THE UNIVERSITY OF CHICAGO

Higher-Order Innovation and New Venture Success

By

Ziwen Chen

July 2021

A paper submitted in partial fulfillment of the requirements for the Master of Arts degree in the Master of Arts in Computational Social Science

Faculty Advisor: James Evans

Preceptor: Sanja Miklin

Higher-Order Innovation and New Venture Success

Abstract

Entrepreneurship literature points out that a balance between exploration and exploitation leads to a good innovative performance of start-ups, yet how this balance can be achieved remains unknown. This study argues that the modularity principle, adopted from the complexity perspective, is the key to successful innovative strategies. By applying dynamic word embedding techniques to millions of documents from 119 business newspapers, magazines and patents, we construct a dynamic landscape of business discourse across 45 years. By locating new venture descriptions in this dynamic business landscape, we can observe how business elements are recombined within a given new venture and how innovative it was in its own time and historical context. Two different categories of recombination exist for each company: the first-order recombination that combines technical elements to approximate application elements to form basic functional blocks, and second-order recombination that recombines these blocks to satisfy more complex demands. Based on these measurements, we model the start-ups' birth, growth, and death with event history analysis. Our analysis reveals that radical second-order recombination is key to entrepreneurial success in the U.S.: the first-order process calls for exploitation while the second-order process requires exploration. Both exploitation and exploration matter for the new ventures, yet exploitation and exploration co-exist in innovative strategies through a specific architecture.

Keywords: entrepreneurship, novelty, recombination, exploration and exploitation, machine learning, high-dimensional embeddings

Introduction

How do new ventures come up with a competitive strategy? Traditional explanations believe a successful strategy is the product of exploration and exploitation, and a wise balance between the two (Duncan, 1976). Exploration broadly refers to practices such as "search, variation, risk-taking, experimentation, discovery," while exploitation is captured by "refinement, choice, efficiency, implementation" (March, 1991). Led by these principles, firms can efficiently recombine relevant elements from local or distant domains, and reap the profits from full utilization of the resources and novelties (Kaplan and Vakili, 2015).

Exploration and exploitation occur at all levels of the economy, yet either one has a price: complete exploitation closes future possibilities, while complete exploration brings too many uncertainties (March, 1991). Therefore, wise management calls for "ambidexterity" (Duncan, 1976; Raisch and Birkinshaw, 2008), the proper balance of the exploration and exploitation that serves long-term benign development of the ventures.

Reasonable as it is, the mechanism of this co-existence still remains a mystery. Entrepreneurship involves complex activities, from producing new technologies, adopting old technologies to new problems, adopting new technologies to serve old problems, to aligning previously distant technologies to generate something totally unimaginable from the old perspective. It is obvious that each move includes some elements of exploration and exploitation, and both of them are essential in value generation. However, it is still unclear how to best distribute resources between the two. The traditional methodological culture, which heavily relies on linear predictions and i.i.d assumptions, has limited our theoretical thinking.

In this paper, we turn to the concept of modularity to think about the mechanism of ambidexterity in new venture strategies. Modularity, proposed by Simon (1962) in *The*

Architecture of Complexity, refers to the form of hierarchy in complex systems—a system should be "composed of interrelated subsystems, each of the latter being, in turn, hierarchic in structure until we reach some lowest level of elementary of the subsystem"—and this should be a common property among diverse forms of complexities, including biological systems, physical world and human society. For example, Adam Smith had argued that division of labor facilitates different functional units that can later be recombined to form broadly observed complexity and variety (Smith, 1776); and recent studies in innovative emergence (Padgett and Powell, 2012) also echoed this thinking.

From this perspective, we try to argue that entrepreneurship, as a unique form of social activity, requires a special form of ambidexterity. It demands the entrepreneurs to focus on exploitation for the first-order inventions—the from-scratch creation of new techno-social blocks that provide a mature formula to solve social problems with technical elements, and exploration for second-order assembly—the recombination of current existing blocks that have existed in the landscape. Business ventures are higher-order social creations that are based on established techno-social clusters: entrepreneurship calls for the exploratory combination of stable, value-creating technologies, not innovation all the way down. Contrary to previous structural approaches and contextual approaches in ambidexterity (Raisch and Birkinshaw, 2008; Alizadeh and Jetter, 2018), we believe that exploration and exploitation can co-exist for the same organization in the same context.

Based on our longitudinal corpus of 119 English newspapers and magazines in the ProQuest database, published in the U.S. and falling into the category of "Business and Economics," we build a dynamic business landscape that allows us to identify established techno-social blocks and measure different levels of recombination for over 200,000 U.S. start-ups during the past 45 years.

This space traces the evolution of the market through history and allows us to measure the modular structure in innovative strategies of enterprises at the time of their genesis. Our empirical analysis shows that second-order recombination is critical for the long-term development of new ventures, while too much attention on the first-order inventions could deprive the vigor of the start-ups.

Literature Review

Chaos of Exploration and Exploitation

In his landmark article *Exploration and Exploitation in Organizational Learning*, March (1991) proposed two fundamental learning activities between which firms divide their attention and resources: exploration and exploitation. Exploration refers to the outward search that brings new elements to the organization, while exploitation aims to take full advantage of existing resources and improve efficiency. Exploitation helps firms better adapt to the current environment and enhance short-term performance, while exploration prepares the firm for possible future change in the environment. Although early studies claimed it is impossible to address efficient exploration and exploitation simultaneously, and there must be a trade-off due to resource limitations (Miller and Friesen, 1986; McGill et al., 1992), March's (1991) argument that successful firms tend to be ambidextrous leads to the more comprehensive thinking in later research. In recent studies, scholars have increasingly adopted the notion of balance between the two principles.

Insightful as it is, the concept "ambidexterity" provides more of a perspective than a mechanism: it explains why, but not how, exploration and exploitation co-exist in the organizations. Although both exploration and exploitation are valuable strategies that bring benefits, they are competitors for resources in practice. The practical contradiction and theoretical compatibility have brought in some theoretical tensions.

One of the commonly used solutions for this problem is to introduce business contexts in ambidexterity studies. Alizadeh and Jetter (2018) categorize these contexts as structural ambidexterity and contextual ambidexterity, and sometimes leadership-based ambidexterity is also mentioned (Raisch and Birkinshaw, 2008). Structural ambidexterity refers to the organizational arrangement in which organizational structures are carefully designed to cope with the competing demands of alignment and adaptability. One instance is differentiation, where some units in the organization work on exploration while other units are devoted to exploitation (O’Reilly and Tushman, 2008). The contextual approach realizes ambidexterity by creating culture, or a specific scheme, that encourages individuals to judge by themselves when and how to divide their time and attention between exploration and exploitation (Gibson and Birkinshaw, 2004). The leadership approach emphasizes the role of the leading team in an organization, hoping leaders can efficiently manage the contradictions between exploration and exploitation (Tushman and O’Reilly, 1997). Other contexts include time (companies focus on exploitation for one stage, exploration for another stage; see Rippa et al., 2019), or space (to keep exploitation within the organization and realize exploration by alliance or spinning-out a unit; see Christensen and Overdorf, 2000).

Even though these studies provide inspiring conclusions, the coexistence of exploration and exploitation is like the coexistence of the sun and moon: they cannot exist at the same time. For example, if a worker decides to explore, he must stop his exploitation work first and switch to an exploration task. Thus, the problem of dualism always lingers, and work can either be exploratory or exploitative, but not both.

The dualism problem arises from our methodological culture of linear thinking. According to linear thinking, exploration and exploitation follow an inverse, straight-line relationship—one goes down and the other must go up. Following this way of thinking, we are unable to see beyond the single, linear dimension and construct nonlinear, multi-dimensional theories. Furthermore, linear thinking is related to the traditional i.i.d assumption, a statistical prerequisite for nearly every quantitative model. According to the i.i.d assumption, individuals and variables are independent and identically distributed, limiting the possibility to measure inter-unit interactions. However, the notion of exploration and exploitation do imply inter-connections between social actors. For example, to explore is to bring in new elements that are created by others. Due to the restriction of linear thinking and i.i.d assumptions, existing research often fails to capture the interconnectedness between exploitation and exploration.

This paper makes a preliminary attempt to break the linear methodology and bring a high-dimension structure into organization learning theories. We propose an "architecture of entrepreneurship search", which is built on Simon's (1962) concept of modularity from Architecture of Complexity. By modeling the business landscape as a complex system with specific structures, we argue that the co-existence of exploration and exploitation goes beyond heterochronicity: they happen at the same time in the architecture of innovative strategies. The perspective of complex system, and recent development of natural language processing, allow us to carry out empirical testing.

"Architecture of Entrepreneurship Search"

A complex system refers to a system with interacting components, nested relationships, and changing collective behaviors. From this perspective, almost all human societies are complex systems. Even though conventional, neoclassical economics assumes static equilibrium and rational agents, the real-world economic system is one of the most chaotic and complex systems, with new actions, beliefs, and strategies emerging every day (Arthur, 2021). The non-equilibrium feature of the economy has attracted much attention and discussion. Kauffman (2002) believed "the co-constructing behaviors of autonomous agents will spill over to the economy, with surprising implications for the foundations of economies, for economic growth, and for the development of adaptive firms that co-evolve in the corporate ecosystems." Arthur (2009) expressed the same principle in his work about technological evolution: "the economy is not a simple system; it is an evolving, complex one, and the structures it forms change constantly over time."

One of the most fundamental properties of complex systems is emergence, the phenomenon that new behavior patterns emerge from the interactions of individual components. This makes it naturally suitable for exploring the economic innovation, the emergence of novel technology and capacity that facilitate economic growth. In fact, complexity theories have indeed become increasingly influential in innovation studies. Studies shows that the innovation process is accompanied by different dimensions of complexities (Garud et al., 2013) that happen on different levels of the economy (Poutanen et al., 2015). In their masterpiece, *The Emergence of Organizations and Markets*, Padgett and Powell (2012) develop a theory about the emergence of market and organization novelties: novelty comes from spillovers across intertwined networks in different domains which evolve ceaselessly.

Despite previous efforts in studying innovation from complexity perspective, the architecture of complex innovation remains a mystery. For this puzzle, Simon's (1962) concept of modularity, a fundamental property of complex system, provides a possible entry point. Modularity refers to the structure where elements are combined into lower-level subsystems in the first place, and high-level components are established on the existing blocks; higher-level architectures again are built on these sub-components, and this process can repeat until a complicated unit is created. This hierarchical structure reveals particularly interesting inferences about the innovation dynamics. For example, innovations are often top-down—when the old blocks reach their limitations and cannot stretch to adapt to the environment, they are replaced by new ones, and some of their unique sub-components are abandoned as well, a chain of collapses that is named "avalanches of destruction" by Arthur (2009). Based on this perspective, many important modern innovations are, in fact, higher-order recombination.

Modularity serves the evolution of the system in multiple ways. First, combinations in the system continue to generate new blocks for future use, thus increasing the diversity of elements in the system. When new needs appear, these "spare parts" can be conveniently utilized to produce potential solutions, and the adaptability of the whole system will be promoted hereby (Hidalgo and Hausmann, 2009). Second, the emergence of useful components highly accelerates the production of new solutions and technologies (Tria et al., 2014; Fink and Teimouri, 2019). Simon (1962) vividly explained this point with an example about opening a safe[1]. Suppose we want to open a safe whose lock has ten dials, each with one hundred possible settings, numbered from 0 to 99. Then there are about 100^{10} possible answers. However, when a click can be heard when any one dial is turned to the correct setting, the total number of trial and error can be decreased to 100*10, trivial compared to the original number. For Simon, existing stable components that can be used for further creation fulfill the same function as a recognizable dial setting: they both guarantee

---

[1] Another interesting example of tying buttons can be found in Kauffman (2002) p.36.

recognizable progress towards the ultimate goals. By decreasing uncertainties in the underlying architectures, modular structures accelerate discovery and problem-solving in the system.

This perspective soon stimulates a stream of hypothesis-testing studies, many of which are based on computational simulations. Arthur and Polak (2006) use logic functions as building blocks of more complex logic structures, and attempt to produce specific logic outputs by thousands of steps of recombination. Fink and Teimouri (2019), on the other hand, use data from three practical domains—language, gastronomy, and drug production—to simulate the recombinational process to attain existing complex structures. Both studies found evidence for modularity: sophisticated structures are not built up from scratch, but are results of new combinations of existing sub-components. With time, early appearing blocks can be replaced by more efficient ones that appear later. This phenomenon is discussed as "modularity," "recursiveness," and "technical recursion," etc. based on the intellectual contexts.

Literature in organizational studies has long adopted the modularity principle in organization building (Albert, 2018). However, this principle is relatively less frequently mentioned in studies about business innovations, partly because studies of a single innovative aspect (technology, business model, marketing, social relation, etc.) tend to render multidimensional complexities irrelevant. Traditional dualism, like exploration vs. exploitation (March, 1991; Raisch and Birkinshaw, 2008), incremental innovation vs. radical innovation (Acemoglu, 2020), and component innovation vs. architectural innovation (Henderson and Clark, 1990), often reflect some outcomes of modular structures: for example, higher-order recombination is more likely to appear as architectural innovation than component innovation, as the sub-components are kept unchanged in this dynamics. However, modularity is more about component structure than component distance, and more than two levels of modular structures can hardly be expressed properly with dualist concepts. Using the idea of modularity from complex theory, we aim to dissect the architecture of entrepreneurship search, which had been overshadowed by traditional dualism thinking.

Second-Order Recombination: The Key to Entrepreneurial Success

To accurately describe a modular structure in the economic landscape, it is necessary first to choose a proper function as a basic unit. Kauffman (2002) endows the autonomous agent (basic individual unit in a complex system) with an elegant definition: “a self-reproducing system able to perform at least one thermodynamic work cycle”. As open thermodynamic structures, autonomous agents are driven by outside sources of matter and energy (like food), and always in a non-equilibrium status to complete a circular process where they solidify free energy into the system. All autonomous agents can act on their own behalf, even influence the environment.

Following this definition, we argue that the basic work cycle of the economy is to serve some human needs with technical tools, and a basic unit in the business system should be the smallest module that can complete such a cycle. Here, both “technical” and “needs” might seem a little ambiguous. Although sociological theories have revealed the complexities of human desires, here, for brevity, we simply defined technical tools as all recombination of possible productive means to generate any economic benefit, and render satisfied needs as successful applications of these technical means in reality.

Based on such defined function units, many interesting theories about organization forms are born. Williamson (1979) believes most organizations in the market fall onto some point on the continuum between market partners and hierarchies, depending on the transaction cost of activities between the basic functional units--every unit has produced some outputs that might satisfy the needs of another party, and whether this product should be attained by the market or by bureaucracy simply depends on the cost and efficiency for different situations. The transaction cost theory focuses more on the modularity in the organizational structure instead of innovative strategies, yet it takes the same underlying assumption and adopt basic function units as we defined here.

Following a similar logic, for innovation strategies, there are also two levels of recombination existing in the landscape: the first-order recombination to align technologies and applications to form a basic functional unit, and the second-order recombination to align these functional units to solve more complicated problems. Jones and Summers (2020) argue that every bit of growth in the economy comes from science and technological advances. We broaden this point, as we measure the higher-order capacities by linking together the basic modules in a firm and distinguish the growth from basic technological solidification and from creative recombination. Multi-dimensional interactions in the opportunity space bring new possibilities.

Traditionally, first-order recombination falls in the scientific and technological sphere, where scientists explore radically new technologies, like radar, the turbojet, or the polymerase chain-reaction (Arthur, 2007). In contrast, new ventures are often established based on relatively mature technologies, which are brought into business activities by entrepreneurs with long-term working experiences in the field. The realization of a new business idea, especially in the initial phase, heavily relies on the exploitation of existing elements rather than exploring the unknown (Rippa et al., 2019). One main reason for this is that new ventures face resource limitations in nearly every aspect: finance, human resources, reputation, etc., due to their newness and small size (Jaroslaw, 2021). This scarcity puts constraints on their ability to explore. When intensive technical support for deep knowledge is in need, it is usually realized through collaborative agreements with universities and research institutions, instead of building from scratch (Galloway et al., 2017; Guerrero et al., 2019). This helps the new ventures to get a solid basis for future development. Thus, we argue that, for new ventures, high-level first-order recombination leads to less sufficient exploitation and higher probability of failure.

However, exploration does exist in business activities, from hiring highly diverse employees, attracting new investors, to opening new business sectors. In this work, we hypothesize that innovative strategies are embedded in modular structure: while the first-order module calls for exploitation, second-order recombination requires exploration.

Exploration is the source of new possibilities. It brings in new perspectives, insights, and knowledge, triggers the "creative conflict" (Stark, 2009; Kaplan and Vakili, 2015), brings in rich resources, including expertise, funding, market shares, and new ties in social networks (Aggarwal et al., 2017; Stanko and Henard, 2017). Partners and insights from a remote area often offer new possibilities that cannot be imagined before. In addition, even if the new findings from the exploration process are invalid, exploration generally adds to the diversity of spare elements in the system (Simon, 1962; Arthur, 2009).

Most of these benefits come from cross-domain interactions. Just as most information and resources in a closed community tend to be redundant (Burt, 1995), within-module knowledge lacks the necessary heterogeneity to produce efficient debate, and can hardly increase diversity in the system as most elements are so similar. In fact, most of the within-module elements are naturally approximate to each other, so that first-order modules can be formed by probability. Based on this logic, we argue that exploration is more important in second-order recombination, and high-level second-order recombination leads to a higher success rate and lower probability of failure.

Computational Methodologies: Dynamic Word Embedding for Landscape and Discourse Atom for Modularity

Early innovation studies often base their measurements on an existing patent or business classification systems, and take a novel combination of sub-classes as a sign of novel emergence (Fleming, 2001; Taylor and Greve, 2006). Despite the insight, these methods rely on a strong assumption that the classification and citing systems are naturally accurate and kept constant across domains and time. This assumption becomes suspicious in the highly dynamic market scenario, where new elements emerge every day and industry can be revolutionized within a decade. For this consideration, recent studies try to take advantage of the latest NLP techniques and extract information from technological and business-related corpus to generate more refined measurements (Kaplan and Vakili, 2015; Shiller, 2019).

Two NLP methods are most used in current innovation and novelty studies: topic modeling and word embedding. Topic models use a probability model to analyze a corpus and generate a representation of the latent "topics", a sparse distribution over words that tend to co-occur in the text (Blei, Ng, and Jordan 2003; Mohr and Bogdanov, 2013), while word embedding refers to a set of models that represent words with low-dimensional vectors and capture complex semantic relations geometrically (Mikolov et al., 2013; Pennington et al., 2014; Joulin et al., 2016). The geometric relationship between word vectors serves as a proxy of elementary relationships for the underlying social facts, and a useful tool to replicate the landscape for reality. See Figure 1 for the conceptual example of the word embedding model.

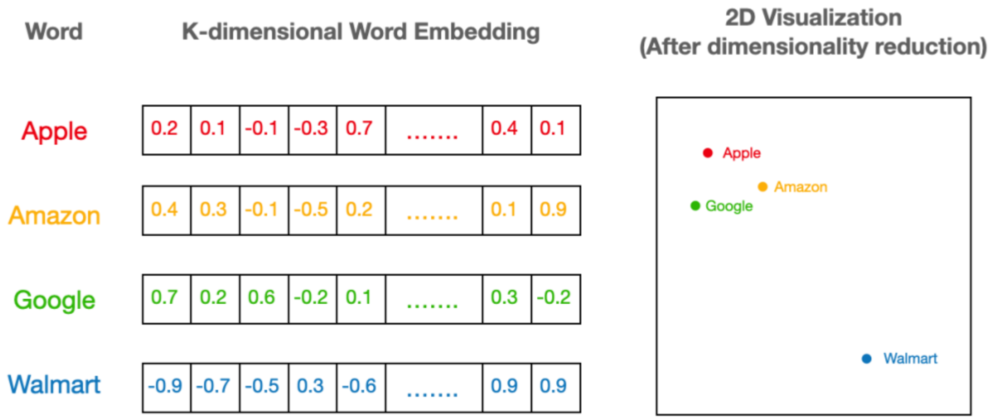

Figure 1. Conceptual Figure for Word Embedding

The family of word embedding algorithms aims to solve the optimal mathematical expression of words that best preserves distances between words across their semantic contexts. The main output of these algorithms is an n-by-k matrix, where n denotes the number of the words and k denotes the dimensions of word vectors. Vectoral expressions of words can be subsequently fed into various distance functions, mostly cosine similarity or Euclidean distance, to calculate the semantic relations. A low distance between word vectors shows close meanings or even identification as synonyms, while a high distance in the semantic space is the sign for irrelevance (when the angle between vectors is 90 degrees) or even contraries (180 degrees). The inner structure of the semantic space also allows researchers to solve analogy problems by applying simple linear algebra. A well-known example is given by Mikolov et al. (2013), where word vectors $\overrightarrow{king}$ - $\overrightarrow{man} + \overrightarrow{woman}$ leads to the vector for $\overrightarrow{queen}$. Classic word embedding model started in the 90s and relied heavily on mathematical calculation, like singular value decomposition (SVD), which constrain the number of documents and lower limits on the size of semantic contexts they could factorize. Later approaches, including word2vec (Mikolov et al., 2013) and GLOVE

(Pennington et al., 2014), have greatly improved both training efficiency and performance of word representation accuracy.

Powerful as they are, these techniques do not usually consider the temporal change of semantic space, assuming a static meaning structure across time. For time-comparison scenarios, dynamic word embeddings models have been developed to compute time-aware word embedding spaces that are comparable across history.

Most dynamic word embedding studies follow one of two patterns: either they align the word embeddings across time slices, or they learn the embeddings across time jointly. The first set of methods are often based on the linear transformation between time-slicing spaces, which can be achieved by solving a d-dimensional least-squares problem of k nearest neighbor words (Kulkarni et al., 2014), solving a d-dimensional Procrustes problem between every two adjacent times (Hamilton et al., 2016), or just assuming there are some “anchor words” whose meaning does not change between slices and can be foundations for a coherent spatial structure (Zhang et al., 2016). However, the aligning process often leads to loss of words, and thus information loss in the model. In the business world, where revolutionary innovation often comes from some previously unknown, marginal elements, this problem can cause inaccurate estimates.

For the current study, we use the method developed by Yao et al. (2018), which trains the temporal embeddings jointly. Algorithms such as negative sampling are proven to be equivalent low-rank factorization of PMI (Levy and Goldberg, 2014). The key idea behind the traditional word embedding methods, including word2vec and GLOVE, can be taken as to find embedding vectors *Uw* and *Uc* such that $u_w^T u_c \approx \mathrm{PMI}(\mathcal{D}, L)_{w,c}$, . Here, for the dynamic embedding task, we add an L2 norm of the difference between adjacent embeddings to the cost function, which allows us to keep adjacent word embedding space relatively comparable:

$$\min_{U(1),\ldots,U(T)} \frac{1}{2} \sum_{t=1}^{T} \|Y(t) - U(t)U(t)^T\|_F^2 + \frac{\lambda}{2} \sum_{t=1}^{T} \|U(t)\|_F^2 + \frac{\tau}{2} \sum_{t=2}^{T} \|U(t-1) - U(t)\|_F^2.$$

By optimization, we get a *t***n***k* matrix, where *t* is the number of time slices, *n* is the number of words, and *k* is the dimensions of word vectors. This method keeps the words in slices where they are actually absent, keeps adjacent semantic space structurally close to one another, yet represents the time-specific feature of each slice. Through this process, dynamic word embedding generates a temporal mathematical description of the environmental landscape. For each year, relationships of elements are presented in a semantic space, a slice of a long semantic sequence.

Word embedding helps us to attain a mathematical description of the business landscape, yet the hierarchical structures in space have not been revealed. To attain a realistic model for modular structures, we apply the algorithm of discourse atoms (Arora et al., 2018).

The discourse atom method is developed for word sense detection. It accepts word2vec or GLOVE vectors as inputs, and clusters words into "atoms" that have highly correlated meanings. Table 1 shows some of the examples of atoms in our corpus. The discourse atom algorithm is better than the topic model in the aspect of cluster separation: from table 1, it is obvious that elements within each atom are closely related, yet cross-atom correlation has been reduced to the utmost extent.

Table 1. Examples for Discourse Atoms

Discourse Atoms	Words in the atoms
Atom 1	business, buy, company, consumer, corporation, firm, industry, etc.
Atom 2	Acylcarnitine, albendazole, albuterol, alprenolol, etc.
Atom 3	Aluminum, americium, alkanethiols, alloy, angstrom, etc.
Atom 4	Afghanistans, Africa, ambassador, algerias, angola, etc.

The highly consistent semantic blocks from the discourse atom algorithm help us identify the system's basic sub-components. Based on the high within-atom consistency and inter-atom difference, each atom can be regarded as a basic functional unit (in the following article, we will call it “module”), and within-module interactions serve as the basis for higher-order structures. Examples and details can be further found in the *Data and Measurements* section. For robustness check, we (1) change the total number of atoms, and (2) replace discourse atoms with k-means clusters, another traditional and commonly used method for clustering (MacQueen, 1967). The results of the robustness check can be found in the Appendix.

Data and Measurements

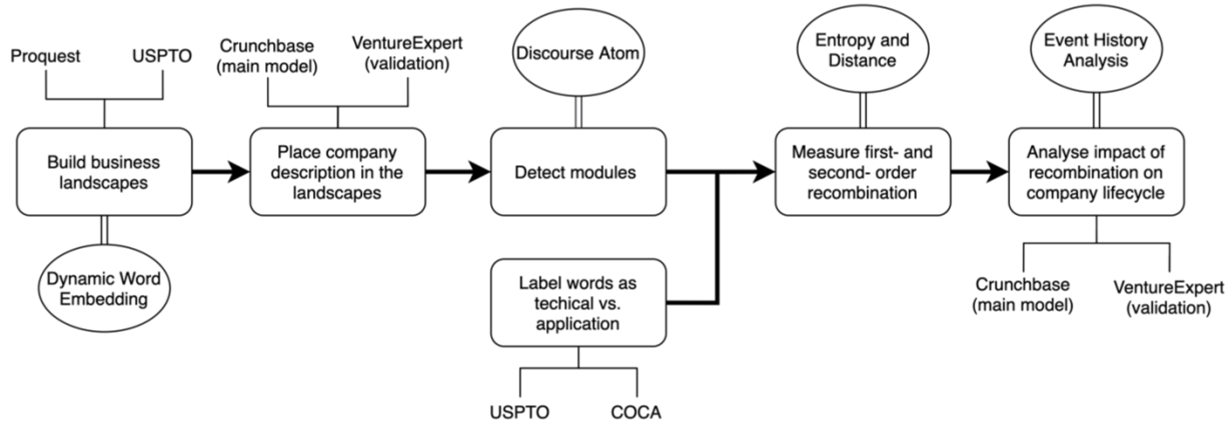

Figure 2. Data Analysis Process

Our research analyzed several large-scale datasets to observe the impact of high-level innovation on new venture success. We firstly constructed a time-varying business landscape using the dynamic word embedding algorithm, as mentioned in the previous part. Two datasets, the ProQuest dataset, and the USPTO dataset are used to provide the full-text contextual information of business newspapers, magazines, and patents from 1976 to 2020. We then placed textual descriptions of start-up companies in the constructed landscape, to locate strategies of these companies in the corresponding historical contexts. The start-up descriptions come from the CrunchBase dataset and VentureExpert dataset, analysis results of the formal are reported by this article, and results of the latter are used as validation. We applied the discourse atom algorithm to the annual landscape to reconstruct the modular structures of the business sphere. We also labeled the word elements within the business landscape as either “technical” or “application”, by developing a frequency filter using the USPTO dataset and COCA dataset. With the modular structure detected and element characteristics labeled, we then build entropy and distance measurement to capture the first- and second- order recombination that happened within each company. Finally, we analyzed how different levels of recombination predict the success and failure of new ventures. Data for the major success/failure events of start-up companies are also retrieved from CrunchBase (main results) and VentureExpert (validation). Figure 2 visualizes the entire analysis procedure, with the rectangle blocks representing the analysis steps and the eclipse blocks representing the methods.

Build Business Landscape

For our task to construct the modularity structures, we first constructed a time-varying business landscape by a dynamic word embedding algorithm. To build this context space, we use 119 English newspaper and magazine full-text articles in the ProQuest database, published in the U.S. and falling into the category of “Business and Economics”. This corpus contains major business publications, including the Wall Street Journal, Bloomberg BusinessWeek, the Economist, and American Banker. It provides temporal information about relationships between business elements, based on how reportage discusses them in the context of contemporary and imagined future business. In total, our dataset includes over 6 million business news records between 1976 and 2020. The business publications provide a temporal tracing of commercial knowledge and market emotions; however, while the media contains much valuable information about business models and applications, it might lose sight of the technology frontiers. To solve this problem, we complement this corpus with relevant patent full-texts from the USPTO (1976-2020) to ensure that both technical elements and business models are included in the analysis. Weights have been adjusted to keep the business media and technological corpora balanced. For a brief description of the ProQuest corpus, please turn to the Appendix.

As mentioned in the methodology section, we adopted the algorithm developed by Yao et al. (2018), who calculate temporal word embedding by adding an L2 norm to the difference between adjacent embeddings to the cost function. This algorithm allows us to compute a 45-slice word embedding, with each slice represents the distribution of business elements (i.e., words) within a given year, and revealing the year's social and economic features.

Figure 3 shows the closest elements to "Amazon" across time in this space. Although both the business landscape and the business identification of Amazon have changed, we can see a broad trend that Amazon has transformed from tropical forest to internet mogul.

In Figure 4, we validate this landscape with the performance of the real estate industry in the stock market. We first build a dimension of profit or loss, based on the method from Kozlowski et al. (2019). We project the vectors for three words, "estate", "homeowner" and "house", onto this dimension. The profiting score from this vector projection algorithm shows the market emotion to the real estate industry, based on news coverage from major business publications. Results show a high correlation between this projection series and the NAHB housing market index, a ground truth index based on surveys collected by the national association of home builders. For 2008, we can see a collapse of the real estate industry, which leads to a systematic crisis in the financial market (and collapse of the S&P 500 index).

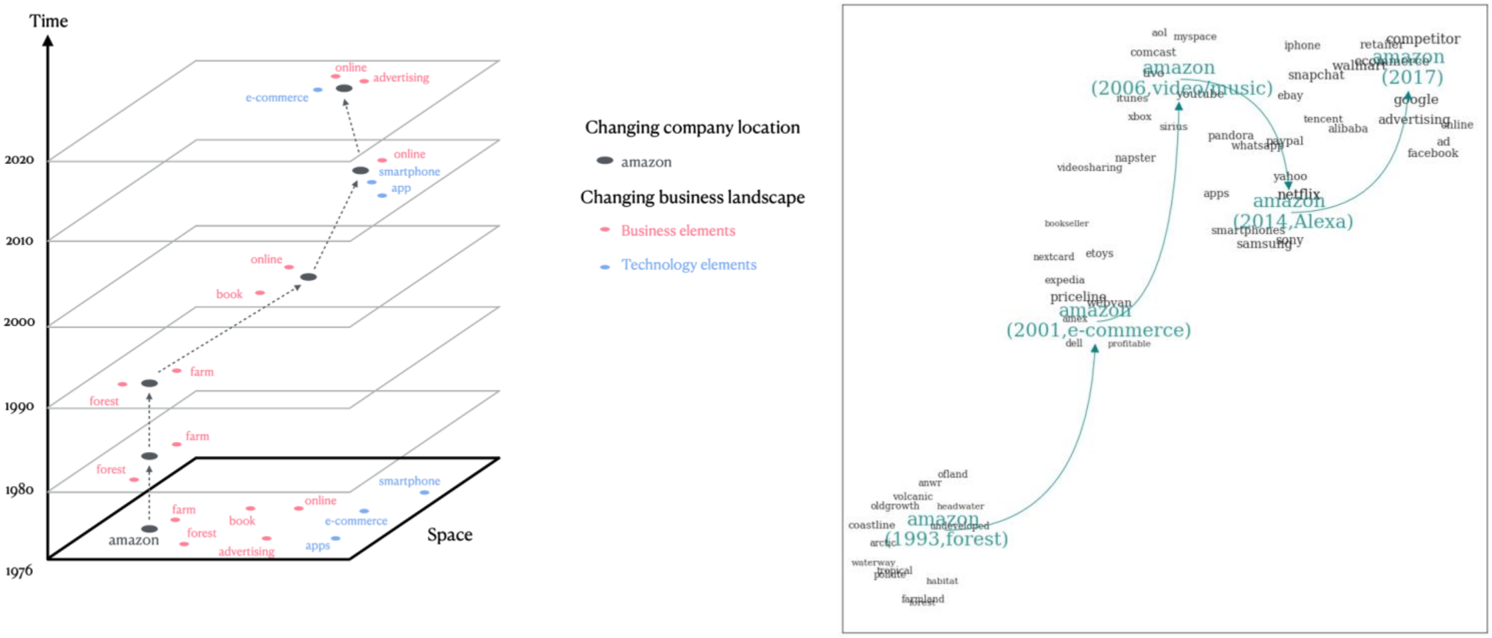

Figure 3. Amazon in Dynamic Business Landscape

(left: conceptual; right: U-MAP visualization;)

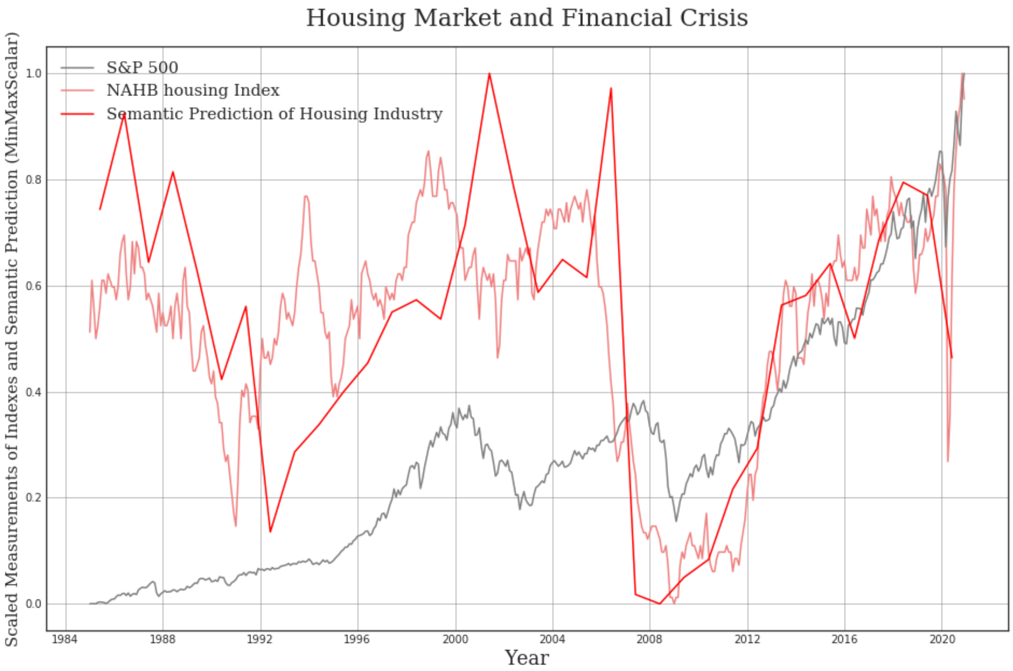

Figure 4. Validation of Dynamic Word Embedding Landscape with the Stock Market

Map Startups in business landscape

After constructing the temporal business landscape, we map individual start-ups within the market contexts. We use two databases for start-ups: CrunchBase (CB), a commercial database serving entrepreneurs and investors (N=298,915 by 2020), and VentureXpert (VX) from Thomson Reuters (N=63,492 by 2018). Both datasets provide a long description of each company, which provides detailed information about its major markets, relevant technologies, and strategies. As companies often use online platforms to establish their public images, attract potential investors and find collaborators, we assume that most companies have highlighted the most relevant information about their market niche and have been keeping them accurate. In the analysis, we match each year's company descriptions in the corresponding business landscapes, identifying meaningful description words that appear in the landscape (e.g., remove stop words), and viewing these words as the container of company innovative strategies.

CB and VX also record major events in company lifecycles, including founding date, funding time and amount, initial public offering (IPO), acquisition, company closure etc, which we will use in later event history analysis. Descriptions about companies and some examples of business descriptions can be found in the Appendix. Although there is a small overlap of companies between CB and VX, the descriptions for the same company are generally different in the two databases. Hence, we treat the two datasets separately in our analysis—CB produces the major results, and VX repeats the entire analysis for validations.

Detect Modular Structure and Label Word Elements

We then move on to reconstruct the modular structures within companies' innovative strategies. We apply the discourse atom algorithm on the word vectors of each slice of business landscape (space for a single year) to cluster the space into 80 modules. We also use another clustering algorithm, the K-means algorithm, for validation (see Appendix for details). We then check how individual company descriptions, which were matched in the landscapes earlier, fall into different modules. For each company, the recombination of elements within a module is taken as the first-order recombination, while cross-module recombination is defined as second-order ones.

Among all words within a company description, some words focus on the application of the new venture (like “retailing”, “platform”), while others focus on technological components and processes that allow the venture to compete in the market (like “dopamine”, “pathophysiology”). We separate these two kinds of words by building a frequency-based filter. To build the filter, we first calculate the word frequency in USPTO patent and in COCA (Corpus of Contemporary American English). COCA is a linguistic corpus of daily-used English, which includes common newspaper full-text, like the New York Times (but not the business newspapers in the ProQuest dataset). We assume that words from the technical description section of patents with a lower than 500 frequencies in COCA are “technical”, as they appear in patents but not very often in common English. For the non-technical words with more than 500 frequencies in COCA, if they also appear in patent claim texts, then we label them as “application”-related, meaning that they describe the application of technical elements. The basic assumption is that technological words appear more frequently in the patent corpus than in the news corpus, while business model words are more commonly used and appear with similar frequency in both. After we build the frequency filter, we use it to label description words of individual companies. We also try to analyze the distribution of technical and application words within modules.

This filtering method might look a little arbitrary at first glance, as technical elements are not always equal to infrequent terms. To validate this measure, we apply two other methods for technical/application word separation. The details of these methods can be found in the Appendix; overall, results from the three methods show similar patterns.

First- and Second-Order Recombination

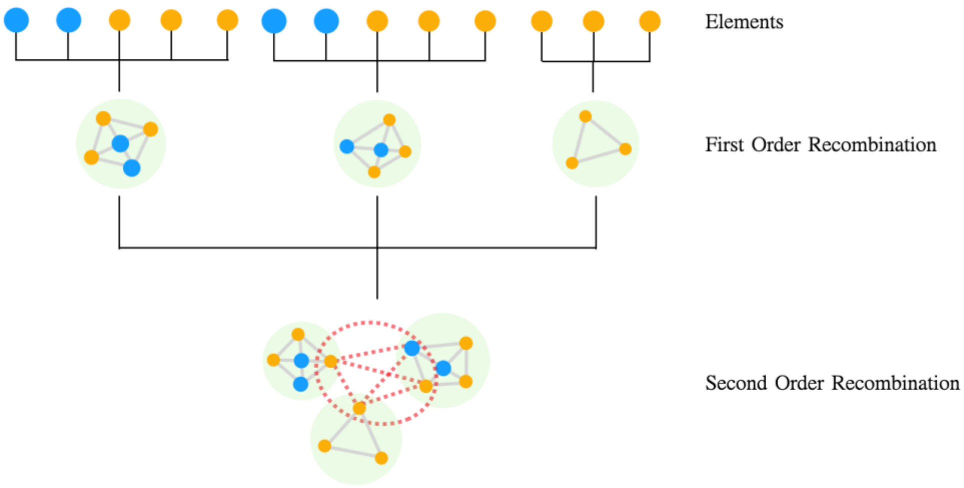

Figure 5. Conceptual Figure of First- and Second- Order Recombination

Now we can construct our measurements for our theoretical concepts. Figure 5 shows the conceptual graph of the first- and second- order recombination. In the figure, blue dots represent technical elements, orange dots represent application elements, and green circles represent modules. For first- and second-order recombination, we adopt the measurement of Shannon entropy, a widely used measure for information complexity. The equation is given by:

$$H(X) = -\sum_{x} P(x)\, log_2\, [P(x)]$$

Based on this equation, we measure the first-order and second-order recombination as follows:

First-Order Recombination: For each firm, with its elements distributed in multiple modules, we calculate the entropy for relevant technical/application elements in each module, and calculate a weighted average across modules. A high average entropy reveals the firm is built on balanced blocks of technical and application elements where technical elements are closely combined with practical capacities. To bring close elements together into the new venture strategy shows a strong signal of adopting past commercial experiences. Here, the entropy measures the maturity of a basic functional unit, and is named as "**within entropy**" in later text. Figure 6 shows an example of a company having high vs. low within entropy.

Second-Order Recombination: For each firm with its elements distributed in multiple modules, we first generate a distribution of elements across modules. **A global entropy** is calculated based on this element-module distribution, to measure the extent to which all elements concentrate in one domain, or spread across domains. This measure reveals the extent to which the firm has recombined technical application blocks from different fields. A high global entropy reveals a high level of high-order recombination. Figure 6 shows an example of a company having high vs. low global entropy.

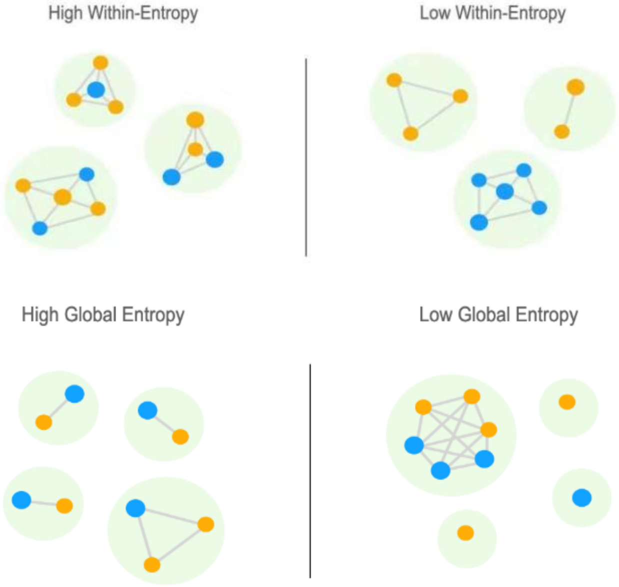

Figure 6. High vs. Low Entropy

We also complement the entropy measures with distance measures, which reflect similar concepts but from a different perspective. **Inner closeness** measures the average Euclidean distance between relevant technical elements and application elements in each module. A low value shows that the technical and application elements are closely combined in the business history, thus the technical application block has been frequently used in the past. If the inner closeness is high, the focal firm should have taken advantage of many such commonly used technical applications, thus has a lot of exploitation. **Outer distance** measures the average distance between centroids of all relevant clusters. A high distance in the semantic space shows that the elements are never aligned, actually never imagined to be related, in history. Thus, if a firm has brought previously remote elements in the space, there is a strong signal for extremely novel connections across domains compared with historical records. Figure 7 shows conceptual examples of high vs. low inner closeness and outer distance. Results with distance measures are also provided in the result section. For empirical examples of firm components, please refer to the Appendix.

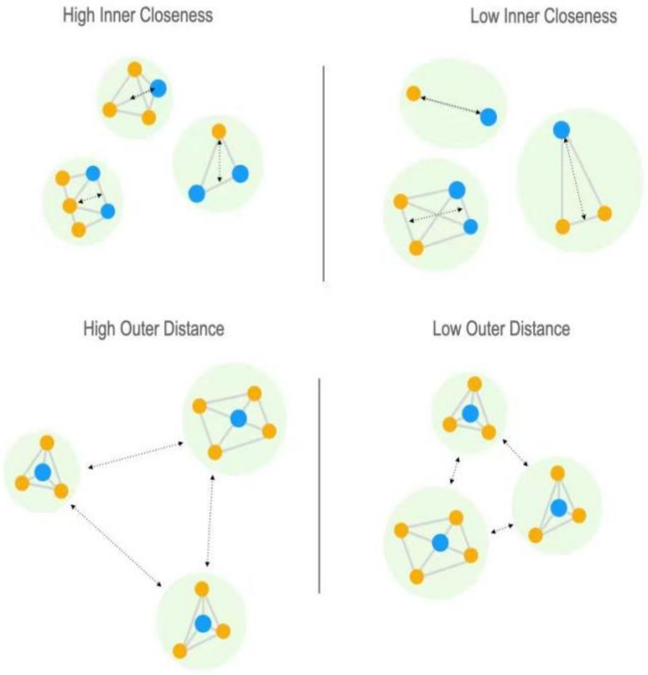

Figure 7. High vs. Low Distance

Event History Analysis

After calculating the scores for first- and second-order recombination, we try to use them to predict the business success of new ventures. Several major events may occur in the life course of a start-up: obtaining funding, being acquired by other companies, doing public with an IPO, or ignominiously closing down, each a milestone that reveals the competitive strengths and development potential (or ultimate failure) of an individual company and the success of its investors (Ljungqvist et al., 2004).

For this life cycle structure, we adopt the multi-outcome event history model to explore how innovation shapes the future of new ventures. As many of the outcomes are mutually exclusive (for example, a start-up can't become successful while go bankrupt at the same time), we apply the cause-specific flexible parametric survival model (FPM) by Mozumder et al. (2017). This method allows us to identify the factors that lead to a specific outcome, taking all other possible outcomes into consideration.

Besides the four major independent variables (i.e., entropy and distance), we also use some controls in our event history model to rule out possible alternative explanations. The controls are: (1) text features of the business description, including text length (numerical), the existence of extremely rare words (dummy), and lack of technical elements (dummy); (2) features of the company, including the funding history (number of historically accepted fundings in the observation period, numeric), founders' identity (whether the founder is a woman or minority, dummy), and media usage of the company (whether the company uses Facebook and Twitter); (3) the market environment in the year when the venture was founded, measured by the growing rate of companies in the same industry; (4) temporal-spatial features, including the time for observation (year group, categorical), and region of the company (midwest, west, northeast, and south, categorical).

Results

Descriptive Analysis

We firstly calculate the average entropies for each industry, classified by the Crunchbase database. Figure 8 shows the level of within entropy (x-axis) and global entropy (y-axis) of different industries. The intersection point of x-axis and y-axis represents the average within entropy and global entropy. From the figure, we could see that high within-entropies tend to appear in the technologically intensive industries, including biotechnology and science, whereas military, AI, cyber security, data science etc. have the highest global entropies. It is also interesting to note that most industries are either explorative and exploitive in both first- and second- order recombination.

The entropy and distance measures are numeric measurements, yet we can transform them into categorical variables show a simplified pattern of correlation between first- and second- order recombination and performance outcomes. We turn each of the numeric measurements into ten groups, with percentiles as breakpoints, and leave those with a 0 value as a single group. The correlation patterns are shown in Figure 9 and Figure 10.

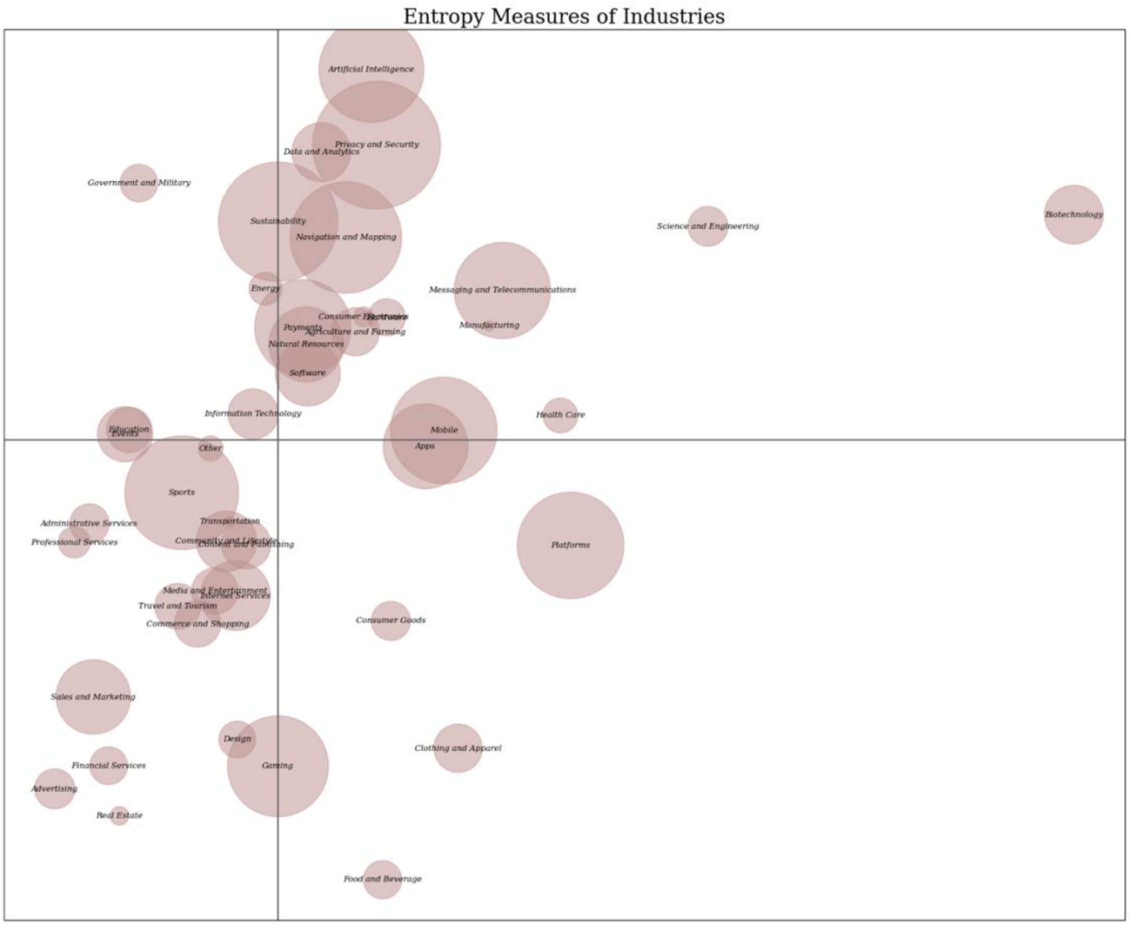

Figure 8. Entropy measures of different industries

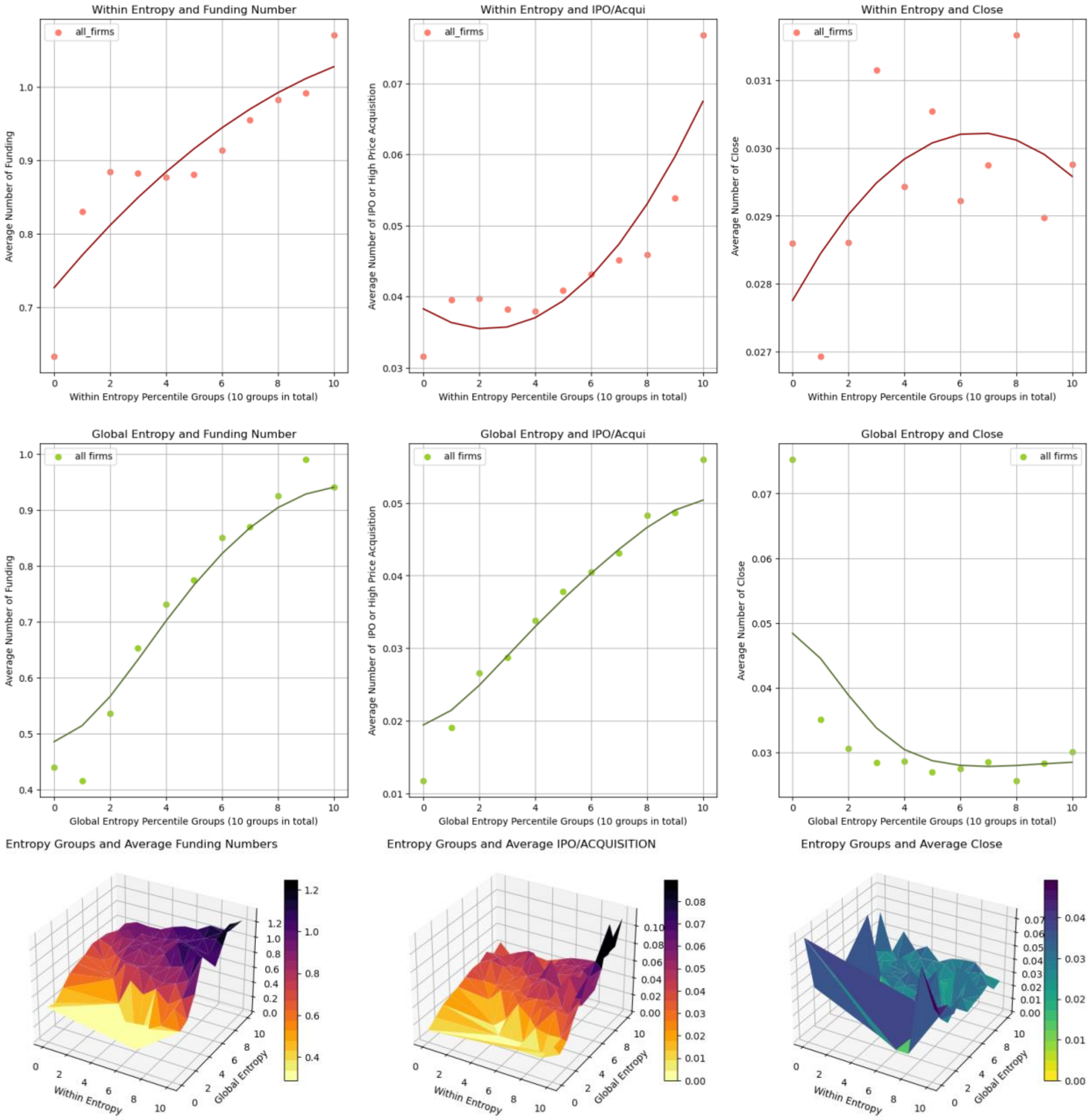

Figure 9. Relationship between Entropy Measures and Venture Outcomes

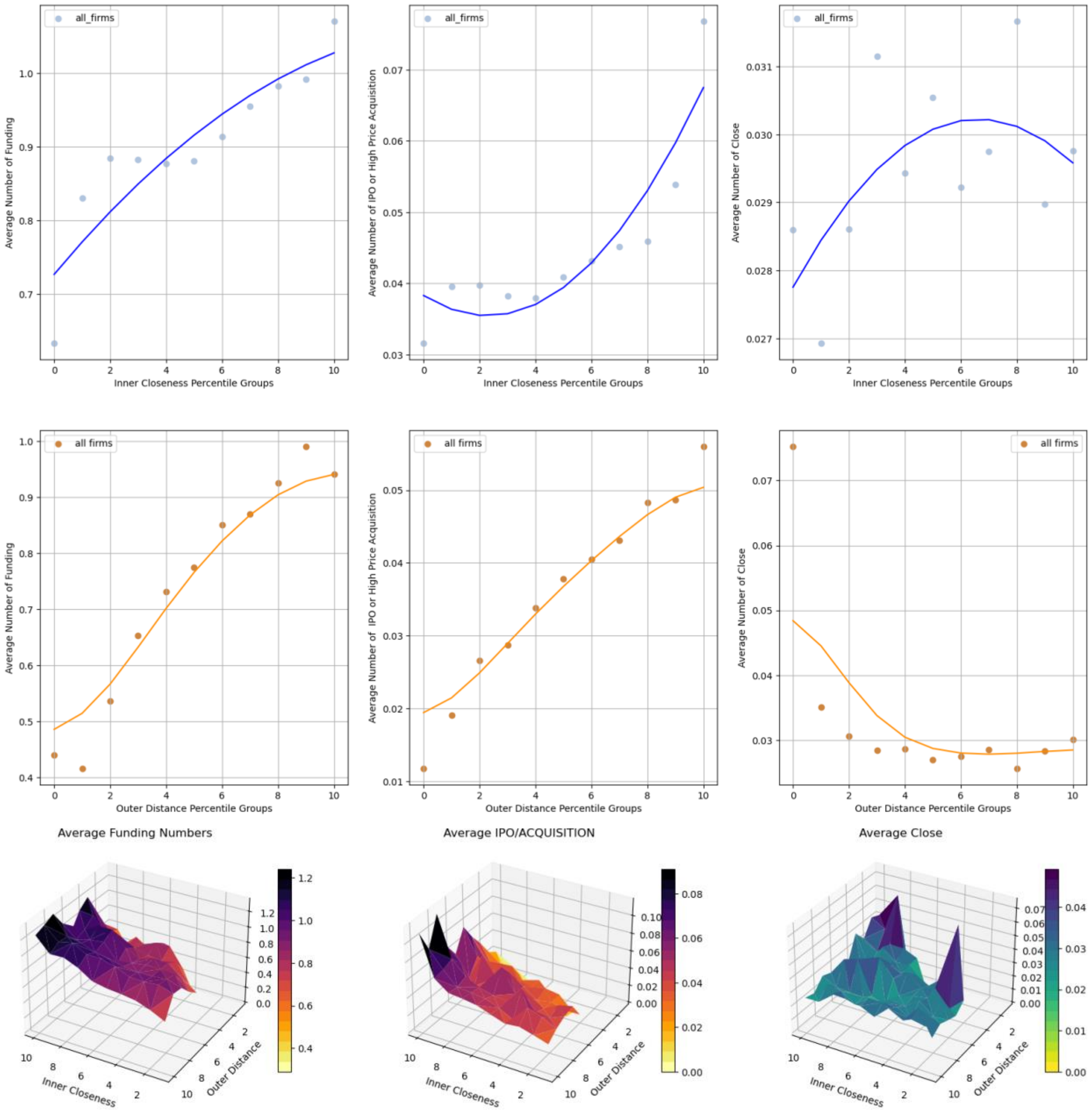

Figure 10. Relationship between Distance Measures and Venture Outcomes

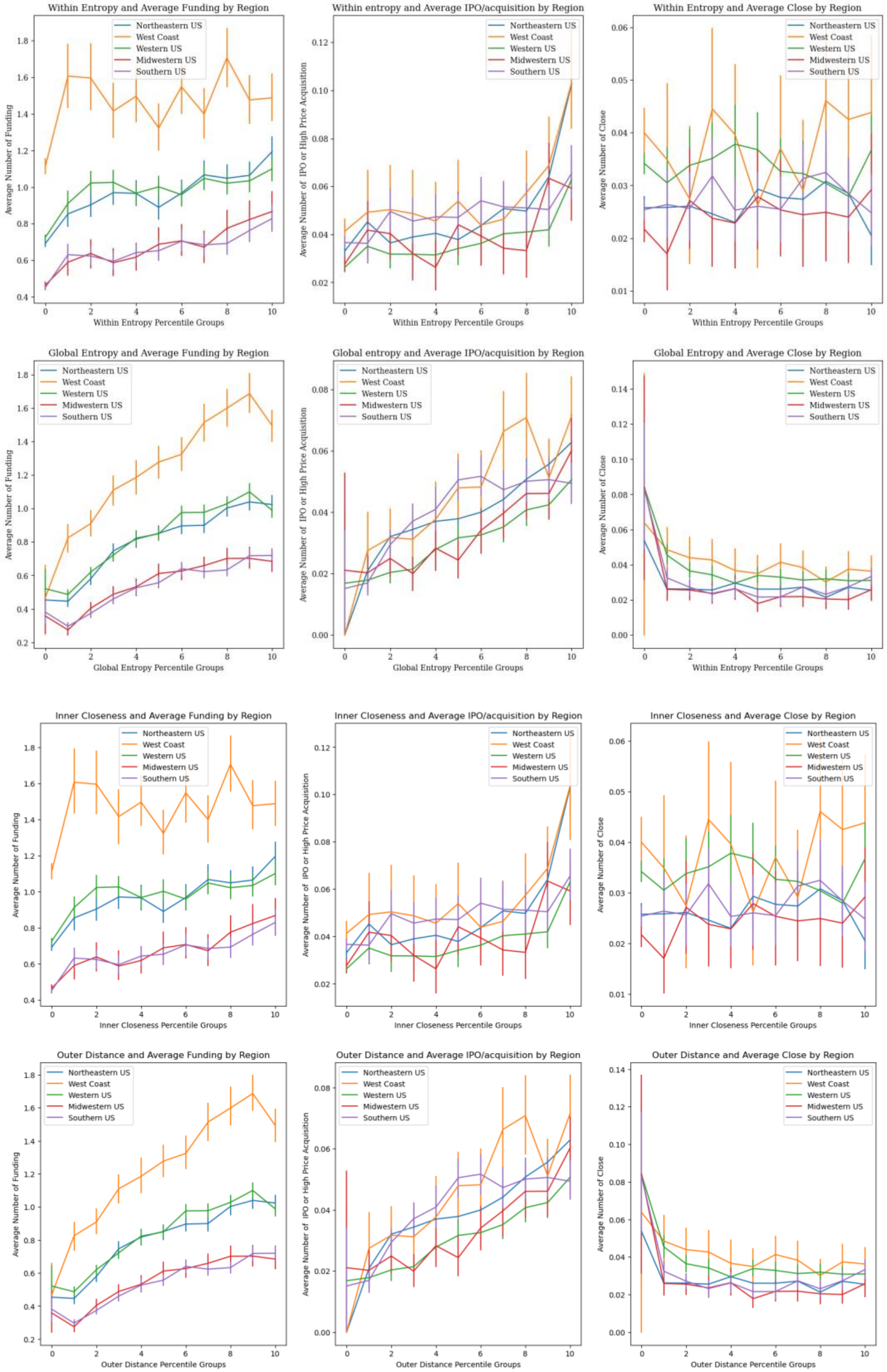

Figure 11. Correlation between Entropy/Distance Measures and Venture Outcomes by Regions

Evidence from Figure 9 and Figure 10 supports our hypotheses. Involvement in a mature module with closely combined technical elements and application elements (high within entropy and high inner closeness) provides the company with a solid basis for long-term development, and grants them a higher probability of receiving new funding or going IPO one day. High global entropy, which shows the breadth of modules, also has a positive influence on new venture success and prevents the start-up from being closed. It should be noted that outer distance shows a contrary effect on number of times funding is received--although a high outer distance does help the start-up to go IPO, it is not helpful when the venture needs to collect new investment. It reminds us of the dark side of innovation: wide exploration often brings in high costs and uncertainties that small companies cannot easily bear. Yet, for large companies close to IPO, it might be an advantage.

We group samples by regions to show more detailed patterns. The results are shown in Figure 11-12. Overall, the figures show similar patterns as Figure 9-10, while revealing additional interesting patterns: the west coast is most successful in breeding new ventures, yet it also contains the highest uncertainties and risks, as captured by the average number of companies being closed. Early established companies have the highest probability to be public and lowest probability to be closed, showing a potential survivorship bias.

To test the validation of these results, we change several significant parameters in our calculation to see if the results are robust. These parameters include the method to identify technical and application words, the method for space division, and the number of total clusters by space division. Results of these robustness checks can be found in the Appendix.

Figure 9 and Figure 10 provide a preliminary testification of our hypotheses, yet the temporal feature of the data calls for better-designed models to take the right censoring into consideration. Under current observation, recently established companies may not have enough time to show their potential for success (or failure). Here, we apply a multi-outcome event history model to solve this problem. Specifically, we apply the competing risk model from Mozumder et al. (2017), as it allows us to estimate the hazard rate of a specific outcome when taking all other possible outcomes into consideration. We also apply the cox proportional hazards model as validation, which is another method recommended for competing risks (Blossfeld and Rohwer, 2007).

Event History Analysis

Table 2 shows the descriptive statistics of major variables that will be included in the event history analysis, including entropy, distance, and company performance. From the statistics in Table 2, we could see that the two databases, CB and VX, have similar value distribution for most variables. The only difference is that VX on average has a higher number of accepted funding and IPO, this could be because that CB, which mainly relies on volunteer reporting and scraping, might have more data missing. In general, statistics in table 2 guarantee that the two databases are comparable, and the analysis results are robust.

Before estimating the model, we also generate a correlation table of the variables (Table 3). In Table 3, we find a very high correlation between within-entropy and inner-closeness. The average distance between application and technical elements tends to be lower for those more mature modules, consistent with our intuition. To take this potential collinearity into consideration, we only include within entropy in our models and leave out the inner closeness measure.

Table 2. Descriptive Statistics of Major Variables

	crunchbase(n=197978)		ventureXpert(n=63018)	
	mean	std	mean	std
Within entropy	0.069	0.094	0.102	0.125
Global entropy	3.948	0.894	3.739	1.058
Inner Closeness	0.054	0.077	0.085	0.114
Outer Distance	5.383	0.855	5.609	1.042
Number of accepted funding until the current day	0.758	1.471	2.455	2.626
Number of IPO or high price acquisition until the current day	0.038	0.202	0.099	0.299
Number of middle or low price acquisition until the current day	0.173	0.378	0.172	0.377
Number of closing until the current day	0.029	0.168	0.13	0.336

The results of the event history model based on CrunchBase are shown in Table 4. Similar to other event history analysis, a company can have observations in the model, but IPO and closing will be taken as final exits, and any other events after a final exit (like revival after closing) will be ignored.

Results from event history models are consistent with our previous findings. While conservative strategies in the first-order recombination provide a solid basis for innovation, wide exploration in second-order recombination helps start-ups realize their full potential and leads to a higher probability of positive events. Reverse trends (exploration on the first order and exploitation on the second order) lead to a higher probability of selling cheaply or even close.

It is interesting to see an inverse-U relationship between outer distance and hazard rate of getting new funding, while this relationship remains monotonic for the outcome of IPO/high price acquisition. One possible explanation is the ability limitations of new ventures. Combining the element far away from each other brings high risk and return at the same time. While this risk can be handled easily by mature companies near IPO, it might be out of control for smaller ventures and harm their potential to be accepted by the market.

Two models are estimated for robustness check. First, we adopt the Cox proportional hazard model as a replacement for FPM. The Cox model serves as one of the most commonly used cause-specific hazard models, and provides a decent estimation for hazard rate for a single event (Blossfeld and Rohwer, 2007). For the second check, we use data from VentureXpert and replicate the same model as in table 4. Note that adding the global entropy measure in the VentureXpert model prevents the model from convergence, possibly because it leads the model to generate a total CIF larger than 1 (Austin et al., 2021). For this reason, we turn the global entropy into a binary (larger than median vs. smaller than the median) to replace the original variable. Another difference is that VentureXpert data did not provide detailed information about closing dates, so we have not modeled firm closing as a competitive outcome.

Results from VentureXpert show very similar signals as Crunchbase when we include quadratic terms in the model. The only difference is the relationship between outer distance and middle-low price acquisition, which is also quite fuzzy in the original results. Thus, our hypothesis is mostly supported.

Table 3. Correlation Table

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13
(1) within entropy	1												
(2) global entropy	0.251***	1											
(3) inner closeness	0.942***	0.249***	1										
(4) outer distance	-0.103***	0.115***	-0.107***	1									
(5) number of fundings once accepted	0.111***	0.115***	0.101***	-0.042***	1								
(6) usage of social media	-0.028***	0.145***	-0.035***	-0.047***	0.108***	1							
(7) dummy: founded by women or minorities	0.015***	0.030***	0.010***	-0.040***	0.086***	0.119***	1						
(8) market growth rate in the year of establishment	-0.012***	0.029***	-0.012***	0.099***	-0.269***	-0.047***	-0.123***	1					
(9) dummy: no technical element at all	-0.386***	-0.447***	-0.373***	0.068***	-0.074***	-0.021***	-0.003*	-0.004**	1				
(10) dummy: more than 3 technical elements	0.481***	0.505***	0.458***	-0.122***	0.083***	-0.003*	-0.008***	0.023***	-0.421***	1			
(11) dummy: more than 40 application elements	0.134***	0.536***	0.164***	-0.089***	0.031***	0.072***	0.001	0.047***	-0.192***	0.334***	1		
(12) uncommon application words	-0.023***	-0.133***	-0.013***	-0.191***	-0.011***	-0.027***	-0.006***	-0.011***	0.031***	-0.023***	-0.013***	1	
(13) extremely uncommon technical words	-0.116***	-0.083***	-0.112***	0.057***	-0.038***	-0.024***	-0.027***	0.028***	-0.070***	-0.123***	-0.048***	0.001	1

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4. Cause-Specific Flexible Parametric Survival Model (FPM) Based on CrunchBase

	Linear Model				Quadratic Model			
	IPO/high price acquisition	Getting new funding	Middle-low price acquisition	close	IPO/high price acquisition	Getting new funding	Middle-low price acquisition	close
within entropy	1.424***	0.529***	-0.377***	-0.606***	-0.119	0.589***	-0.0324	0.605
global entropy	0.143***	0.154***	-0.241***	-0.195***	-0.00614	0.0957***	-0.0632**	-0.137**
outer distance	0.188***	-0.0187***	0.0687***	-0.0357*	0.378***	0.150***	-0.0812**	-0.190**
within entropy squared					3.093***	-0.814***	-0.0081	-2.245*
global entropy squared					0.0342***	-0.0146***	-0.0954***	0.00925
outer distance squared					-0.0124	-0.0135***	0.272	0.00828
usage of social media	-0.0485***	0.0772***	-0.0828**	-0.242***	-0.0453***	0.0857***	-0.215***	-0.260***
dummy: founded by women or minorities	-0.465***	0.288***	-0.194***	-0.209***	-0.456***	0.292***	-0.452***	-0.221***
market growth rate in the year of establishment	-0.038	0.0464**	-0.436***	-0.188*	-0.0247	0.0494***	-0.0762*	-0.191*
Year and Location	controlled	controlled	controlled	controlled	controlled	controlled	controlled	controlled
Text Features	controlled	controlled	controlled	controlled	controlled	controlled	controlled	controlled
Funding History	controlled	controlled	controlled	controlled	controlled	controlled	controlled	controlled
_rcs_c1_1	0.999***	1.023***	1.558***	1.249***	1.000***	1.022***	1.557***	1.249***
_rcs_c1_2	0.103***	0.430***	-0.201***	0.247***	0.103***	0.430***	-0.200***	0.246***
Constant	-5.994***	-4.119***	-3.626***	-6.560***	-6.101***	-4.118***	-3.863***	-6.612***

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5. Models for Robustness Check

	Cox Model for CrunchBase				VentureXpert Model (FPM)					
	IPO/high price acquisition	Getting new funding	Middle-low price acquisition	close	IPO/high price acquisition	Getting new funding	Middle-low price acquisition	IPO/high price acquisition	Getting new funding	Middle-low price acquisition
within entropy	1.85**	0.74**	0.40**	0.00	-0.717***	0.381***	-0.542***	-1.396***	0.498***	-0.408
global entropy	0.20**	0.15**	-0.22**	-0.15**	0.354***	0.0692***	-0.218***	0.345***	0.0664***	-0.220***
outer distance	0.21**	-0.01**	0.08**	-0.04*	0.0325*	0.0359***	-0.0186	0.254**	0.130***	0.202**
within entropy squared								1.454*	-0.284*	-0.349
outer distance squared								-0.0162*	-0.00836***	-0.0212**
usage of social media	-0.12**	0.01*	-0.27**	-0.40**						
dummy: founded by women or minorities	-0.38**	0.29**	-0.32**	0.03						
market growth rate in the year of establishment	0.06	0.01**	0.03	-0.04						
Year and Location	controlled	controlled	controlled	controlled	controlled	controlled	controlled	controlled	controlled	controlled
Text Features	controlled	controlled	controlled	controlled	controlled	controlled	controlled	controlled	controlled	controlled
Funding History	controlled	controlled	controlled	controlled	controlled	controlled	controlled	controlled	controlled	controlled
_rcs_c1_1					0.733***	1.017***	1.030***	0.733***	1.018***	1.030***
_rcs_c1_2					0.351***	0.417***	0.487***	0.351***	0.417***	0.487***
Constant					-2.858***	-1.485***	-7.786***	-3.549***	-1.743***	-8.322***

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Conclusion and Discussion

The results of the correlation and event history model both support our hypothesis of modularity. While closely combined elements constitute mature blocks and a solid basis for entrepreneurship, for successful new ventures, the innovation and exploration often happen on the second order, where existing modules are aligned to generate innovative niche. Business ventures, by their essence, are higher-order social creations that are based on established techno-social clusters.

Our study shows that exploration and exploitation can happen simultaneously in the organizational strategy, and achieve architectural ambidexterity that is not segmented by unit functions or individual time split. By adopting the complex system perspective in place of traditional linear thinking, we make the first attempt to break the dualism in organizational learning theories and generate a multi-dimensional understanding of organizational strategies. Exploration and exploitation are not essentially contrary, as is observed from a linear perspective, but co-exist in a symbiosis among all innovative possibilities. Our study is only the first attempt in this direction, yet it reveals the great theoretical potential of ambidexterity.

By adopting the state-or-art word embedding models, our study helps expand knowledge of innovation in two ways. First, our model allows us to study innovation from the perspective of a global business conceptual space. This provides us with the chance to trace how technologies and production capacities co-evolve in history, and model their interdependence across the economy. Second, our model is a temporal model that breaks the assumption of a static, ahistorical market. By including patents and media publications, we capture both the technological frontiers and Zeitgeist of the business era, and locate ourselves in the historical contexts where an entrepreneur decides how to construct a business in her unique time and context. Based on this landscape, we reconstruct the modular structures in the complex economy of a specific time.

Our work urges entrepreneurs, especially technical experts, to think about optimal business strategies and potential challenges in their career paths. Although technical expertise is valued in the financial markets, it alone does not support the success of a new venture. Technical experts can pay more attention to inter-domain knowledge and collaborate broadly with people from other areas. For the public policymakers, policies that help technical experts broaden horizons and develop business insights may generate more small venture successes and keep the market in vitality.

There are several limitations of our analysis. First, although the description texts have contained key information about business strategies of start-ups, they typically do not contain clues about other aspects of business backgrounds, including supply chain, customers, products, brands, etc. These elements, however, also play an important role in innovation and cannot be disentangled from the strategies of a new venture. Second, our samples are limited to start-ups, which are typically challengers in an existing industry. Only some of these new ventures grow up to be incumbents. Due to data limitations, we cannot reconstruct the whole picture of a business revolution process: the incumbents fall, the challenger rises, and a new era is born. The current version of the one-sided story provides us with some knowledge, but a dynamic description of interactions between the incumbents and challengers will be more informative for the whole cycle of creative destruction.

Our study is a preliminary attempt to model the innovative structures in the complex system of economy, yet the potential of this perspective is far from being fully realized. Future research may further explore the mechanism of organizational evolution in economic complexity, the financial, social and mental outcomes of these mechanisms, and change of optimal strategies in the life cycle of a market participant. Radical as it might sound, one ambition of the earliest social scientists is to find the laws that apply to the natural world and human world similarly (Cohen, 1994). A mathematical model or computational simulation of our conclusion, as well as a comparison between business functions and natural science functions, would be a possible way to serve this purpose.

References

Acemoglu, D., Akcigit, U., & Celik, M. A. (2020). Radical and Incremental Innovation: The Roles of Firms, Managers and Innovators. *American Economic Journal: Macroeconomics*.

Aggarwal, V. A., Posen, H. E., & Workiewicz, M. (2017). Adaptive capacity to technological change: A microfoundational approach. *Strategic Management Journal*, *38*(6), 1212–1231.

Albert, D. (2018). Organizational Module Design and Architectural Inertia: Evidence from Structural Recombination of Business Divisions. *Organization Science*, *29*(5), 890–911.

Alizadeh, Y., & Jetter, A. (2018). Pathways for Balancing Exploration and Exploitation in Innovations: A Review and Expansion of Ambidexterity Theory. *International Journal of Innovation and Technology Management, 16(05), 1950032.*

Arora, S., Li, Y., Liang, Y., Ma, T., & Risteski, A. (2018). Linear Algebraic Structure of Word Senses, with Applications to Polysemy. *arXiv:1601.03764 [cs, stat]*.

Arthur, W. B. (2007). The Structure of Invention. *Research Policy*, 36(2), 274-287.

Arthur, W. B. (2009). *The Nature of Technology: What It Is and How It Evolves* (Reprint edition). Free Press.

Arthur, W. B. (2020). *Complexity Economics: Proceedings of the Santa Fe Institute's 2019 Fall Symposium* (E. D. Beinhocker & A. Stanger, edited). SFI Press.

Arthur, W. B., & Polak, W. (2006). The evolution of technology within a simple computer model. *Complexity*, *11*(5), 23–31.

Austin, P. C., Steyerberg, E. W., & Putter, H. (2021). Fine-Gray subdistribution hazard models to simultaneously estimate the absolute risk of different event types: Cumulative total failure probability may exceed 1. *Statistics in Medicine*, *1-13.*

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *The Journal of Machine Learning Research*, *3*(null), 993–1022.

Blossfeld, H.-P., & Rohwer, G. (2007). *Event History Analysis With Stata* (1st edition). Psychology Press.

Burt, R. S. (1995). *Structural Holes: The Social Structure of Competition* (1st Paperback Edition). Harvard University Press.

Christensen, C. M., & Overdorf, M. (2000). Meeting the Challenge of Disruptive Change. *Harvard Business Review*.

Cohen, I. B. (1994). An Analysis of Interactions between the Natural Sciences and the Social Sciences. In I. B. Cohen (edited), *The Natural Sciences and the Social Sciences: Some Critical and Historical Perspectives* (page 1–99). Springer Netherlands.

Duncan, R. B. (1976). The ambidextrous organization, designing dual structures for innovation. *The management of organization design, 1*.

Erwin, D. H., & Krakauer, D. C. (2004). Insights into Innovation. *Science, 304*(5674), 1117–1119.

Evans, J. A. (2007). Start-Ups in Science: Entrepreneurs, Diverse Backing, and Novelty Outside Business. In M. Ruef & M. Lounsbury (edited), *Sociology of Entrepreneurship* (volume 25, page 261–305). Emerald Group Publishing Limited.

Fink, T. M. A., & Teimouri, A. (2019). The mathematical structure of innovation. *arXiv:1912.03281 [cond-mat, physics:physics]*.

Fleming, L. (2001). Recombinant Uncertainty in Technological Search. *Management Science*, *47*(1), 117–132.

Galloway, T. L., Miller, D. R., Sahaym, A., & Arthurs, J. D. (2017). Exploring the innovation strategies of young firms: Corporate venture capital and venture capital impact on alliance innovation strategy. *Journal of Business Research*, *71*, 55–65.

Garud, R., Tuertscher, P., & Van de Ven, A. H. (2013). Perspectives on Innovation Processes. *Academy of Management Annals, 7*(1), 775–819.

Gibson, C. B., & Birkinshaw, J. (2004). The Antecedents, Consequences, and Mediating Role of Organizational Ambidexterity. *The Academy of Management Journal*, *47*(2), 209–226.

Guerrero, M., Urbano, D., & Herrera, F. (2019). Innovation practices in emerging economies: Do university partnerships matter? *The Journal of Technology Transfer*, *44*(2), 615–646.

Hamilton, W. L., Leskovec, J., & Jurafsky, D. (2016). Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 1489–1501.

Henderson, R. M., & Clark, K. B. (1990). Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms. *Administrative Science Quarterly*, *35*(1), 9–30.

Hidalgo, C. A., & Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences*, *106*(26), 10570–10575.

III, C. A. O., & Tushman, M. L. (2002). *Winning Through Innovation: A Practical Guide to Leading Organizational Change and Renewal* (Revised edition). Harvard Business Review Press.

Jaroslaw, K. (2021). Process Ambidexterity in Startups Innovation. *Management Systems in Production Engineering*, *29*, 27–32.

Jones, B. F., & Summers, L. H. (2020). *A Calculation of the Social Returns to Innovation* (Working Paper; Working Paper Series). National Bureau of Economic Research.

Joulin, A., Grave, E., Bojanowski, P., & Mikolov, T. (2017). Bag of Tricks for Efficient Text Classification. *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, 427–431.

Kaplan, S., & Vakili, K. (2015). The double-edged sword of recombination in breakthrough innovation. *Strategic Management Journal*, *36*(10), 1435–1457.

Kauffman, S. A. (2002). *Investigations*. Oxford University Press.

Kozlowski, A. C., Taddy, M., & Evans, J. A. (2019). The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings. *American Sociological Review*, *84*(5), 905–949.

Kulkarni, V., Al-Rfou, R., Perozzi, B., & Skiena, S. (2015). Statistically Significant Detection of Linguistic Change. In *Proceedings of the 24th International Conference on World Wide Web* (page 625–635). International World Wide Web Conferences Steering Committee.

Levy, O., & Goldberg, Y. (2014). Neural Word Embedding as Implicit Matrix Factorization. *Advances in Neural Information Processing Systems*, *27*.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*, 281–297.

March, J. G. (1991). Exploration and Exploitation in Organizational Learning. *Organization Science*, *2*(1), 71–87.

McGill, M. E., Slocum, J. W., & Lei, D. (1992). Management practices in learning organizations. *Organizational Dynamics*, *21*(1), 5–17.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. *arXiv:1301.3781 [cs]*.

Mikolov, T., Yih, W.-T., & Zweig, G. (2013). Linguistic regularities in continuous space word representations. *Proceedings of NAACL-HLT*, 746–751.

Miller, D., & Friesen, P. H. (1986). Porter's (1980) Generic Strategies and Performance: An Empirical Examination with American Data: Part I: Testing Porter. *Organization Studies*, *7*(1), 37–55.

Mohr, J. W., & Bogdanov, P. (2013). Introduction—Topic models: What they are and why they matter. *Poetics*, *41*(6), 545–569.

O'Reilly, C. A., & Tushman, M. L. (2008). Ambidexterity as a dynamic capability: Resolving the innovator's dilemma. *Research in Organizational Behavior*, *28*, 185–206.

Padgett, J. F., & Powell, W. W. (2012). *The Emergence of Organizations and Markets*. Princeton University Press.

Pennington, J., Socher, R., & Manning, C. (2014). GloVe: Global Vectors for Word Representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1532–1543.

Poutanen, P., Soliman, W., & Ståhle, P. (2015). The complexity of innovation: An assessment and review of the complexity perspective. *European Journal of Innovation Management*, *19*.

Raisch, S., & Birkinshaw, J. (2008). Organizational Ambidexterity: Antecedents, Outcomes, and Moderators. *Journal of Management*, *34*(3), 375–409.

Rippa, P., Ponsiglione, C., Bocanet, A., Capaldo, G., & Zollo, G. (2019). Do new ventures explore, exploit or both? A case-based analysis of six innovative Italian start-ups. *International Journal of Entrepreneurial Behavior & Research*, *25*(7), 1515-1536.

Shiller, R. J. (2019). *Narrative Economics: How Stories Go Viral and Drive Major Economic Events* (1st edition). Princeton University Press.

Simon, H. (1962). Architecture of Complexity, *Proceedings of the American Philosophical Society*, *106.*

Stanko, M. A., & Henard, D. H. (2017). Toward a better understanding of crowdfunding, openness and the consequences for innovation. *Research Policy*, *46*(4), 784–798.

Stark, D. (2011). *The Sense of Dissonance: Accounts of Worth in Economic Life*. Princeton University Press.

Taylor, A., & Greve, H. R. (2006). Superman or the Fantastic Four? Knowledge Combination and Experience in Innovative Teams. *The Academy of Management Journal*, *49*(4), 723–740.

Smith, A. (1776). The Wealth of Nations. London: W. Strahan and T. Cadell, London.

Tria, F., Loreto, V., Servedio, V. D. P., & Strogatz, S. H. (2014). The dynamics of correlated novelties. *Scientific Reports*, *4*(1), 5890.

Williamson, Ó. (1979). Transaction-Cost Economics: The Governance Of Contractual Relations. *Journal of Law and Economics*, *22*, 233–261.

Yao, Z., Sun, Y., Ding, W., Rao, N., & Xiong, H. (2018). Dynamic Word Embeddings for Evolving Semantic Discovery. *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, 673–681.

Zhang, Y., Jatowt, A., S Bhowmick, S., & Tanaka, K. (2016). The Past is Not a Foreign Country: Detecting Semantically Similar Terms across Time. *IEEE Transactions on Knowledge and Data Engineering*, *28*, 1–1.