KAUFFMAN

The Foundation of Entrepreneurship

KAUFFMAN DISSERTATION EXECUTIVE SUMMARY

Part of the Ewing Marion Kauffman Foundation's Emerging Scholars initiative, the Kauffman Dissertation Fellowship Program recognizes exceptional doctoral students and their universities. The annual program awards up to fifteen Dissertation Fellowship grants of $20,000 each to Ph.D., D.B.A., or other doctoral students at accredited U.S. universities to support dissertations in the area of entrepreneurship.

Since its establishment in 2002, this program has helped to launch world-class scholars into the exciting and emerging field of entrepreneurship research, thus laying a foundation for future scientific advancement. The findings generated by this effort will be translated into knowledge with immediate application for policymakers, educators, service providers, and entrepreneurs as well as high-quality academic research.

Kauffman Dissertation Fellowship Program KDFP Kauffman Dissertation Fellowship Program KDFP

www.kauffman.org/kdfp

ENTREPRENEURIAL DISCOVERY AND INFORMATION COMPLEXITY IN KNOWLEDGE-INTENSIVE INDUSTRIES

Amol M. Joshi
Ph.D. in Business Administration
Strategy & Entrepreneurship Department, Kenan-Flagler Business School
University of North Carolina at Chapel Hill

April 2011

(under the direction of Atul Nerkar)

ABSTRACT

I investigate how the inherent complexity of the information associated with a newly introduced product affects the likelihood that the product is subsequently replicated, imitated, or both. Using information theory, I introduce a model and methods for quantifying the complexity of any product that is representable as an algorithm. I apply this methodology to construct and analyze a historical dataset of 91 digital signal processing firms and 853 product introductions (1974-2009). The empirical results support previously untested predictions from earlier simulation studies. The generalized model is extensible to many knowledge-intensive industries and has important implications for researchers, managers, and policymakers.

Category: entrepreneurship, strategy, management

Keywords: discovery, opportunities, complexity, replication, imitation, algorithm

Acknowledgement:

This research is funded in part by a Dissertation Fellowship grant from the Ewing Marion Kauffman Foundation. The contents of this publication are solely the responsibility of Amol M. Joshi.

Amol M. Joshi, Strategy & Entrepreneurship
Kenan-Flagler Business School
University of North Carolina at Chapel Hill

ENTREPRENEURIAL DISCOVERY AND INFORMATION COMPLEXITY IN KNOWLEDGE-INTENSIVE INDUSTRIES

Research Question:

Why are some firms better able than others to exploit new opportunities?

I posit that differences in the type and level of complexity of the information obtained through the entrepreneurial discovery process may be a meaningful indicator of the likelihood that a firm is able to exploit a new opportunity.

Theoretical Lens:

Specifically, I investigate knowledge reproduction processes for product replication (internal copying) and imitation (external copying) as a means of exploiting opportunities and building competitive advantage. Integrating concepts from information theory and the knowledge-based view (KBV) of the firm, I introduce a generalized model and quantitative methods for estimating the inherent complexity of any unit of knowledge, such as a strategy, technology, product, or service, as long as the unit is represented in algorithm form.

Methodology and Data:

Modeling organizations as information processing systems, I develop measures of the information complexity of an algorithm representing a unit of knowledge in terms of the minimum amount of data (algorithmic complexity) and the minimum number of instructions (computational complexity) required to fully describe and execute the algorithm. I apply this methodology to construct and analyze a unique historical dataset of 91 firms (diversifying and *de novo* entrants) and 853 new product introductions (1974-2009), in a knowledge-intensive industry, digital signal processing.

Key Findings:

(1) information complexity is negatively and significantly related to product replication and imitation;

(2) replicators have the greatest advantage over imitators at moderate levels of information complexity;

(3) intellectual property regimes strengthening the patentability of algorithms significantly increase product replication, without significantly decreasing imitation;

(4) outbound licensing of patented technologies decreases product replication and increases imitation;

(5) products introduced by *de novo* entrants are less likely to be replicated and more likely to be imitated than products introduced by diversifying entrants; and

(6) diversifying entrants have the greatest advantage over *de novo* entrants at high and low levels of information complexity; neither type of entrant has a significant advantage at moderate levels of complexity.

Contributions:

These empirical findings support and extend previously untested propositions from earlier simulation studies and contribute to the literature by expanding our understanding of how different types and levels of information complexity relate to competitive advantage. For researchers, the generalized model of information complexity, replication, and imitation offers new methods for analyzing the inherent complexity of any unit of knowledge. For managers, this model estimates the extent to which changes in product complexity lead to competitive advantage. For policymakers, this model captures the effectiveness of patent policy changes in encouraging knowledge replication or deterring knowledge imitation.

RESEARCH OVERVIEW

Entrepreneurial Discovery

Entrepreneurial discovery is a dynamic competitive process through which market participants obtain *information* about *new opportunities* (Fiet, 1996; Kirzner, 1997; Shane, 2000). As prospective market participants, *de novo entrants* (startup and spinoffs) and *diversifying entrants* (established firms) alike may exploit entrepreneurial opportunities (Companys & McMullen, 2007; Ganco & Agarawal, 2009) by "carrying out new combinations" of technical or managerial innovations (Schumpeter, 1934). *Ex ante* validation of entrepreneurial opportunities is challenging for firms because "the only reliable confirmation that a previously unseen or unknown valuable opportunity does in fact exist occurs when a market has been created for the new item" (Eckhardt & Shane, 2003). Thus, market participation in the form of a new product introduction by a firm may indicate the successful exploitation of the information obtained through the process of entrepreneurial discovery. Previous research (Shane & Venkataraman, 2000) suggest that a firm's ability to engage in entrepreneurial discovery and exploit new opportunities is essentially a function of the *availability of prior information* and the *capacity to process this information* (Galbraith, 1977; Hayek, 1945).

Information Complexity

But what if the prior information that is available to firms is extremely complex, as is the case in many knowledge-intensive industries? Then does the *inherent complexity* of the available information in some way affect the ability of the firm to handle and use it in identifying and pursuing new opportunities? The term *complexity* has a broad definition which is commonly used to describe the extent to which an entity exhibits the characteristics of intricacy, interdependency, complicatedness, multiplicity, etc. (Anderson, 1999; Gell-Mann, 1995; Simon, 1962). In this study, complexity specifically refers to the concept of *information complexity* (Boisot, MacMillan, & Han, 2007; Sorenson, 2005), which is based on information theory (Blum, 1967; Kolmogorov, 1968; Shannon, 1948) and describes the properties of *algorithms*.

What is an Algorithm?

> *Simply put, an algorithm is a set of instructions, a code. Though a seemingly basic concept, it is what made the modern world possible, for without the algorithm, there would be no computer.* (Berlinski, 2001)

An *algorithm* is defined as a set of simple instructions for carrying out a task (Sipser, 1997). If a unit of knowledge such as a strategy, technology, product, or service can be fully specified as a finite and complete set of data and instructions, then it is representable in the form of an algorithm and the information complexity of the algorithm may be estimated (MacKay, 2003). Based on the computational resources required, Gell-Mann (1995) describes two types of information complexity —— algorithmic and computational. As a simple analogy, if we characterize a unit of knowledge as a procedure or a "recipe" (Nelson & Winter, 1982), then the *algorithmic complexity* of this recipe corresponds to the number of words required to fully represent the entire recipe, while its *computational complexity* corresponds to the number of steps required to fully complete the recipe.

Why are Algorithms Important?

> *Algorithms sound scary, of interest only to dome-headed mathematicians. In fact they have become the instruction manuals for a host of routine consumer transactions. Browse for a book on Amazon.com and algorithms generate recommendations for other titles to buy. Buy a copy and they help a logistics firm decide on the best delivery route. Ring to check your order's progress and more algorithms spring into action to determine the quickest connection to and through a call-centre. From analysing credit card transactions to deciding how to stack supermarket shelves, algorithms now underpin a large amount of everyday life.*
>
> Source: "Algorithms - Business by numbers," The Economist, Sept. 13, 2007

Beyond the consumer examples described above, within the knowledge-intensive sector of the economy, there are a number of emerging, high-growth industries which are based primarily on scientific and engineering advances in fundamental algorithms. In addition to digital signal processing (DSP), which is the empirical context for this study, these *algorithm-based industries* include: Internet search engines, the semantic web, bioinformatics and computational genomics, cryptography, encryption, networking, telecommunications, fraud detection, and actuarial sciences. These examples reflect the diversity of segments of the knowledge-based economy where firms pursue innovation in algorithms in order to build competitive advantage.

Knowledge Replication and Imitation

In the previously described examples of algorithm-based industries, firms purposefully use algorithms to represent a unit of knowledge (Boulding, 1966). Replication (internal copying) and imitation (external copying) are two ways to describe the knowledge reproduction activities that take place in firms; “imitation is simply replication performed by a competitor” (Teece & Pisano, 1994). The “paradox of replication” embodies the concept that “efforts by a firm to grow by the replication of its technology enhances the potential for imitation” (Kogut & Zander, 1992). For example, products based on technologies that are easy for a firm to replicate may also be easy for the firm’s competitors to imitate, while products based on technologies that are difficult for a firm to replicate may also be correspondingly difficult for its competitors to imitate (Rivkin, 2000). Thus, to assess competitive advantage, it is essential to understand the relative difference in the ease of replication versus imitation associated with the reproduction of a unit of knowledge (Grant, 1996). The magnitude of the difference between the ease of replication and imitation may be influenced by the strength of the intellectual property (IP) regime (Lerner, 2002), which may be a substantial imitation barrier for firms in knowledge-intensive industries, such as digital signal processing, which is the focus of this study.

Empirical Setting: The Digital Signal Processing (DSP) Industry, 1974-2009

Digital Signal Processors (DSPs) are physical devices or semiconductor "chips" that are used to digitally convert, transform and analyze audio, voice, image, video and other sensory signals from one form to another (Smith, 1997). DSPs are essential components (Eyre, 2000; Frantz & Papamichalis, 1996) used in a variety of consumer applications (cell phones, DVD players, stereos) and industrial solutions (telecommunications network switches, satellite transmission, global positioning). In 2007, worldwide DSP revenues were $27.2 billion, representing about 10% of the total semiconductor industry revenue (Strauss, 2007).

The DSP industry and other algorithm-based industries described earlier all exhibit a unique characteristic —— their product introductions can be mapped to a finite set of algorithms of known or computable complexity (Deka, 1995), which makes the proposed type of information-theoretic analysis possible. Also, the algorithms underlying these product introductions are public knowledge and extensively codified (Nebeker, 1998). The generally low levels of tacitness of product knowledge for all firms in the DSP industry means that the process of reverse engineering competitors' DSP products is a feasible and prevalent practice. This enables me to control for tacitness as a source of competitive advantage and causal ambiguity (Reed & Defillippi, 1990) and focus on the inherent complexity of the products themselves as the primary source of advantage and ambiguity.

I hand-assembled a comprehensive historical dataset of firm-level and product-level market entry data and technical specifications for DSP chips from 1973 to 2009 (from industry inception to the most recent full year in which new product introduction data is available). I obtained the data from seven independent sources including: Berkeley Design Technology, Inc. (BDTI), Electronic Design News (EDN), Forward Concepts (FC), the Institute of Electrical and Electronics Engineers (IEEE) Signal Processing Society, the Smithsonian Institution's Chip Collection as well as information provided by the most active firms such as Analog Devices (AD) and Texas Instruments (TI). Using this extensive archival data, the combined sample was found to have 91 unique firms and a total of 853 new product introductions over a 37-year period. Each firm has at least one product introduction, and each product is assigned to only one firm. There are no jointly produced products in the sample.

In this study, the dependent variables are the subsequent occurrences of replication or imitation of newly introduced products. The independent variables are the information complexity (algorithmic and computational) of each product and the change in IP regime. The control variables capture firm fixed effects, geographic differences, licensing availability, the number of previous entrants and the number of previous times a product was replicated and imitated. Because my primary interest is in establishing the effect of the theoretical variables (algorithmic and computational complexity) at the product-level on the dependent variables, i.e., the likelihood of a product being replicated and imitated, across all of the products in the sample, I used the generalized estimating equations (GEE) method to estimate the logistic regression model. About 77% of the products were replicated by the originating firm and about 45% were imitated by competitors. Thus, some products were both replicated and imitated.

Interpretation of Results and Potential Implications

Using the empirical results obtained from my analyses of the DSP industry, I compiled a list of questions and answers for researchers, managers, and policymakers to use when estimating the likelihood of replication and imitation of products introduced by different firms. This list is intended to serve as a set of rules of thumb for making optimally imperfect decisions based on whether a product is more or less likely to be replicated or imitated (see table below).

Rules of Thumb for Estimating the Likelihood of Product Replication and Imitation

Main Question	**Follow-up Question**	**Answers Based on the Empirical Results of This Study**
Which product, A or B, is more likely to be replicated?	What is the algorithmic complexity of each product?	The product with the *greater* algorithmic complexity is *less* likely to be *replicated*. If the algorithmic complexity is the *same* for both products, then the products are *equally* likely to be *replicated*.
Which product, A or B, is more likely to be imitated?	What is the computational complexity of each product?	The product with the *greater* computational complexity is *less* likely to be *imitated*. If the computational complexity is the *same* for both products, then the products are *equally* likely to be *imitated*.
Should the respective firms that introduced product A and product B, pursue a replicator or an imitator strategy?	What is the overall level of information complexity of each product? LOW, MODERATE, or HIGH	At *moderate* levels of information complexity, replicators have the *greatest* advantage over imitators. Therefore, the firm introducing the product should pursue a *replicator* strategy. At *low* and *high* levels of information complexity, replicators have the *smallest* advantage over imitators. Therefore, the firm introducing the product should pursue *either* strategy without significantly affecting how likely the product is to be replicated or imitated.
How does a change in IP regime affect whether a product more or less likely to be replicated or imitated?	Does the new IP regime support the patentability of algorithms? YES or NO	If the IP regime supports the patentability of algorithms, then the product is *more* likely to *replicated* but *not necessarily less* likely to be *imitated*.
How does technology licensing affect whether a product more or less likely to be replicated or imitated?	Is the product available via outbound licensing? YES or NO	If the product is available via outbound licensing, then the product is *less* likely to be *replicated* and *more* likely to be *imitated*.
How does the type of firm affect whether a product more or less likely to be replicated or imitated?	Is the firm that introduced the product a diversifying entrant (established firm) or a *de novo* entrant (startup or spinoff)?	At *low* and *high* levels of information complexity, a product introduced by a diversifying entrant is *more* likely to be *replicated* and *less* likely to be *imitated* than a product introduced by a *de novo* entrant. At *moderate* levels of information complexity, there is *no significant difference* between diversifying and *de novo* entrants in terms of how likely a product is to be *replicated* or *imitated*.

REFERENCES

Anderson, Philip. (1999). Complexity Theory and Organization Science. *Organization Science, 10*(3), 216-232.

Berlinski, David. (2001). *The advent of the algorithm : the 300-year journey from an idea to the computer*. San Diego: Harcourt.

Blum, M. (1967). A Machine-Independent Theory of the Complexity of Recursive Functions. *Journal of the ACM, 14*(2), 322.

Boisot, Max, MacMillan, Ian C., & Han, Kyeong Seok. (2007). *Explorations in information space : knowledge, agents, and organization*. Oxford; New York: Oxford University Press.

Boulding, Kenneth E. (1966). The Economics of Knowledge and the Knowledge of Economics. *The American Economic Review, 56*(1/2), 1-13.

Companys, Yosem, & McMullen, Jeffery. (2007). Strategic Entrepreneurs at Work: The Nature, Discovery, and Exploitation of Entrepreneurial Opportunities. [10.1007/s11187-006-9034-x]. *Small business economics, 28*(4), 301-322.

Deka, Rabin. (1995). A comprehensive study of digital signal processing devices. [doi: DOI: 10.1016/0141-9331(95)91860-7]. *Microprocessors and Microsystems, 19*(4), 209-221.

Eckhardt, Jonathan T., & Shane, Scott A. (2003). Opportunities and Entrepreneurship. *Journal of management., 29*(3), 333.

Eyre, J. J. (2000). The evolution of DSP processors. *IEEE Signal Processing Magazine, 17*(2), 43-51.

Fiet, James J. O. (1996). The informational basis of entrepreneurial discovery. *Small business economics, 8*(6), 419-430.

Frantz, Gene, & Papamichalis, Panagiotis. (1996). Introduction to DSP Solutions. *Texas Instruments Technical Journal, 13*(2), 5-16.

Galbraith, Jay R. (1977). *Organization design*. Reading, Mass.: Addison-Wesley Pub. Co.

Ganco, M., & Agarawal, R. (2009). Performance Differentials between Diversifying Entrants and Entrepreneurial Start-Ups: a Complexity Approach. *Academy of Management Review, 34*(2), 228-252.

Gell-Mann, M. (1995). What is complexity. *Complexity (New York, N.Y.), 1*(1), 16.

Grant, Robert M. (1996). Prospering in Dynamically-Competitive Environments: Organizational Capability as Knowledge Integration. *Organization Science, 7*(4), 375-387.

Hayek, F. A. (1945). The Use of Knowledge in Society. *The American Economic Review, 35*(4), 519-530.

Kirzner, Israel M. (1997). Entrepreneurial Discovery and the Competitive Market Process: An Austrian Approach. *Journal of Economic Literature, 35*(1), 60-85.

Kogut, Bruce, & Zander, Udo. (1992). Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. *Organization Science, 3*(3, Focused Issue: Management of Technology), 383-397.

Kolmogorov, A. A. N. (1968). Three approaches to the quantitative definition of information. *International journal of computer mathematics, 2*(1), 157-168.

Lerner, Josh. (2002). 150 Years of Patent Protection. *The American Economic Review, 92*(2, Papers and Proceedings of the One Hundred Fourteenth Annual Meeting of the American Economic Association), 221-225.

MacKay, David J. C. (2003). *Information theory, inference, and learning algorithms*. Cambridge, UK; New York: Cambridge University Press.

Nebeker, Frederik. (1998). *Signal processing : the emergence of a discipline, 1948 to 1998*. [New Brunswick, NJ]: IEEE History Center.

Nelson, Richard R., & Winter, Sidney G. (1982). *An evolutionary theory of economic change*. Cambridge, Mass.: Belknap Press of Harvard University Press.

Reed, Richard, & Defillippi, Robert J. (1990). Causal Ambiguity, Barriers to Imitation, and Sustainable Competitive Advantage. *The Academy of Management Review, 15*(1), 88-102.

Rivkin, Jan W. (2000). Imitation of Complex Strategies. *Management science, 46*(6), 824-844.

Schumpeter, Joseph Alois. (1934). *The theory of economic development; an inquiry into profits, capital, credit, interest, and the business cycle*. Cambridge, Mass.: Harvard University Press.

Shane, Scott. (2000). Prior Knowledge and the Discovery of Entrepreneurial Opportunities. *Organization Science, 11*(4), 448-469.

Shane, Scott, & Venkataraman, S. (2000). The Promise of Enterpreneurship as a Field of Research. *The Academy of Management Review, 25*(1), 217-226.

Shannon, Claude Elwood. (1948). *A mathematical theory of communication*. New York: American Telephone and Telegraph Co.

Simon, Herbert A. (1962). The Architecture of Complexity. *Proceedings of the American Philosophical Society, 106*(6), 467-482.

Sipser, Michael. (1997). *Introduction to the theory of computation*. Boston: PWS Pub. Co.

Smith, Steven W. (1997). *The scientist and engineer's guide to digital signal processing*. San Diego, Calif.: California Technical Pub.

Sorenson, O. (2005). Social Networks, Informational Complexity and Industrial Geography. *INTERNATIONAL STUDIES IN ENTREPRENEURSHIP*(6), 79-96.

Strauss, Will. (2007). DSP CHIP STRATEGIES ’07: A Study of IC Markets Driven by Digital Signal Processing Technology: Forward Concepts.

Teece, David, & Pisano, Gary. (1994). The Dynamic Capabilities of Firms: an Introduction. *Industrial and Corporate Change, 3*(3), 537-556. doi: 10.1093/icc/3.3.537-a