SKEWED OPPORTUNITIES: HOW THE DISTRIBUTION OF ENTREPRENEURIAL INPUTS AND OUTCOMES RECONCEPTUALIZES A RESEARCH DOMAIN

ABSTRACT

Over the last four decades, Academy of Management Review has devoted a great deal of attention to the scholarly debate about the theoretical nature of entrepreneurs, entrepreneurship, and entrepreneurial opportunities. Most recently, an entire Dialogue section of the journal was devoted to four articles that provided alternative ontological, epistemological, and philosophical views of "opportunity." Inasmuch as the domain appreciates the effort to advance entrepreneurship theory, these arguments appear to constitute what past AMR editor-in-chief, Roy Suddaby, termed "fetishism," where "theory becomes an exercise in writing and interpretation but is detached from the empirical world" (2014: 408). That reality was demonstrated in the Crawford, Aguinis, Lichtenstein, Davidsson, & McKelvey (2015) study, which discovered highly skewed power law distributions in all of the domain's theoretically relevant input variables and all generalizable outcome measures. The significant number of outliers in these distributions provide necessary and sufficient cause for a paradigm shift in the domain. In response, this paper uses the empirical reality of power law distributed phenomena for 1) developing historical and empirical justification for the difficulties in building theory about opportunities and entrepreneurship, 2) identifying how the seemingly antithetical perspectives of discovery and creation theories can be synthesized, and 3) proposing a generalizable framework—of Endowments, Expectations, Engagement, and Environments—around which new entrepreneurship theory can be developed.

THE LACK OF REPRESENTATIVE REALITY IN ENTREPRENEURSHIP THEORY

Throughout ancient history and in modern academic literature, the bold and daring entrepreneur has taken center stage in the public's eye. Risking personal sacrifice and facing great odds, these individuals of intellect and invention have been known by terms like "Hero", "Adventurer", "Growth Entrepreneur", and "Explorer" (Hebert & Link, 2009). These entrepreneurs are often captains of industry who create massive organizations and societal change, braving uncertainty in previously unseen and unimagined new territories, without regard to their immediate resources at hand. Though they failed more often than not, they aspired—for themselves and for others—to be more than their current circumstances, and to grow as large as their dreams would take them. Thus, high-impact growth has always been a primary measure of entrepreneurial performance.

In contrast, history has not been as kind to the entrepreneurs whose deeds and outcomes are not so grand. Shop owners, and other self-employed individuals who create and maintain very small organizations, were thought to bring marginal benefit to society at large, create nominal wealth, and risk very little. They have been relegated to benign terms like "Lifestyle Entrepreneur" and "Merchants" (Hebert & Link, 2009). Whereas society and academia would like to view entrepreneurship as a Schumpeterian 'creative destruction' activity, scholars have identified that growth and profit maximization are not the sole—or even one of the top four—motivations of the individuals who engage in entrepreneurial action (c.f., Carter, et. al., 2003). Therefore, any comprehensive theory of entrepreneurship and, by extension, opportunities, must include explanation and prediction for outcomes on both ends of the spectrum.

Indeed, the inherent rarity of high-growth ventures, coupled with an overabundance of low-expectation founders, leads to highly skewed distributions of both inputs and outcomes, and has stifled the development of a comprehensive theory for decades (Leitch, Hill, & Neergaard, 2010). As evidenced by extant studies, the development of accumulated knowledge on entrepreneurship and opportunities has been extremely fragmented due to conflicting empirical results (Parker, 2009), inconsistent measures across studies (Shepherd & Wiklund, 2009), cross-sectional studies that exhibit survival bias (Aldrich & Ruef, 2006), and small, selection-biased samples that are of limited generalizability to the full realm of entrepreneurial activity (Aldrich, 2012). For every theoretically relevant input construct—e.g., human capital, growth expectations, financial capital—proposed to have a significant relationship with an outcome, there has been at least one null or opposite finding (Parker, 2009). Moreover, the focal levels of analyses are also fragmented: some theories point to micro-level resource endowments or the cognitive properties of founder(s); other theories point to macro-level conditions; still others point to the seemingly random actions of the entrepreneurs.

This is an important problem for entrepreneurship scholars. The consistently inconsistent empirical findings and the lack of a dominant theoretical framework to explain the full range of antecedents and consequences of entrepreneurial action at multiple levels reduces the field's ability to make claims of causal inference—claims that would be of benefit to practice, policy, pedagogy, and domain legitimacy (Busenitz et al. 2003; Shane 2012). Though these theoretical and methodological difficulties have pervaded the domain of entrepreneurship research since its

inception, new findings may proffer an empirical foundation upon which a comprehensive theory could be built.

> "*Out of clutter, find simplicity. From discord, find harmony.*
> *In the middle of difficulty lies opportunity*." Albert Einstein

In an inductive empirical investigation of four entrepreneurship[1]-related data sets—including the Panel Study of Entrepreneurial Dynamics (PSED) II, Comprehensive Australian Study of Entrepreneurial Emergence (CAUSEE), Kauffman Firm Survey (KFS), and Inc. 5000 (N=12,000+)—Crawford, Aguinis, Lichtenstein, Davidsson, & McKelvey (2015) found power law distributions (PLDs) in all resource-, cognition-, action-, and environment-based input variables, as well as all revenue-, employee-, and growth-based outcome measures. The authors conclude that all theoretically relevant *input*[2] variables—those that make up constructs from theories borrowed from outside entrepreneurship, like the resource-based view, institutional theory, and self-efficacy—and all generalizable *outcome* variables—those that apply to all ventures, regardless of size, including annual revenue and number of employees—are decidedly *not* distributed like a bell-shaped curve. Instead, all inputs and all outcomes in entrepreneurship are power law distributed. In these distributions, high-value outliers disproportionately influence the statistical and behavioral properties of all observations in the system.

In these distributions, the data are highly skewed to the right. Below, Figure 1a depicts a typical PLD when data are plotted on normal scales. Here, the majority of observations (values high on the Event Frequency Y-axis) are at the lowest values of the X-axis, while very few observations (values low on the Y-axis) are out at the far end of the X-axis—making the distribution look like a child's slide on a playground. The same data are shown in Figure 1b, but plotted on logarithmic scales. Of particular note in Figure 1b is the negative slope of the power law tail (to the right of the horizontal dotted line) which, in empirical data, always displays as a straight line. Though the power law distribution's signature tail is so ubiquitous in social systems it has been called "spooky" (Krugman 1994), the Crawford et al. (2015) study was the first to identify their empirical distributions and their importance to the development of theory in the domain of entrepreneurship.

----Insert FIGUREs 1a & 1b about here----

All known perspectives of entrepreneurship share a common set of assumptions—assumptions that underlie virtually all analyses of entrepreneurial emergence and growth, and which pervade economics and the social sciences in general. These have been expressed as the belief in General Linear Reality (Abbott 1988), namely that outcomes in the social world—like the emergence of

[1] The understanding of "entrepreneurial" varies in the literature (cf. Gartner, 1990; Mitchell, 2011). Our conceptual and empirical use of the term captures a broad variety of criteria such as "new entrant' and "founder-managed," as well as "high growth", which is likely to reflect underlying "innovation" along some dimension.

[2] We use the general term "input variables" to encompass items that precede and purportedly explain the outcomes of interest in entrepreneurship; these variables have also been called "antecedent," "explanatory," "independent," "explanantia," "predictor," and "determinant" variables in the literature. Similarly, we use "outcome variables" as a general term to encompass "consequent," "explananda," "criterion," or "dependent" variables.

new companies, or their rapid growth in revenue or employees—can be explained using linear frameworks that rely on normal Gaussian distributions. Linearity assumes that the sum of a system's inputs is directly proportional to its outcomes, and that change is incremental and predictable. Although these assumptions may be appropriate for systems close to equilibrium, they are unable to explain outcomes in emerging systems or environments in a constant state of flux, where nonlinear analysis is more suitable (Meyer et al. 2005). Even though scholars have explicitly warned against linear assumptions when building theory (c.f., Delbridge & Fiss 2013), its pervasiveness is baked into the shared mental framework of the entrepreneurship domain. Thus, while entrepreneurship scholars struggle to explain the emergence and the creation of new order—whether those are ventures, innovations, technological breakthroughs—in conditions that are inherently nonlinear and far-from-equilibrium (McKelvey 2004), there are very few theories and methods that can be used to accomplish this goal.

A similar challenge is found with most quantitative methodology; nearly all the statistical techniques used in the domain rely on assumption of normality and linear relationships as the foundation of testing theoretical hypotheses (Dean et al. 2007). In particular, this requires that samples are normally distributed, where each independent observation can be accurately described by its position relative to some deviation from the mean. This requirement leads to the assumption that outliers—cases which are more than three standard deviations from the mean—should be viewed as random statistical anomalies that must be cleansed from the data to obtain statistical significance. Here again these GLR assumptions do not enable a realistic view of entrepreneurship, for outlier firms—Facebook or Instagram or Google, for example—are often the most interesting and influential cases in a data set. The rare outliers are the ones that change the competitive landscape of an industry and spur continued global innovation.

As O'Boyle & Aguinis (2012) suggest, using Gaussian assumptions and methods on power law-distributed data can lead to incorrect conclusions, misspecified theories, and misleading normative recommendations. Even more importantly, these techniques cannot accurately model emergent, threshold-based phenomena like entrepreneurship (Gimeno, Folta, Cooper, & Woo, 1997), where systems transition from non-existence to existence (Katz & Gartner, 1988; Lichtenstein, Carter, Dooley, & Gartner, 2007). Indeed, when Gaussian statistics are used on outlier-laden data, it changes the "substantive conclusions including the presence or absence, direction, and size of effect or relationship" (Aguinis, Gottfredson, & Joo, 2013: p. 272). Based on all these points, we propose that the existence of power law distributions in entrepreneurship form a justifiable explanation for the domain's conflicting empirical results.

For entrepreneurship scholars, then, the empirical reality of power law-distributed inputs and outcomes present a significant dilemma in the formulation of cumulative theory-based knowledge. First, when power laws exist, outliers are much more prevalent than previously assumed, and even *one* "freak" outlier can change the substantive conclusion of hypothesis tests that use Gaussian methods. To wit, Aguinis & Boyle (2013) identify a Peterson, Smith, Martorana, & Owens (2003) study that found statistical significance for 17 hypothesized relationships among CEO personality, team dynamics, and firm performance; Hollenbeck, DeRue, and Mannor (2006) reanalyzed the data and found that a single data point influenced *16* of those relationships enough to make them *insignificant*. Other domains, like social psychology and sociology, have also had research results publically challenged due the mishandling of outliers (Aguinis, Gottfredson, & Joo, 2013).

Finding these outliers, though, is a good thing for entrepreneurship scholars, as Andriani & McKelvey (2009:16) suggest, "the analysis is…totally meaningless…if the sampling of outliers is insufficient" and "…no statistical finding should be accepted…if it gains significance [where] extreme events…are ignored."

Second, Crawford et al.'s (2015) discovery of power law-distributed inputs and outcomes in two of the domain's most representative samples of emerging entrepreneurs in the nascent stage of venture organizing, the PSED II in the US and CAUSEE in Australia, provides a foundation for challenging nearly every empirical study ever completed. Combined, nearly 200 (Xref Casey Frid) studies, published in the domain's most elite journals, have analyzed PSED-based data. The random-digit-dialing sampling procedure of the PSED was monumental because every individual in a given country had a non-zero probability of inclusion. This provided entrepreneurship scholars with longitudinal data that overcame the significant limitations of previous research designs, including those with data biased by left-truncated survival, cross-sectional observations, convenience sampling, and limited generalizability (Yang & Aldrich, 2009). In those two samples, however, about 17% of all observations are in the tail (i.e., would be considered an outlier in Gaussian statistics). Publicly challenging these findings, however, would only constrain our accumulated knowledge-building efforts, discredit scholars, and reduce the legitimacy of entrepreneurship as a valid research domain. Instead, the findings from the Crawford et al. (2015) study can serve as a foundation for constructing domain-specific theory with assumptions that can distinctly differentiate entrepreneurship from strategic management and other disciplines (Shane, 2012; Shane & Venkataraman, 2000).

Next, we draw from seminal guidance to synthesize a power law framework to subsume alternative explanations about the nature of entrepreneurs, the nature of opportunities, and the nature of the decision-making context. Here, we leverage Boisot & McKelvey's (2010:426) idea that "a power law distribution…reconciles two antagonistic ontologies into one overarching ontology." In the following section, we deduce meta-constructs from the Crawford et al. (2015) study, and arrange them in an epistemologically and ontologically consistent framework that uses these skewed distributions as a foundation to potentially unite the field.

SYNTHESIZING A POWER LAW THEORY OF ENTREPRENEURSHIP

We begin by building off the primary tenets that a theory is a statement of relationships between units observed or approximated in the empirical world (Bacharach, 1989; Dubin, 1969), with the primary function of organizing that world in a manner which comprehensively and parsimoniously explains the phenomena of interest (Whetten, 1989). Given that theory exists, as Abbott (1988:169) claims, "to provide a comprehensible and logically rigorous accounts of facts," the development of theory to understand these phenomena becomes imperative for research, practice, policy, and pedagogy (Sorenson and Stuart 2008). Philosophy of science scholars suggest that the ultimate test of a theory is its verisimilitude, its truth-likeness, in modeling empirical reality. Most importantly, then, the legitimacy, usefulness, and verisimilitude of entrepreneurship as a distinct domain of research hinges on offering "a theoretical framework to explain and predict phenomena that are neither explained nor predicted by other fields" (Shane, 2012: 12x). The

power law distributions in representative samples of nascent ventures provides an empirical foundation upon which such a framework can be constructed.

Regardless of the fetishistic debates about the ontological or epistemological "realness" of opportunities, regardless of whether the domain views the decision-making context as risk or uncertainty, the primary sub-components that comprise opportunities and all ventures are different classifications of resources. To clarify the notion that an opportunity is not equivalent to a business idea, Shane (2012: 15) draws from the Schumpeterian (1934) conceptualization of entrepreneurs exploiting "potentially profitable opportunities by creatively recombining resources." These resources, some of which exist endogenously within entrepreneurs and some of which exist exogenously within different environments, exist as ontologically real resources—whether people, money, or information. The problem is that opportunities are made up of in the environment, whether they are people, money, or information. Making it problematic for entrepreneurship scholars is the fact that these real resources are not normally distributed, nor are they randomly distributed.

A power law-based framework can more accurately explains the dynamics of influential factors on new venture outcomes. Ecology scholars use an evolutionary perspective to study population dynamics—the combination of resource munificence and the competitive density of existing firms (Aldrich 1990). Population dynamics create opposing tensions on new ventures. As one tension, the abundance of perceived resources (e.g., opportunities like quality or quantity of labor available, gross metropolitan product, total innovations) pulls individuals in to the market. Empirical studies show that environmental resources like these are not only highly skewed, they are also highly correlated, with those in the tail of the distribution exhibit the same underlying dynamics. Bettencourt et al. (2010) find that Gross Metropolitan Product (GMP)—the value of all goods and services produced within a given region—is power law distributed in the United States. The size of cities and the income of individuals within each GMP are power law distributed. Those environments in the tail of the distribution have a lot more people available in the labor pool, leading to more total creative people in the labor pool, leading to more patents filed in that city. The authors find that the GMP explains 65%-97% of variance in all of the sub-distributions, and all of these variables have power law exponents of ~2.15. Stating the obvious: there are cities in the tail—like New York and Los Angeles—that have underlying interaction dynamics that are qualitatively and quantitatively different than smaller cities. Thus, resources in the environment, no matter how they are measured, are power law distributed, and those areas in the tail have qualities similar to gravity: they pull in additional resources.

A power law perspective acknowledges and emphasizes uncertainty (Lewin, 1999). In fact, some studies equate nonlinearity and interdependent human interaction with uncertainty (Buckley, 1968) where, as defined in the domain of entrepreneurship, the possible outcomes are unknown and the probabilities for those outcomes are incalculable (Davidsson, 2005; Knight, 1921) and the future is unknown and unknowable (Wiltbank, Dew, Read, & Sarasvathy, 2006). This kind of perspective would be appropriate for Hero-type ventures, like a company that flies passengers to Mars, and would exist in the Paretian world of uncertainty, as shown in figure 2. However, a power law framework can also incorporate risk. Risk should be included in any generalizable framework because it is part of "the truth as known today" in the domain of entrepreneurship research (Shepherd & Sutcliffe, 2011: 363), and because of overwhelming empirical evidence, as shown in

the sheer volume of data points in Figure 2's shaded Gaussian World, the area of insurance-company-assumed, calculable financial risk-assessment (McKelvey & Andriani, 2010). Thus, if a founder opens a street-corner pizzeria (a type of venture with limited variation in both inputs and outcomes), an insurance company could use risk-based econometric techniques, regressing against historical data to assess the probability of the shop's exit from the market and write an insurance policy against it—what Knight (1921:231) calls "statistical probability uncertainty." Thus, a power law perspective encompasses both risk and uncertainty to provide a more generalizable and ontologically agnostic explanation for the antecedents and consequences of entrepreneurial activity.

----Insert FIGURE 2 about here----

This perspective is also consistent with Berglund & Korsgaard's (2017) call for a "social mechanism brand of causal analysis" (732). Indeed, Bar-Yam (1997) posits that power law distributions are the outcome of a deep, underlying pattern of emergence, with significant points of criticality in between. These distributions are important to entrepreneurship scholars because the theoretical constructs 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. Dooley & Van de Ven (1999) proffer that power law distributions are indicators of critical, threshold-based phenomena, which would be consistent with Gimeno, Folta, Cooper, & Woo's (1997) threshold-based view of entrepreneurship. Most interestingly, Simon (1968) and Newman (2005) suggest that it is possible that a "simple set" of mechanisms drive the emergence of PLDs in a domain.

Consistent with this view, using a critical realism perspective that suggests the ontologically actual and ontologically empirical are outcomes and "manifestations of…generative mechanisms," Ramoglou & Tsang (2016) may come the closest to addressing the highly skewed inputs and outcomes of entrepreneurship. Underlying generative mechanisms (e.g., self-organized criticality, preferential attachment) have been purported to cause power law distributions in fields like physics, economics, and behavioral psychology. Though Ramoglou & Tsang (2017) reject mechanism-based explanations under the presumption that "entrepreneurial processes often end up transforming the social systems in which mechanisms are supposedly embedded, thus undermining stability," a power law perspective suggests that emergence and transformation are inherent underlying assumptions of distribution.

TOWARD A RECONCEPTULAIZED RESEARCH DOMAIN

Because of the unique and potentially disproportionate effects that outliers have on other components of the population, developing theory in domains where power laws are known to exist has proven difficult. Simon (1955) was the first to identify this difficulty and further refined a development protocol in 1968. Theory development focused, first, on pursuing generalizable outcomes through empirical investigation and, subsequently, attempting to explain the 'stylized facts' that emerge. Second, Simon (1968) and others (c.f., Anderson, 1999; Andriani & McKelvey, 2009) propose that if domain outcomes are distributed in consistently similar patterns, then it is

likely that similarly skewed sub-distributions exist at multiple levels. Third, these sub-distributions constitute a "simple set" of mechanisms that drive all outcomes of interest—this set is simple because the constructs are readily understood by those who have experience in the domain (whether scholars, practitioners, policy makers, or teachers). These mechanisms apply to all agents in a domain, and the micro-level interaction of these agents drive the subsequent emergence of macro-level outcome patterns. Fourth, to more deeply identify causal relationships among constructs, there will be statistically significant long-range nonlinear correlations between the antecedent constructs and subsequent outcomes (Sornette, 2006). Finally, building theory about how the interactions among antecedents emerge into outcomes requires both rich qualitative data to understand the underlying processes and representative samples of the phenomenon to identify quantitative generalizable outcomes. From there, computational agent-based simulation models are built and multiple experiments are conducted, with model inputs and outcomes validated with large, representative empirical datasets to build more robust theory (Boisot & McKelvey, 2010).

The Crawford et al. (2015) study of multiple stages of venture emergence, encompassing multiple definitions and conceptualizations of entrepreneurship, completes the first and second steps. To follow, we derive a simple set of mechanisms that can explain and predict the emergence of power law distributions (and outliers) in entrepreneurship.

Simple Mechanisms: Endowments, Engagement, Expectations, and Environments

As Andriani and McKelvey (2009) proffer, the (meta)constructs to explain PLD phenomena must be scale-free and apply to multiple levels of analysis to account for the mechanism(s) driving the distributions. Additionally, meta-constructs must be *scale-invariant*, where all lower-level constructs are composed of variables which are PLD. Using a complexity science perspective, where the primary drivers of outcomes consist of 1) an agent's initial conditions, 2) its rules for engagement with other agents based on expectations for future outcomes, and 3) environmental resources (Anderson, 1999), we propose four simple mechanisms that align with seminal entrepreneurship research.

First, we suggest that a venture's initial conditions consist of resource-based constructs analyzed in the Crawford et al. (2015) study: human capital, social capital, intellectual capital, and financial capital—all of which are PLD. These are an individual's, a team's, and a venture's initial *endowments*. Studies from other domains also identify highly skewed distributions of human capital (O'Boyle & Aguinis, 2012), social capital (Barabasi, XXX), intellectual capital (Fleming & Singh XX), and financial capital (XxY). Second, drawing from Bandura's theory of self-efficacy—where the expected level of confidence one has in successfully completing a specific task is caused by one's direct experience with that task—we propose that self-efficacy is a proxy for human experience (i.e., human capital) and suggest that individual *expectations* for growth will be PLD. This was confirmed in the Crawford et al (2015) study. Third, drawing from Song et al. (2010) and González et al. (2008), human action is PLD, as are the number of activities undertaken and the time spent by entrepreneurs in pursuit of creating a new venture (Fleming 2007)—this is the recursive pattern of interaction, this is *engagement*. Finally, Andriani and McKelvey (2009) identify more than 200 empirical studies which depict organizational *environments* as PLD, including corporate growth rates, industry market capitalization, and the size of corporate supply

chains; Crawford et al. (2015) find PLDs in all industry segments. If viewed through a power law lens, entrepreneurial opportunities, consequently, are most readily discovered and created—and ventures are most likely to emerge—when any of the three mechanisms exist beyond some critical threshold; conversely, when the constructs are sub-critical, ventures are more likely to experience "exponentially decaying activity, always dying out" (Sornette 2006:396). In other words, when individuals and teams and ventures have inputs above critical thresholds, extreme outcomes are more likely; when inputs are below thresholds, systems are more likely to be selected out of the system. Using the analogy of a single grain of sand dropped onto the top of a sand pile, self-organized criticality (SOC) suggests that most drops only cause avalanches the size of one or two grains—this represents linear change. As the slope becomes steeper, it reaches a point of criticality, after which each grain has the potential to cause avalanches the size of one or two *hundred* grains—this is nonlinear change.

In all of these systems, activity at the most micro-level aggregates to higher-order activity whereby outcomes at a macro-level are a result of lower-level accumulation (Lewin et al. 1999); the macro-level patterns of outcomes, thus, look similar to micro-level patterns. Therefore, evidence of a power law at one level is an indication that similar dynamics may be acting at the preceding level (Andriani & McKelvey 2009). These patterns are recursive and feedback-driven. Without a significant top-down intervention, like system constraints (e.g., a socialist government) or cross-level negative feedbacks (e.g., suppliers can't deliver product as instructed), expectations for future outcomes drive repeated interactions with other agents in the system, behaviors perpetuate, and micro-level interactions generate emergent order at higher levels that reflect (and are reflected by) the same statistical dynamics across multiple levels (Sornette, 2006; West & Deering, 1995). Emergence, in the classic treatment of complex adaptive systems, occurs when order spontaneously appears at a higher level, without a centralized controller, as a result of the actions and interactions of myriad agents pursuing their own interests at a lower level in a system (Buckley, 1968; Stacey, 1995). In an entrepreneurship sense, a new venture emerges as a synergistic accumulation of resources inherently endowed in some individual, in potential stakeholders, and in existing environments. Subsequently, through aggregation, the entire scope of the phenomenon can be driven by a single process, where an agent recursively interacts with other agents based on their expected outcomes (Drazin & Sandelands 1992).

Expectations are a forecast of an outcome that are based on current information that the predicting party has at the time the prediction is made (Muth 1961). We propose that an entrepreneur's current information is generated from endogenous knowledge and historical context of both current resource endowments (i.e., education, experience) and the availability of exogenous environmental resources. Since human performance is power law distributed (Aguinis & O'Boyle 2013; O'Boyle & Aguinis 2012), it is likely that those in the tail of the distribution will expect consistent levels of high performance. Similarly, environmental resources like city population, innovations within cities, and venture capital availability are also power law distributed (Bettencourt et al. 2010), suggesting that, for example, individuals in California's Silicon Valley or Boston's VC area may have asymmetrical knowledge of potential resources that others in much smaller cities may not have. Together, expectations for future outcomes are likely to be highly skewed because of past experience or environmental munificence. As evidence, Crawford et al. (2015) found nascent expectations for growth to be PLD in two representative samples: The Panel Study of Entrepreneurial Dynamics II and the Comprehensive Australian Study of Entrepreneurial Dynamics.

In a self-organized criticality environment, once a construct emerges beyond some critical threshold, potential interactions are more likely to become nonlinear and cascade to other areas. With nascent entrepreneurial expectations for growth, for example, once expectations are above 25 employees, the entrepreneur cannot simply operate out of her basement, where she could run the business on her existing savings and current cognitive capacity (a linear outcome); instead, because expectations are high, they have a greater probability of intrinsically driving the entrepreneur to interact differently, possibly cascading out to find a larger facility or ask a bank for money or hire experienced managers to run the company (a nonlinear outcome). In a more general sense, in a self-organized criticality environment, once an agent has inputs or outcomes beyond some critical threshold, there is an increased probability of capturing resources compared to an agent that is still below that threshold. As an example, consider the entrepreneur in the previous paragraph. Ceteris paribus, potential stakeholders like property owners, lending officers, or employees on the job market are more likely to share resources when there is higher potential for growth.

A power law perspective posits that small differences in a system's initial conditions, over time and over multiple interactions with the environment, have the potential to emerge into outcomes of infinite variance; the statistical signature of these systems is the power law distribution (McKelvey 2004). Using a power law framework affords the integration of assumptions and conceptual definitions vital to entrepreneurship, namely: heterogeneity, nonlinearity, coevolution, uncertainty, and emergence (Davidsson, 2005; McKelvey, 2004).

Extant studies have proposed indicators of what a power law theory of new order emergence might look like. First, constructs of theoretical interest will be power law distributed (Barabasi 2002)—these will be evident in both inputs and outcomes. Second, as Thietart (2015) explains, the power law slope (as shown in Figure 1b) will be less than 2. This slope, called *alpha* (α), is an indicator of how influential outliers are in the system. As alpha gets closer to 1, it indicates that the outliers in the distribution become more influential on the statistical and behavioural properties of other observations. Here, outliers in the tail make up an increasingly disproportionate total of the entire distribution. An *alpha* above ~2.5 indicates relatively stable environment, where the addition of one outlier will be less likely to destabilize the system. Third, a self-organized criticality environment will exhibit consistent alphas within constructs across time (Andriani & McKelvey 2009). Fourth, as Sornette (2006: 223) suggests, "Studying the correlations and their consequences is an essential part of the analysis of a system, since the correlations are the signatures that inform us about the underlying mechanisms." Though expectations (e.g., cognitive processes), engagement (e.g., total time and interactions working toward a specific goal), and endowments (e.g., resources) have been shown to be causal explanations for behaviors and outcomes in extant research (Aldrich and Ruef 2006) a complexity perspective proposes that the nonlinearity of human interaction renders outcomes as non-deterministic (Rahmandad and Sterman 2008)—where the same action can produce disproportionately positive or negative effects. Additionally, consistent with our complexity argument, we do not suggest the relationship between endowments and outcomes is linear, where more advantageous initial conditions lead to proportionately higher outcomes. Instead, we propose that since observations in the tail of the distribution are nonlinear and, thus, have the potential to lead to both superlative achievement and cataclysmic catastrophe, the relationship between inputs and outcomes will be significant, positive, and nonlinear. The general framework for these relationships is shown in Figure 3.

----Insert FIGURE 3 about here----

From a Darwinian selectionist perspective, the theory with the most utility—that which most accurately describes and reliably predicts—wins. With its robust and growing cadre of methods, complexity science is making a case for the theoretical perspective with the highest utility for describing generative mechanisms, extreme events, and the creation of new order—the changes that drive the evolution of social systems. The simple set of mechanisms we propose is consistent with Simon's (1968: 449) conceptualization:

> "It should be evident that the mechanisms incorporated in the explanatory theory were not motivated by their falsifiability. They were introduced in order to provide "plausible" premises from which the generalization summarizing the observed data could be deduced. And what does "plausible" mean in this context? It means that the assumptions about [power law-distributed inputs and outcomes] are not inconsistent with our everyday general knowledge of these matters. At the moment they are introduced, they are already known (or strongly suspected) to not be far from the truth."
>
> (On Judging the Plausibility of Theories, 1968 :449)

This framework offers empirical evidence that assumptions of normality in entrepreneurship research are untenable more often than not. Accordingly, these empirical results call for an important shift in terms of future entrepreneurship theory and research. Power law distributions suggest that more attention needs to be given to those outliers that make a disproportionate contribution, rather than the more "average" units. For example, 95% of all U.S. businesses are small (employing 20 people or fewer), more than 60% of all new jobs are created by a mere .03% of all entrepreneurial start-ups. These high-influence firms drive innovation in whole sectors of the economy (Shane 2008); they are the ones that change the competitive landscape of an industry, spur continued global innovation, and are the ones that are of most interest from a practice perspective. If entrepreneurship research continues to focus on the mean, it may continue to achieve statistically significant results, but the domain of entrepreneurship is unlikely to make important theoretical progress, and results in academic publications will likely have little value for practitioners. Instead, researchers need to examine the entire distribution of a phenomenon; in particular by focusing on extreme cases, rather than explaining them away as anomalies. Our proposed research agenda—which emphasizes how important it is to know whether a particular distribution is normal or not, and explaining the emergence of power law distributions and the outliers therein—offers clear directions for future study.

While many popular press portrayals of outliers suggest that these individuals constitute an N=1 study (c.f., Gladwell, 2008), the power law-distributed inputs in the Crawford, et al. (2015) study suggest that there are many more outliers than Gaussian statistics would lead us to believe. The proposed simple set of mechanisms of endowments, expectations, engagement, and environments, therefore, represent a unique theoretical framework that is unique to the domain of entrepreneurship. Founded on the empirical reality of power law-distributed inputs and outcomes, with associated underlying assumptions supported by extant research domains, the four mechanisms proposed here satisfy the are empirically tractable, realistic, and fit the purpose at

hand. The mechanisms are empirically derived in a manner that is easily translatable to pedagogy, practice, and policy.

Combining extant research, all mechanisms are scale-free: they are power law distributed no matter what variables are used to measure them at the individual-, team-, and venture-level of analysis. The framework also suggests that outliers and extreme events drive Schumpeterian change by motivating action from the bottom up, creating competitive imperfections, instigating network feedback effects, while recombining and transforming resources from the environment. To follow, we discuss some ways that our framework might influence future theory building in the domain of entrepreneurship.

IMPLICATIONS, FUTURE INTERESTS, AND CONCLUSIONS

The framework here lends insight to research and practicing entrepreneurs, suggesting that ventures with engagement or endowments below critical thresholds of the power law distribution must engage with potential stakeholders at a quality and quantity in the tail of the distribution to substantively scale the venture. This provides a foundation for future causal inference claims in entrepreneurship, where success—for both scholars and practitioners—can be assessed in relation to: 1) the founder's *subjective* expectations for growth, and 2) the venture's *objective* outcomes in the power law distribution. Additionally, future studies can investigate whether individuals or companies with outlier endowments or expectations for growth engage in ways that are qualitatively different than 'normal' agents.

Our framework provides two perspectives on conceptualizing outliers in entrepreneurship. The first is to focus on the mechanisms that drive differences among outliers and non-outliers. A power law is called a scale-free distribution because it looks the same regardless of the scale used to measure it. The power law distribution is the only function satisfying the scale-free criterion (Sornette, 2006; Newman, 2005; Clauset, et al., 2009). Thus, a power law perspective provides entrepreneurship scholars a framework for theorizing at multiple levels. As one example, because of their disproportionately influential outcomes, an individual outlier (i.e., superstar) agent has the potential ability to do things that 'normal' agents cannot: successfully push back on selection forces at higher levels of analysis to accumulate resources. As examples, the framework suggests that, for an individual, superstar performance has the potential to change 'normal' employment policies and create iDeals—idiosyncratic, customized deals/work arrangements that could increase satisfaction for that individual, aimed at keeping her contented and keeping her as an employee (Aguinis and O'Boyle 2014); or, for an organization, superstar performance has the potential to push back on the bargaining power of suppliers (one of Michael Porter's five 'forces') and successfully obtain special treatment not afforded 'normal' organizations (e.g., Walmart demanding that suppliers reduce prices by 10% and stock their own products on a just-in-time schedule). However, if the superstar agent does not *ask* for those special concessions, does not engage potential stakeholders in a manner that increases the probability that additional resources will be acquired, it is unlikely that selection forces will pre-emptively change. In a direct way, the simple act of pushing back on selection forces is outlier engagement—most 'normal' agents simply do not ask. Thus, future research can investigate whether outlier agents engage in ways that are qualitatively different than 'normal' agents.

Recently, Joo, Aguinis, & Bradley (2017) used refined statistical techniques to question some of the human capital findings in the Crawford et al. (2015) study, suggesting that some of the

distributions may not be power laws, but instead, alternative highly right-skewed, heavy-tailed distributions (e.g., log-normal or exponential) In all these alternative distributions, though, outliers still dominate; outliers still drive aggregate system outcomes; outliers still disproportionately influence the statistical and behavioral properties of the population. More interestingly, we propose that the evidence of alternative, non-normal distributions and generative mechanisms suggest that the meta-construct of endowments subsumes these alternative explanations and, thus, holds more theoretical utility than any specific mechanism.

The second perspective on conceptualizing outliers is the importance they play in the understanding of how social systems progress. In most research domains, outliers are often viewed as "freak" observations, data points far outside the normal that skew the behavioral and statistical properties of the system under study. They are often viewed as "mistakes," substantively different than the whole of the population, and consequently, do not play a role in the explanation of the phenomena under study (Rasmussen, 1988). Troublingly, even with the many non-parametric techniques available, coupled with the ubiquitous fact that outliers and power law distributions have been discussed for nearly a century (c.f., Jeffreys & Wrinch, 1921 or Simon, 1955), scholars continue to simply delete outliers from their analysis as a means of gaining statistical significance and publication (see, for example, Chakrabarti & Mitchell's (2016: p. 679) *Strategic Management Journal* article or Pfeffer & Carney's (2017: p. 8) *Academy of Management Discoveries* article). In the current state of social dissention, where the "1%ers" are persecuted by the other 99% of the population, it is important to note that—in entrepreneurship, most especially—these outliers are the ones who create and innovate and push our society forward. Undeniably, outliers *are* different: in this analysis, we identify that outliers *think* differently and *act* differently and *grow* differently than the rest of the population. Our research demonstrates that, in many ways, outliers enrich entrepreneurship theory by helping us understand phenomena like opportunity identification, evaluation, and exploitation in a deeper, more nuanced manner. Indeed, outliers are vital in providing scholars with a complete understanding the entrepreneurial process. Indeed, we need more freaks.

As the development of a comprehensive theory of entrepreneurial growth moves forward, we evoke Albert Einstein's quote, "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" (*Life Magazine*, January, 1950). The primary assumption underlying entrepreneurship, the domain of new order creation—that nearly all inputs and outcomes in entrepreneurship are PL distributed—can help fulfill the grand aim of management research and facilitate moving the field beyond what was previously thought possible. In contrast, without the acknowledgement of power law distributions, the entrepreneurship discipline may never be able to construct a theory that applies to the creation and emergence of *all* ventures—such a theory would be important for the advancement of domain legitimacy, as well as significant for pedagogy, policy, and practice. Without acknowledging, explaining, and understanding power laws, outliers will continue to wreak havoc on both theory and method. In sum, the field's reliance on the 'average' may hinder theory development for all stakeholders. We hope our work will serve as a catalyst for a fruitful research agenda that is likely to have important implications throughout the domain.

REFERENCES

Abbott, A. 1988. Transcending general linear reality. ***Sociological Theory***, 6: 169-186.

Aguinis, H., Gottfredson, R. K., Joo, H. 2013. Best-practice recommendations for defining, identifying, and handling outliers. ***Organizational Research Methods***, 16: 270-301.

Aguinis, H. & O'Boyle, E.H.. 2014. Star Performers in Twenty-First-Century Organizations. ***Personnel Psychology***, 67: 313-350.

Aldrich, H. E., & Ruef, M. 2006. ***Organizations Evolving***. Thousand Oaks: Sage Publications.

Alvarez, S. A., Barney, J. B., McBride, R. & Wuebker, R. 2017. On opportunities: Philosophical and empirical implications. ***Academy of Management Review***, 42: 726-730.

Anderson, P. 1999. Complexity theory and organization science. ***Organization Science,*** 10: 216-232.

Andriani, P., & McKelvey, B. 2009 From Gaussian to Paretian Thinking: Causes and Implications of Power Laws in Organizations. ***Organization Science,*** 20: 1053-1071.

Bacharach, S. B. 1989 Organizational Theories: Some Criteria for Evaluation. ***Academy of Management Review***, 14: 496-515.

Bar-Yam, Y. (1997) ***Dynamics of Complex Systems***. Addison-Wesley, Reading, MA.

Berglund, H. & Korsgaard, S. 2017. Opportunities, time, and mechanisms in entrepreneurship: On the practical irrelevance of propensities. ***Academy of Management Review***, 42: 730-733.

Bettencourt, L.M, Lobo, J. Strumsky, D. & West, G. 2010. Urban Scaling and Its Deviations: Revealing the Structure of Wealth, Innovation, and Crime across Cities. ***PLoS ONE*** 5: 1-9.

Boisot, M. & McKelvey, B. 2010. Integrating modernist and postmodernist perspectives on organizations: A complexity science bridge. ***Academy of Management Review***, 35: 415-433.

Buckley, W. 1968. Society as a Complex Adaptive System. In W. Buckley (Ed.), ***Modern Systems Research for the Behavioral Scientist***, Vol. 10: 490-516. Chicago, IL: Institute for the Study of Coherence & Emergence.

Busenitz, L. W., West III, G. P., Shepherd, D., Nelson, T., Chandler, G. N., & Zacharakis, A. 2003. Entrepreneurship research in emergence: Past trends and future directions. ***Journal of Management***, 29: 285-308.

Chakrabarti, A., & Mitchell, W. 2016 The role of geographic distance in completing related acquisitions: Evidence from US chemical manufacturers. ***Strategic Management Journal***, 37: 673-694.

Colquitt, J. A. & Zapata-Phelan, C. P. 2007. Trends in theory building and theory testing: A five-decade study of the *Academy of Management Journal*. ***Academy of Management Journal***, 50: 1281-1303.

Crawford, G. C., Aguinis, H. A., Lichtenstein, B., Davidsson, P., & McKelvey, B. 2015 Power law distributions in entrepreneurship: Implications for theory and research. ***Journal of Business Venturing***, 30: 696-713.

Crawford, G. C., Dimov, D., & McKelvey, B. 2016. Realism, empiricism, and fetishism in the study of entrepreneurship. ***Journal of Management Inquiry***, 25: 168-170.

Davidsson, P. 2005. ***Researching Entrepreneurship***. New York, NY: Springer.

Dean, M. A., Shook, C. L., & Payne, G. T. 2007. The past, present, and future of entrepreneurship research: Data analytic trends and training. ***Entrepreneurship Theory and Practice***, 31: 601-618.

Delbridge, R., Fiss, P. 2013. Styles of theorizing and the social organization of knowledge. ***Academy of Management Review***, 38: 325-331.

Dooley, K. J. & Van de Ven, A. H. 1999. Explaining complex organizational dynamics. ***Organization Science***, 10: 358-372.

Drazin, R., & Sandelands, L. 1992. Autogenesis: A perspective on the process of organizing. ***Organization Science****, 3:* 230-249.

Dubin, R. 1969. ***Theory Building***. Toronto, Ontario: Collier-Macmillan Canada, Ltd.

Foss, N. J. & Klein, P. G. 2017. Entrepreneurial discovery or creation? In search of the middle ground. ***Academy of Management Review***, 42: 733-736.

Gladwell, M. 2008. **Outliers: The Story of Success**. Little, Brown and Company, New York.

Gimeno, J., Folta, T. B., Cooper, A. C. & Woo, C. Y. 1997. Survival of the fittest? Entrepreneurial human capital and the persistence of underperforming firms. ***Administrative Science Quarterly***, 42: 750-783.

Hebert, R. F. & Link, A. N. 2009. ***A History of Entrepreneurship***. New York, NY: Routledge

Jeffreys, H. & Wrinch, D. 1921. On Certain Fundamental Principles of Scientifique Inquiry. ***Philosophy Magazine***, 42: 369-390.

Joo, H., Aguinis, H., & Bradley, K. J. 2017. Not All Non-normal Distributions Are Created Equal: Improved Theoretical and Measurement Precision. ***Journal of Applied Psychology***. http://dx.doi.org/10.1037/apl0000214

Katz, J. & Gartner, W. B. 1988. Properties of Emerging Organizations. ***Academy of Management Review***, 13: 429-441.

Leitch, C., Hill, F., & Neergaard, H. 2010. Entrepreneurial business growth and the quest for a "comprehensive theory": Tilting at windmills? ***Entrepreneurship Theory and Practice***, 34: 249-260.

Lichtenstein, B., Carter, N., Dooley, K., & Gartner, W. 2007. Complexity dynamics of nascent entrepreneurship. ***Journal of Business Venturing***, 22: 236-261.

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica, 29*(3), 315-335Newman, M. E. J. 2005. Power laws, Pareto distributions, and Zipf's law. ***Contemporary Physics***, 46: 323-351.

O'Boyle, E. H. & Aguinis, H. 2012. The best and the rest: Revisiting the norm of normality of individual performance. ***Personnel Psychology***, 65: 79-119.

Parker, S. 2009. ***The Economics of Entrepreneurship***. New York, NY: Cambridge University Press.

Ramoglou, S. & Tsang, E. W. K. 2017. In defense of common sense in entrepreneurship theory: Beyond philosophical extremities and linguistic abuses. ***Academy of Management Review***, 42: 736-744.

Ramoglou S., & Tsang E. W. K. 2016. A realist perspective of entrepreneurship: Opportunities as propensities. ***Academy of Management Review***, 41: 410–434.

Rousseau, D. 2006. Presidential address: Is there such a thing as "evidence-based management"? ***Academy of Management Review***, 31: 256-269.

Sarasvathy, S. D. 2001. Causation and effectuation: Toward a theoretical shift from economic inevitability to entrepreneurial contingency. ***The Academy of Management Review***, 26: 243-263.

Schumpeter, J. (1934) *The Theory of Economic Development*. Oxford: Oxford University Press.Shane, S. 2008. ***The illusions of entrepreneurship***.

Shane, S. 2012 Reflections on the 2010 AMR Decade Award: Delivering on the promise of entrepreneurship as a field of research. ***Academy of Management Review,*** 37: 10–20.

Shane, S. & Venkataraman, S. 2000. The Promise of Entrepreneurship as a Field of Research. ***Academy of Management Review***, 25: 217-226.

Shepherd, D. & Wiklund, J. 2009 Are we comparing apples with apples or apples with oranges? Appropriateness of knowledge accumulation across growth studies. ***Entrepreneurship: Theory and Practice,*** 33: 105-123.

Simon, H. A. 1955. On a class of skew distribution functions. ***Biometrika*** 42: 425-440.

Simon, H. A. 1968. On judging the plausibility of theories. v. Rootselaar, Staal, eds. ***Logic, Methodology and Philosophy of Sciences III***. North-Holland, Amsterdam.

Sornette, D. 2006. ***Critical Phenomena in Natural Sciences: Chaos, Fractals, Self-Organization, and Disorder: Concepts and Tools***. Springer-Verlag Berlin, Heidelberg.

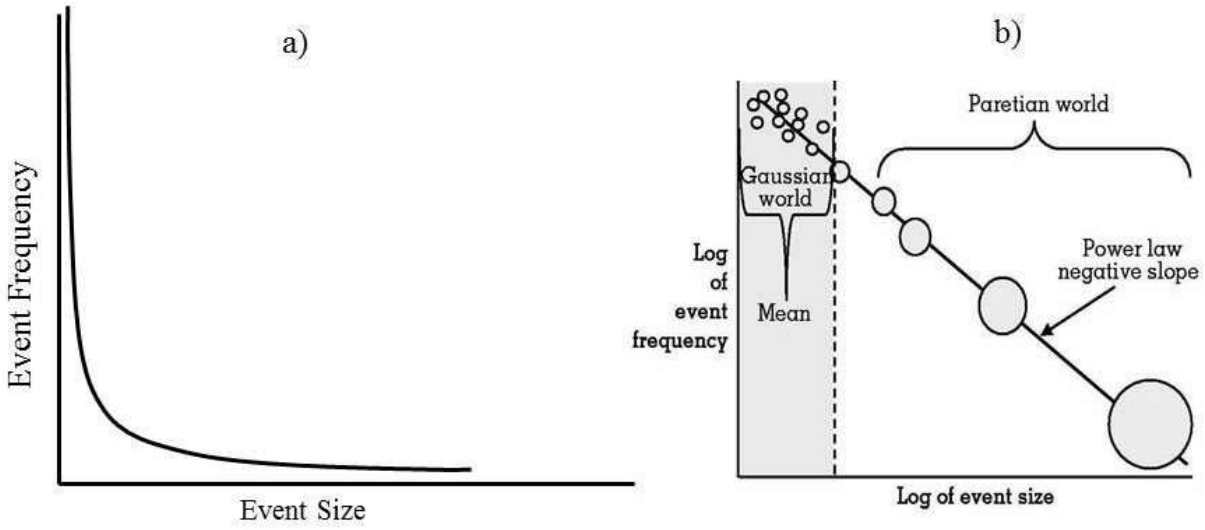
Stacey, R. D. 1995. The Science of Complexity: An Alternative Perspective for Strategic Change Processes. ***Strategic Management Journal***,16: 477-495.

Suddaby, R. 2014. Editor's comments: Why theory? ***Academy of Management Review***, 39: 407-411.

Whetten, D. A. 1989. What Constitutes a Theoretical Contribution? ***Academy of Management Review,*** 14: 490-495.

FIGURE 1

a) Power Law Distribution on Linear Scales; b) Power Law Distribution Plotted on Log-Log Scales



From Boisot and McKelvey (2010) Integrating Modernist and Postmodernist Perspectives on Organizations: A Complexity Science Bridge, *Academy of Management Review*, 35, p. 417.

FIGURE 2
A Power Law Distribution Synthesis of Seminal Entrepreneurship Antimonies

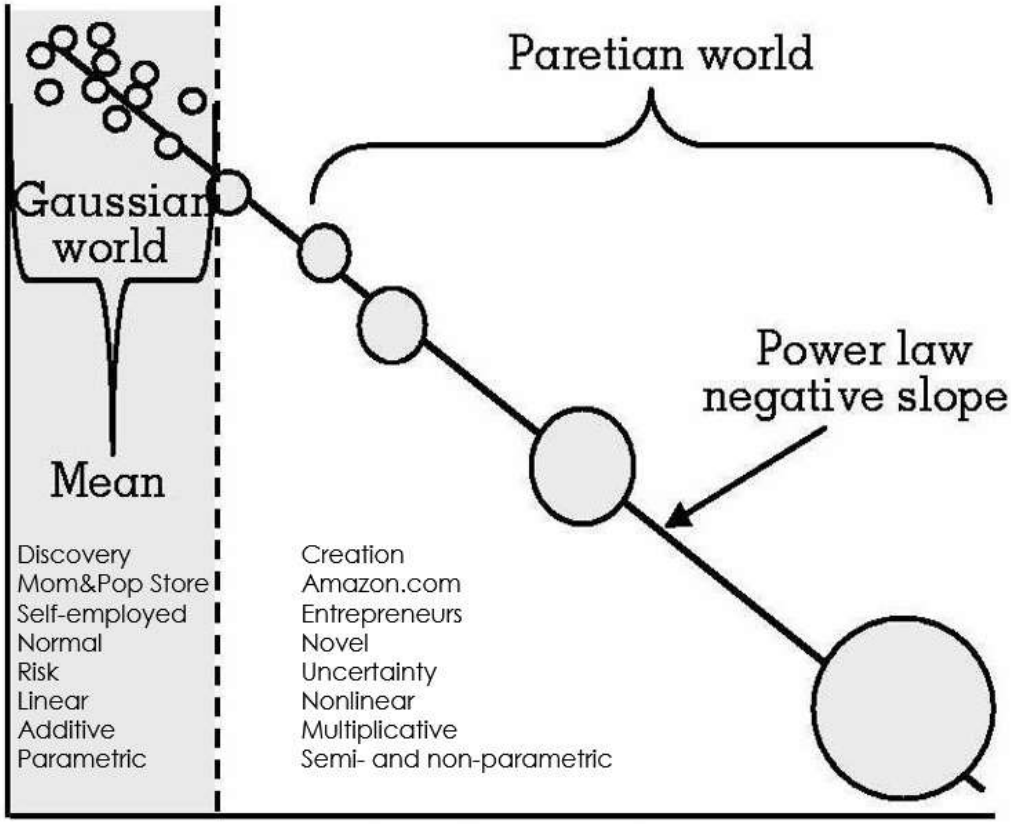


FIGURE 3
A Power Law Framework for Entrepreneurship:
Causal Influences on Outcomes at Multiple Levels