate from market competitions.

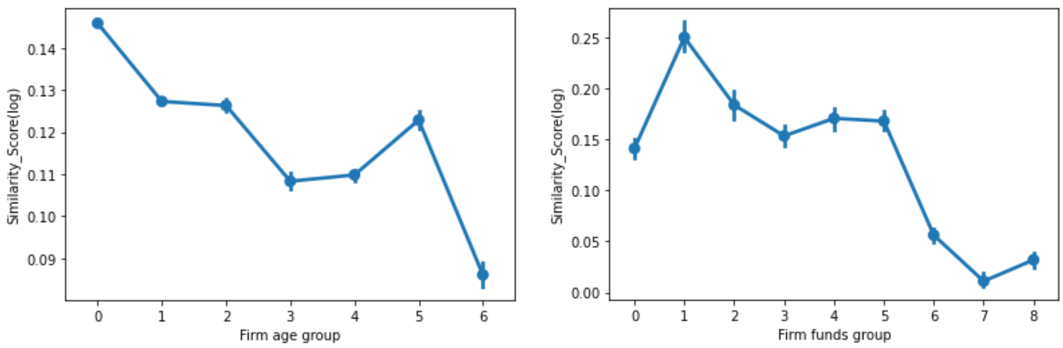

Figure 8. (a) On the left, a graph of firm lifespan and similarity score. (b) On the right, a graph of firm funding group (grouped by 10 percentile) and similarity score.

VARIABLES	(1) Similarity Score	(2) Similarity Score
did	0.00355***	0.00321***
	(0.000650)	(0.000652)
Funds		-4.6e-13 ***
		(0)
Constant	0.0969***	0.0975***
	(0.000634)	(0.000638)
Observations	298,700	298,700
R-squared	0.000	0.000

Figure 9. A table showing two difference-in-difference models that estimate how differentiation strategy would be changed before and after IPO.

5. CONCLUSION

In summary, based on computational methods, I measure and investigate how does business strategies in early-stage firms influence future success. Digital text methods enable scholars in organizational theory to measure the dynamics of conceptual spaces—innovation diffusion and process of institutionalization within markets. I find that a general trend for companies to change the way they describe their business and products overtime. To hone in on the idea of a substantive pivot from changes in linguistic features, I also find that companies that are younger and larger with more past funding are more likely to pivot. Interestingly, there seems to be a positive feedback loop between pivoting and funding: companies with more funding from the past are more likely to pivot, and companies that pivot tend to acquire more money raised in the future years. From the regression analysis in Figure 5, it shows the positive effect of past funding on pitch changes; and from the regression analysis in Figure 6, the effect of pitch changes positively predicts the future funds supports that if a firm received more funds in the past, then it will more likely to change its pitch; after the pitch has been updated, the firm is more likely to raise its funds.

We find that a general trend for companies to change the way they describe their business and products overtime. In terms of linguistic features, company pitches become increasingly concrete, professional, and specific, though they also incorporate buzzwords in business. However, the linguistic features do not have significant effects on firms' fund raising.

For the central question of fitting in or standing out, we find that public companies are less likely to pivot, while younger startup companies are more likely to "stand out" and shift the industry paradigm. With the rich text corpus on business self-descriptions for both early-stage startups and public companies, scholars can use text analysis methods to further investigate the nuances in business pitch changes and pivots, their patterns, drivers, and consequences.

Specifically, there are several future directions for this project: as the business strategy only measured by differentiation and similarity, substantive changes in contents have not been examined. From reading the pitches manually, I found overtime pitches are more likely to mention investments, investors and manager teams, instead of presenting a board and abstract business idea. The effect between content changes is unclear. Finally, the current results indicate the process of institutionalization—that is, early-stage companies will benefit from imitating the public listed companies. If firms become more similar over time, one question could be where does the innovation comes from? How does legitimization of innovation happen in marketplace? And what's the effect of innovation on firms' growth?

6. FUTURE RESEARCH

There are a few limitations to this project, where future research can elaborate on and further the investigation into the question of standing out or fitting in.

First, the time series text corpus of startup business descriptions was scraped from the Wayback Machine. A drawback of using the Wayback Machine is that we have more data points on companies that are more recognized and visible because the Wayback Machine scrapes more frequently sites that are more popular. To further study the changes in text description of early-stage companies, future research can scrape a set growing list of early-stage companies within an industry of interest regularly to capture business descriptions on the same time intervals.

Second, the business description text data from 10-K's of public companies and from Crunchbase of startups have syntax and semantic differences. Future research can tap into this well of difference and employ causal inference models to explore whether public companies releasing descriptions causes changes in startup descriptions.

Third, the public companies text data used in this paper is a randomly selected sample of 25% Edgar 10-K's. The R package employed to scrape and parse the business descriptions from 10-K's took up much memory and time to run, while the Python code we used to parse the business description ourselves did not yield very accurate results. We decided to use a sample of accurately parsed data, but future research should seek to use all of the 10-K data for more robust results.

7. BIBLIOGRAPHY

Organizational Theory

Askin, N., & Mauskapf, M. (2017). What makes popular culture popular? Product features and optimal differentiation in music. *American Sociological Review*, *82*(5), 910-944.

Bian, J., Greenberg, J., Li, J., & Wang, Y. (2022). Good to go first? Position effects in expert evaluation of early-stage ventures. *Management Science*, *68*(1), 300-315.

Caldwell, Dalton. 2022. "All About Pivoting." Y Combinator Startup School. https://www.ycombinator.com/library/6p-all-about-pivoting

Coser, L. A., Kadushin, C., & Powell, W. W. (1982). *Books: The culture and commerce of publishing* (p. 22). New York: Basic Books.

DiMaggio, Paul J., and Walter W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 48 (2): 147–60.

Eisenhardt, Kathleen M., and Jeffrey A. Martin. "Dynamic capabilities: what are they?." *Strategic management journal* 21, no. 10-11 (2000): 1105-1121.

Goldberg, A., Srivastava, S. B., Manian, V. G., Monroe, W., & Potts, C. (2016). Fitting in or standing out? The tradeoffs of structural and cultural embeddedness. *American Sociological Review*, *81*(6), 1190-1222.

Granqvist, N., Grodal, S., & Woolley, J. L. (2013). Hedging your bets: Explaining executives' market labeling strategies in nanotechnology. *Organization science*, *24*(2), 395-413.

Hoberg, Gerard and Gordon Phillips. 2016. "Text-Based Network Industries and Endogenous Product Differentiation." *Journal of Political Economy* 124 (5): 1423-1465.

Huang, L., & Pearce, J. L. (2015). Managing the unknowable: The effectiveness of early-stage investor gut feel in entrepreneurial investment decisions. *Administrative Science Quarterly*, *60*(4), 634-670.

Rosa, J. A., Porac, J. F., Runser-Spanjol, J., & Saxon, M. S. (1999). Sociocognitive dynamics in a product market. Journal of marketing, 63(4_suppl1), 64-77.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic management journal*, *18*(7), 509-533.

Starr, P. (1980). Medical care and the boundaries of capitalist organization. *Unpublished manuscript. Program on Non-Profit Organizations, Yale* University, New Haven, CT.

Porter, M. E. (2011). *Competitive advantage of nations: creating and sustaining superior performance*.

Priem, R. L., & Butler, J. E. (2001). Is the resource-based “view” a useful perspective for strategic management research?. *Academy of management review*, *26*(1), 22-40.

Navis, C., & Glynn, M. A. (2010). How new market categories emerge: Temporal dynamics of legitimacy, identity, and entrepreneurship in satellite radio, 1990–2005. Administrative Science Quarterly, 55(3), 439-471.

Navis, C., & Glynn, M. A. (2010). How new market categories emerge: Temporal dynamics of legitimacy, identity, and entrepreneurship in satellite radio, 1990–2005. Administrative Science Quarterly, 55(3), 439-471.

Nigam, Nirijhar, Cristiane Benetti, and Hareesh Mavoori. 2022. “Access to venture capital: Does gender (still) really matter?” *Strategic Change* 31(2): 239-248.

Kleinbaum, A. M. (2012). Organizational misfits and the origins of brokerage in intrafirm networks. *Administrative Science Quarterly*, *57*(3), 407-452.

Weber, Max 1952 The Protestant Ethic and the Spirit of Capitalism. New York: Scribner. 1968 Economy and Society: An Outline of In- terpretive Sociology. Three volumes. New York: Bedminster.

Williamson, O. E. (1999). Strategy research: governance and competence perspectives. *Strategic management journal*, *20*(12), 1087-1108.

Wu, A. (2016). Organizational decision-making and information: Angel investments by venture capital partners. In *Academy of Management Proceedings* (Vol. 2016, No. 1, p. 11043). Briarcliff Manor, NY 10510: Academy of Management.

Zuckerman, Ezra W. 1999. “The Categorical Imperative: Securities Analysts and the Illegitimacy Discount.” *American Journal of Sociology* 104 (5): 1398–1438

Methodology, Startups, Gender

Almeida, Paul, Gina Dohko, and Lori Rosenkopf. 2003. “Startup size and the mechanisms of external learning: increasing opportunity and decreasing ability?” *Research Policy* 32(2): 301-315

Brysbaert, Marc, Amy Beth Warriner, and Victor Kuperman. 2014. “Concreteness ratings for 40 thousand generally known English word lemmas.” *Behavior Research Methods* 46: 904-911

Caldwell, Dalton. 2022. “All About Pivoting.” Y Combinator Startup School. https://www.ycombinator.com/library/6p-all-about-pivoting

Center for Reading Research. 2022. “Concreteness ratings for 40 thousand English lemmas.” Department of Experimental Psychology of Ghent University

Dalle, J., M. den Besten and C. Menon. 2017. "Using Crunchbase for economic and managerial research", OECD Science, Technology and Industry Working Papers, No. 2017/08, OECD Publishing, Paris

Gender API. 2022. https://gender-api.com

Ghanim Al-Ali, Ahmed, Robert Phaal, and Donald Sull. 2020. “Deep Learning Framework for Measuring the Digital Strategy of Companies from Earnings Calls.” Proceedings from the 28th International Conference on Computational Linguistics, pages 927-935. https://aclanthology.org/2020.coling-main.80.pdf

Guzman, Jorge, and Aishen Li. 2019. “Measuring Founding Strategy.” SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3489585

Huang, L., Joshi, P., Wakslak, C., & Wu, A. (2021). Sizing up entrepreneurial potential: Gender differences in communication and investor perceptions of long-term growth and scalability. Academy of Management Journal, 64(3), 716-740.

Kanze, Dana, Laura Huang, Mark A. conley, and E. Tory Higgins. 2018. “We Ask Men to Win and Women Not to Lose: Closing the Gender Gap in Startup Funding.” *Academy of Management Journal* 61(2).

Khaire, M., & Wadhwani, R. D. (2010). Changing landscapes: The construction of meaning and value in a new market category—Modern Indian art. *Academy of Management Journal*, *53*(6), 1281-1304.

Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., & Chissom, B. S. 1975. “Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel.” Naval Technical Training Command Millington TN Research Branch.

Londoner, Carroll A. 1967. “A readability analysis of randomly selected basic education and vocational education curriculum materials used at the Atterbury Job Corps Center as measured by the Gunning Fox Index.” School of Education, Indiana University

Lin, Yu-Kai, Likoebe M. Maruping. 2021. “Open Source Collaboration in Digital Entrepreneurship.” Organizational Science 33(1): 212-230.

Mclaughlin, G. Harry. 1969. “SMOG Grading - a New Readability Formula.” *Journal of Reading* 12(8): 639-646

Nigam, Nirijhar, Cristiane Benetti, and Hareesh Mavoori. 2022. “Access to venture capital: Does gender (still) really matter?” Strategic Change 31(2): 239-248

Office of Chief Actuary. 2022. “Top 10 Baby Names of 2020.” Social Security Administration. https://www.ssa.gov/oact/babynames/

Vicinanza, Paul, Amir Goldberg, and Sameer B. Srivastava. 2020. “Quantifying vision through language demonstrates that visionary ideas come from the periphery.” SocArXi.