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

CAUSES OF EXTREME OUTCOMES IN ENTREPRENEURSHIP: EXPECTATIONS, ENDOWMENTS, ENGAGEMENT, AND ENVIRONMENTS

G. Christopher Crawford, Ph.D.
Department of Management & Entrepreneurship
University of Louisville, 2013

ABSTRACT

What drives the emergence of new firms that produce massive, creative destruction in the market? I answer this question in three essays, using a complexity science perspective and integrating inductive, deductive, and abductive theory-building methods to construct a comprehensive framework of new venture creation and growth. Finding power law distributions for all theoretically relevant input and outcome variables in three studies (N=11,000+), I develop four multi-level constructs—*expectations, endowments, engagement,* and *environments*—that are mutually exclusive, yet collectively exhaustive, to explain and predict how new firms scale up to extreme outcomes. Challenging the assumptions of Gaussian statistics and normal distributions, my framework is empirically generalizable, methodologically rigorous, and theoretically novel. The findings in this dissertation have significant implications for practice and future theory-building efforts in entrepreneurship.

Dissertation Topic Category: Entrepreneurship

Keywords: New Venture Creation, Emergence, Complexity Science, Outliers, Extreme Outcomes, Growth, Power Law Distribution, Agent-based Model, Theory Development, Philosophy of Science.

DISSERTATION OF G. CHRISTOPHER CRAWFORD

CAUSES OF EXTREME OUTCOMES IN ENTREPRENEURSHIP: EXPECTATIONS, ENDOWMENTS, ENGAGEMENT, AND ENVIRONMENTS

EXECTUTIVE SUMMARY

Empirical and Theoretical Context

Most scholars, academics, and policy makers implicitly want to see new ventures create large, positive externalities in the environment. Whether those effects include the increased competitiveness of a nation, the creative destruction of markets, an increase in personal wealth, or the formation of new jobs, high-influence growth is of primary concern in the study of entrepreneurship. Whereas the study of high-impact entrepreneurial organizations is of primary interest in the domain, little is understood about the causal mechanisms underlying the initial creation and growth of such firms. Indeed, senior scholars suggest that entrepreneurial growth is random, and lament the field's lack of a comprehensive theory to explain and predict the extreme outcomes in the domain—firms like Facebook, Google, or Tesla Motors—that are quantitatively and qualitatively different from the "normal" population of ventures. These firms are outliers for both their rarity and their scale of influence on the environment; they significantly skew the distribution of outcomes for the entire population. Examples of these would include a new firm that generates 1000 new jobs, $100M in revenue, or a 40,000% increase in growth. In all extreme cases, a firm must, in some way or another, emerge beyond the normal. The transcendent research question of this dissertation is: *what are the causes of extreme outcomes in entrepreneurship*?

My question is exceptionally important because understanding the antecedents and processes of growth can benefit theory, practice, pedagogy, and policy. However, as noted

throughout the entire existence of entrepreneurship research, the construction of a generalizable and comprehensive theory of growth is very difficult. These difficulties are the result of many factors: most studies of new firm creation suffer from non-randomized selection and inadequate sample sizes; the inherent rarity of high-growth firms makes them problematic to capture outside the confines of a single, survival-biased study; the typical venture is much smaller in size, capital, and scope when compared to established firms; and, for nearly every study that postulates a variable to have some effect on outcomes, another study reports null or negative effects for the same variable. Additionally, explaining or predicting these outcomes with existing theories and traditional statistical methods becomes impossible because of the assumptions underlying each: namely, that inputs and outcomes of entrepreneurial activity are normally distributed across the population. In response, this dissertation draws from both seminal and contemporary theory-building recommendations to offer solutions that empirically address the aforementioned difficulties.

Methodological Suggestions, Empirical Observations, and Theoretical Solutions

Herbert Simon (1968) outlines a methodology for building theory in domains where distributions are heavily skewed, and where findings are conflicting or inconclusive. This methodology is important for constructing generalizable entrepreneurship theory because it uses inductive and deductive procedures that draw from multiple empirical samples. For the first step, Simon recommends classifying generalizable outcomes in the domain, then identifying any "striking empirical regularities" of the distributions. This is consistent with the top-down inductive theory building posited by Shepherd & Sutcliffe (2011), where scholars gather a full gamut of information to "see" a holistic structure of data and to provide the basis for a coherent story of the domain's "truth as is known today."

Essay one reviews outcome variables in extant entrepreneurship research and argues that annual revenue, number of employees, and the growth (relative and absolute) of each can apply to all new ventures, regardless of size. To encompass all historical definitions of entrepreneurship, I analyze three representative datasets at different states of firm emergence: the Panel Study of Entrepreneurial Dynamics (PSED) II, the Kauffman Firm Survey (KFS), and the Inc. 5000 (INC), which sample nascent founder conditions and emergent venture outcomes over five years, a cohort of recently established firms over six years, and a 2012 cross-sectional survey of the fastest growing private U.S. companies, respectively (total N=11,000+). My interest is in the shape of the outcomes' distribution. Analyzing data in MATLAB, I construct semi-parametric bootstrap estimates for maximum likelihood fit to discover a visually distinct consistency: all revenue-, employee-, and growth-based outcome measures are power law distributed *across all twelve years of data*—not one measure is normally distributed.

To be sure, this is a "striking regularity" in the data. Finding these distinct, highly skewed, fat-tailed distributions overwhelmingly supports the notion that extreme events in a population are much more prevalent than traditional statistics would lead us to believe, that outliers in the tail of the distribution disproportionately influence the behavioral and statistical dynamics of the entire system, and, most importantly, that these extreme outcomes are not random. Together, these three points are important to entrepreneurship scholars because linear Gaussian methods that implicitly assume normal distributions are the primary means of testing theory in the domain. Moreover, *one* significant outlier can change substantive conclusions of research findings—including the presence or absence, direction, and size of effect or relationship among variables and constructs (Aguinis et al., 2013). Since about 20% of the total observations in each power law distribution would be considered outliers in traditional statistics, my findings provide an explanation for the consistently inconsistent extant research in entrepreneurship.

Christopher Crawford
Assistant Professor of Strategic Management
Crawforg@Ohio.edu
Ohio University, Athens, Ohio, USA

A more focused research question then becomes: *what components drive the emergence of power laws in entrepreneurship*?

The dissertation's second essay builds off of the second step of Simon's theory-building outline, where he advises scholars to deductively group all the domain's extant variables used to predict outcomes into broad conceptual concepts. As a starting point, I follow Kaplan (1964), who notes "proper concepts are needed to formulate a good theory, but we need a good theory to formulate proper concepts." As such, I draw from three dominant research streams in entrepreneurship—founder psychology, resource-based view, and evolutionary perspective—to encompass what a new venture founder *thinks*, what she *has*, and where she *is*, respectively; in addition, though human agency (i.e., what a founder *does*) is highly discounted in the literature, I include it to reflect my own entrepreneurship experience that depth of interaction with stakeholders can influence outcomes. Then, I link each of these four broad concepts to a theory established in natural, social, and biological systems: complexity science. I draw from this because the goal of complexity science is to understand emergence in its most fundamental form by searching for commonalities that apply to all systems within a population.

A Plausible Story: Complexity Explanations for Power Laws and Extreme Outcomes

Power law distributions are important to entrepreneurship scholars because the theories that explain them are scale-free, where system characteristics, processes, behaviors, and outcomes appear in self-similar (i.e., fractal) patterns across many orders of magnitude. As well, the constructs within a scale-free theory are called scale-invariant because all the variables that form the construct are also power law distributed, regardless of the measurement used. A complexity perspective suggests that emergence and outcomes are governed (i.e., caused) by a simple set of underlying components that apply to all systems in the population. Seminal complexity science

research proposes that outcomes in a complex system are caused by the interactions among four components: 1) schemata—decision-rules of agents that are self-regulated according to an agent's expected outcomes and responses to feedback loops; 2) initial conditions—an agent's resource endowment; 3) population dynamics—local competition and resource munificence, and 4) interactions among the agents in a system. Combining extant entrepreneurship concepts with a cohesive complexity science framework, I outline a framework of four overarching constructs that apply to all ventures, regardless of size: *expectations*, *endowments*, *engagement*, and *environments*. The interaction among these constructs—which are mutually exclusive, yet collectively exhaustive—provide a plausible story for explaining the how PLs emerge in entrepreneurship. As well, the constructs are purposefully a bit fuzzy to be as inclusive as possible of the extant entrepreneurship research findings, while also remaining applicable to and implementable by practitioners, policy makers, and professors.

Given that outcomes are PL distributed, and scale-free theory requires constructs to be scale-invariant, I hypothesize that variables—at the individual-, team-, and venture-level of analysis—within each of these constructs are PL distributed. I test these using PSED data at the first and second year of data collection. I find PLs (with significant Kolmogorov-Smirnov tests) in *expectations* as a founder's expected future revenue, employees, and growth; in *endowments* (including human, social, intellectual, and financial capital variables) as number of employees supervised, number of owners, start-up team experience, team years of education, and previous ventures founded, number of network weak ties and strong ties, and individual net worth, individual funding, team funding, and venture debt; for *engagement*, measured as total activities and total hours worked; and in all *environment* variables, including eleven disparate industry sectors. The framework developed in this dissertation, thus, adheres to the scale-free and scale-invariant requirements of power law theory.

I also test a seminal complexity science proposition of autogenesis: that individual expectations for future outcomes are the "simple rules" that drive aggregate outcomes in a domain. I hypothesize a universal growth dynamic—the scaling exponent (α) of the maximum likelihood estimation—in three different states of emergence: 1) the nascent entrepreneur's expectation for future growth in the PSED, 2) actual PSED outcomes in the fifth year of operations, and 3) outcomes of hyper-growth firms in the INC. I find significant long-range, non-parametric correlations and a universal scaling exponent (~1.75) to support this. This scaling exponent is identical to those found in disparate domains, including the distribution of scientific networks, human interaction dynamics, and the impact of technological breakthroughs, among others. The consistency of these findings hints that the framework may be universal and apply to all emergent social systems.

Essay three builds an abductive agent-based simulation—the first in the domain—to model the proposed relationships among the theory's components, where inputs in the tail of each construct is a necessary and sufficient condition for an extreme outcome. A model like this is a common and effective tool for proof of concept and, most importantly, for future theorizing. Data from the PSED, KFS, and INC establish the model's input parameters, verify its baseline conditions, and validate its emergent outcomes. The model displays a high degree of utility and validity in support of a new Relation Theory of Entrepreneurship, where "success" can be measured on two different scales: the founder's performance in *relation* to his subjective expectations, and the venture's performance in *relation* to its position in the PL distribution.

Implications for Practice, Policy, and Theory

Though power-law distributions have been shown to be ubiquitous in management and strategy research, my dissertation is the first to approach these distributions in depth. My results suggest

that the underlying law in entrepreneurship is the scale-free nature of *growth expectations* of firms at the nascent, new, and hyper-growth levels. The significance at multiple levels and the universality throughout nascent expectations suggests, like Simon (1968), that the findings, "even if approximate, cannot be accidental, but must reveal underlying lawfulness (443)." Overall, this dissertation's findings, methods, and framework have significant implications for practice, policy, and entrepreneurship theory.

For practice, identifying power laws suggests that variance in both inputs and outcomes is nearly unlimited. Thus, resources—and potential for extreme outcomes—are everywhere; however, without sufficient interaction with the environment (e.g., obtaining feedback from potential clients, networking with fund providers), all this potential goes untapped. Almost all resource *endowments* of new venture founders have an exponent of ~2.12, and *engagement* is exactly the same: 2.11; as well, the exponents for *expectations* and all emergent outcomes in the domain are ~1.75. These underlying dynamics suggest that founders may expect extreme outcomes, and that those expectations drive the highest level of performance, but founders are not engaging in a manner that facilitate those outcomes. Instead, the same exponent for *endowments* and *engagement* can be interpreted as founders acting in a manner that is consistent with what they *have*, not what they *envision*. Practitioners would be best served by building up structural endowments prior to forming the venture as a means of believing that superlative outcomes are possible. If founders begin a venture with *endowments*, *expectations*, and *engagement* in the tail of the distribution, they have the potential to push back on selection forces in the environment, rather than be at the mercy of it.

Similarly, my analysis pinpoints the resources that entrepreneurs can most likely leverage to scale up into an extreme outcome. For example, adding endowments like start-up team members with diversely tacit knowledge, weak ties with networking contacts, proprietary

processes, or venture capital funding are disproportionately likely to facilitate growth, vis-à-vis resources that are more easily inimitable. Finally, by identifying which industry sectors are power law distributed, I provide strategic direction for entrepreneurs entering the market with high growth expectations. Where power laws exist, opportunities (i.e., resources) are plentiful and founders should focus on the selection and capture of those that provide the most appropriate fit with the firm's existing knowledge base.

My findings have several implications for public policy. Interventions in the form of tax breaks, business incubators, or grants, are instituted to spur innovation or create new jobs. These interventions increase environmental munificence and, hypothetically, stimulate new entrepreneurial activity. In order to reduce poverty and substantially increase the standard of living, the Kauffman Foundation's VP of research and policy, Robert Litan, contends that the U.S. needs to produce 30 to 60 new companies annually that will eventually reach $1B in revenue (Inc. Magazine, 2012). My power law findings suggest an alternative perspective. Whereas Litan used $1B as "a proxy for firms that generate substantial externalities (pg. 77)," my analyses identify where the tail of the distribution begins (i.e., where externalities are most likely); in the KFS, this is around $2M and, in the INC, this is around $150M. Most founders don't expect to grow and so they don't do the things that lead to growth—like forming teams with diverse knowledge, asking for external funding, or developing proprietary processes; more established companies don't want to grow to $1B because "it would compromise service, quality, and culture (pg. 78)." Thus, the efficacy of interventions may be enhanced by providing incentives for firms to: (1) begin the venture with nonlinear inputs, like a diverse team or a substantial equity investment; and/or (2) to achieve outcomes in the tail of the distribution. In these scenarios, new firms procure enough resources as a foundation for potential growth, and all firms are given aggressive, yet cognitively achievable performance goals.

Resource providers—whether policy makers, venture capitalists, angel investors, or bankers—should also embrace and encourage outliers in their funding decisions. If inputs and outcomes are PL distributed, then funding based on an increase in average portfolio performance would be misguided. Instead, since extreme inputs beget extreme outcomes, resource providers can determine where the tail of the distribution begins for human, social, intellectual, or financial capital and disproportionately allocate funds to outliers.

For theory, these findings can re-direct entrepreneurship research efforts: future theoretical development will need to account for PLs in the data and examine the mechanisms that drive their emergence. Finding PLs in every input and outcome clearly demonstrate that continued use of linear methods lead to inaccurate and misleading substantive conclusions. For example, the nonlinearity of resource endowments provides an explanation for why there is inconclusive evidence about human capital's role in new venture performance and why a theory of growth appears to be so elusive. As well, since we know that PLs exist in engagement, social capital, and financial capital, any study that did not account for these constructs will suffer from unobserved variable bias. Similarly, since engagement in the tail of the distribution can be a necessary and sufficient condition for an extreme outcome, it has the potential to disproportionately compensate for very low levels of human capital. This suggests that even if an entrepreneur does not begin with sufficient endowments, the venture can still achieve superlative outcomes if the founder(s) and employees can work disproportionately harder than competitors.

For theory, this study makes novel methodological contributions to the domain by introducing semi- and non-parametric techniques that more closely mirror the potential nonlinear dynamics inherent in entrepreneurship. In a like manner, the methodologically robust agent-based model—novel to the domain, but called for more than a decade ago by senior scholars—also simulates probabilistic and nonlinear emergence over time. Additionally, my framework

provides an empirical, theoretical, and epistemological first step in outlining a unifying theory of entrepreneurial action. The findings herein embrace and support seminal entrepreneurship frameworks, and encourage scholars to engage in theory-building efforts that account for differences in both variance *and* process.

Attempting to develop a comprehensive theory of entrepreneurial growth is an important endeavor in the pursuit of knowledge and truth about entrepreneurship and the creation of new order. As Albert Einstein (1950) said: "The grand aim of all science is to cover the greatest number of empirical facts by logical deduction from the smallest number of hypotheses or axioms." My one axiom—inputs and outcomes in entrepreneurship are PL distributed—can provide a foundation for fulfilling a grand aim of organization science and facilitate moving the domain beyond what was previously thought possible. I posit that, without the acknowledgement of power law distributions, it is highly unlikely that the field will never be able to construct a unifying theory of entrepreneurship. Without a wholesale a paradigm shift—acknowledging that the norm of normality does not exist in entrepreneurship—the domain will continue to tilt at windmills.

Christopher Crawford
Assistant Professor of Strategic Management
Crawforg@Ohio.edu
Ohio University, Athens, Ohio, USA