STARTUPS IN A SPIKY WORLD: THREE ESSAYS ON GEOGRAPHY AND ENTREPRENEURSHIP

By

SIDDHARTH VEDULA

B.Sc. Hons., University of Toronto, 2005

M. Eng., McGill University, 2008

A dissertation submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Doctor of Philosophy

Department of Management & Entrepreneurship

2015

This dissertation entitled:
Startups in a spiky world: Three essays on geography and entrepreneurship
written by Siddharth Vedula
has been approved for the Department of Management & Entrepreneurship

____
Sharon F. Matusik, Ph.D.

____
Jeffrey G. York, Ph.D.

Date_____________

The final copy of this dissertation has been examined by the signatories, and we
Find that both the content and the form meet acceptable presentation standards
Of scholarly work in the above mentioned discipline.

ABSTRACT

Vedula, Siddharth (Ph.D., Management & Entrepreneurship)

Startups in a spiky world: Three essays on geography and entrepreneurship

Thesis directed by Professor Sharon F Matusik

Entrepreneurship is a key driver of prosperity within the U.S. economy. It is also a geographically uneven phenomenon; while some regions have cultivated successful entrepreneurial ecosystems others have lagged behind. This regional divergence has important strategic implications for startup firms, in terms of where they choose to locate as well as their subsequent performance. In this dissertation, I therefore use a three-paper model to examine how geographic factors impact a) regional startup formation rates b) startup performance and c) startup decision making processes.

In the first paper I combine theoretical perspectives from sociological institutional theory and knowledge economics. Using the context of the cleantech industry sector, I study how community environmental ideologies moderate (i.e. structure) knowledge spillovers thus shaping regional entrepreneurship rates. The principal finding is that community ideologies are particularly important when the regional knowledge base is less specialized in cleantech. More broadly, I demonstrate that exogenous (i.e. beyond industry boundaries) sociocultural institutions dynamically moderate endogenous industrial dynamics. In the second paper, using data on venture-backed startups and a variance decomposition methodology, I examine the *relative* (i.e. in comparison to firm and industry-specific factors) impact that regions have on differences in startup valuation. I find that, on average, regions explain 2-3 percent of the variance in startup

valuation. However, for startups at the earliest stages of development and for those operating in emerging industries, regions explain 14 and 7 percent of the variance in valuation respectively. The main implication is that regions play a *supporting role* in terms of contributing to startup performance differentials. In the third paper I look at how one specific regional factor, geographic clustering, influences startup decision making processes relating to business closure. The novel finding from this study is that the high rates of startup failure within geographic clusters can be explained, at least in part, by performance premium effects. For instance, startups within clusters are subjectively less satisfied and more likely to terminate operations for comparable, objective levels of performance.

DEDICATION

To my family

To Lucy. Thank you for your infinite love, support, encouragement, and patience throughout this journey. I could not have done this without you, my beautiful wife. You are the love of my life, the reason that I live with a smile on my face, and truly the Khaleesi of the Great Pasture of Willow Creek Drive.

To my sister Indu, who always “told me so”. I am proud to be your little brother. Thank you for always looking out for me. I know you did not think you would see the day, but I’m finally done being a student! And to your sidekick P- thank you for being an awesome brother-in-law. It has been fantastic to talk to an *actual* entrepreneur to keep my hare-brained ideas in check!

To my grandparents. I miss you dearly and wish you could be here. Thank you for your blessings, love, and wisdom. I hope that I have made you proud.

To Bruce and Jackie. I am so fortunate to be a part of your lives. Thank you for accepting me wholeheartedly into the clan McAllister! I am truly blessed to have found such a wonderful home away from home.

And last but not least, to amma and daddy. You are the most wonderful and supportive parents a child could ever wish for. Thank you for your love, infinite patience, blessings, and guidance. Thank you for encouraging my curiosity and answering the unending stream of questions through my childhood. You taught me to value knowledge and pursue it without fearing failure. I am forever grateful.

ACKNOWLEDGEMENTS

This dissertation has been a wonderful and eventful journey. I can still recall the day that I began this program, not entirely sure what I was getting myself into! I am immensely grateful to everyone that has helped me grow both as a scholar and as a person over the past five years.

I would first like to express my sincere gratitude to my dissertation committee (Bret Fund, Tony Tong, Jeff York, and Michael Lenox) and in particular to my chair, Sharon Matusik. Thank you for being such fantastic mentors. I am extremely fortunate to have had the chance to learn from all of you. Bret, you have been a wonderful friend and confidante. Thank you for giving me the opportunity to learn the ropes. Jeff, I am lucky to have had the chance to constantly harass you from down the hall, although I am sure that you feel that the benefits are asymmetric! Your passion and commitment to scholarly work, not to mention your special "you don't know Dixie-isms", is inspiring. And Sharon, you are *the role model* for faculty. I am so proud to have been your student. Thank you for your encouragement, guidance, and willingness to put up with all my random thoughts. I am also forever grateful for your help in keeping things in perspective during the hard times that I went through, and remembering that there is more to life beyond AMJs and SMJs ☺

I would also like to acknowledge my doctoral program cohort: Jenni Dinger, Michael Conger, Carla Bustamante, and Richard Hunt. Thanks for making this fun! I have learned a lot from all of you and am fortunate to have shared this journey with you. Thanks to Jenni in particular for being my official note taker and cheerleader. I'm absolutely thrilled that we are both going to continue our academic careers in the land of stretched out A's…

Lastly, I would also like to thank a number of people at both Leeds and the Deming center outside the M&E department. Jenni Dittenhofer was a godsend to organizationally challenged graduate students such as myself. My friends, Jody Reale and Patty Graff at the Deming Center were always up for an afternoon cuppa or more…their energy was absolutely contagious. Thomas Vossen and Kishen Iyengar similarly listened to me on days where I wanted to crawl into a hole and for that I thank you both immensely. Markus Fitza showed me how to show that things *matter*, but only sometimes. David Balkin made learning fun. Frank Moyes was extremely generous with his time and a savior during my first shock-and-awe experience in the classroom. David Allen gave me the opportunity to gain some valuable experience at the CU Technology transfer office. Alicia Robb introduced me to the good people at the Kauffman foundation, and I look forward to working with them in the years to come. And lastly, Ryan, Brent, and Andy at LTS allowed me to bug them even on weekends with incredibly tedious research server related issues. Thank you all for going out of your way to make my time here that much more special and memorable!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Entrepreneurship is as a critical ingredient of the U.S. economy, driving innovation, economic growth, and overall prosperity (Acs and Szerb, 2007; Audretsch, 2007; Wennekers and Thurik, 1999). It is also a distinctly geographically uneven process (Plummer and Pe’er, 2010; Stam, 2010). While some regions, most notably the Silicon Valley, have achieved notoriety as hot beds of startup activity, others have consistently lagged behind (Goetz and Freshwater, 2001; Guzman and Stern, 2015). This has happened despite the best efforts of policy makers to intervene and engineer entrepreneurial ecosystems around the country (Gilson, 2003; Lerner, 2010). Why does this occur? Why has there not been more regional convergence in entrepreneurial activity in the supposedly “flat world” that we live in (Friedman, 2005)? And in the absence of such convergence, what are the strategic implications for startup firms in terms of where they choose to locate and their subsequent performance? As we increasingly commit ourselves pedagogically (Katz, 2003; Kuratko, 2005) philosophically (Baumol, 1996; Baumol, Litan, and Schramm, 2007), and materially (Audretsch and Thurik, 2001; Obama, 2010) to a startup-based economy, these are undoubtedly important questions to answer.

Not surprisingly, the practical relevance of this topic has meant that there has been a significant body of research on spatial issues in entrepreneurship (see Plummer and Pe’er, 2010 for a detailed review). Somewhat surprisingly however, the majority of research studies have almost exclusively focused on a single question, explaining *regional variations in entrepreneurial activity* (Buhr and Owen-Smith, 2010). Scholars have utilized a range of theoretical perspectives to study this issue, drawing on foundational ideas in different academic disciplines. For instance, economic geographers have focused on the role of location-based, agglomeration externalities in attracting entrepreneurs to regions (Delgado, Porter, and Stern, 2010), sociologists have highlighted the important role of the regional institutional environment in motivating entrepreneurial action (Saxenian, 1996; Tolbert, David, and Sine, 2011), and strategy scholars have focused on the evolution of regional industries through endogenous entrepreneurial dynamics such as spinoff firms (Agarwal, Audretsch, and Sarkar, 2007; Agarwal and Braguinsky, 2014; Klepper, 2010). This body of work has led to a number of useful insights. However, perhaps as a result of the diverse disciplinary foci, there has also been relatively little integration and combination of different perspectives. As a result, our current understanding of the underlying drivers behind regional differences in entrepreneurial activity is somewhat siloed.

In paper 1 of this dissertation I attempt to bridge this gap across disciplinary boundaries. In particular, I develop a model that combines perspectives from both knowledge spillover (Audretsch and Keilbach, 2007) and sociocultural institutional (Tolbert *et al.*, 2011) theories on the geography of entrepreneurship. I do so by studying the interactive influences of knowledge spillovers due to regional innovation (Audretsch and Feldman, 1996), and community sociopolitical ideologies (Simons and Ingram, 2004) on rates of new venture creation within urban metropolitan regions. My core argument is that community sociopolitical ideologies, by shaping both entrepreneurial cognitions (Baron, 2004) and the normative legitimacy of industry sectors (Aldrich and Fiol, 1994), should enhance opportunity recognition and entrepreneurial motivation. Thus ideologies should moderate the extent to which entrepreneurs take advantage of latent opportunities, which arise due to the innovation (i.e. R&D) related activities of incumbents (Agarwal *et al.*, 2007; Audretsch and Keilbach, 2007). Using the emerging cleantech industry sector as a research context, I find that new venture creation rates are higher in MSAs with more incumbent innovation. I also find that the strength of this baseline relationship (Acs *et al.*, 2009; Audretsch and Keilbach, 2008; Plummer and Acs, 2014) is moderated by the strength of sociopolitical ideologies in the urban region where innovation occurs. In particular, knowledge spillovers in the cleantech context are more likely to translate into entrepreneurial action in communities with stronger environmental ideologies. Perhaps the most interesting finding from this study however is that the moderating impacts of ideologies are dynamic and finite, such that they become increasingly insignificant as the knowledge base in a region becomes more specialized in cleantech.

The main theoretical contribution of this paper is to demonstrate that theories of regional entrepreneurship based on the economics of knowledge and industrial evolution (Agarwal *et al.*, 2007; Agarwal and Braguinsky, 2014; Audretsch and Keilbach, 2007; Klepper, 2010) can be enriched by explicitly accounting for the societal context within which industries operate. In effect, by treating entrepreneurship as a largely endogenous market-driven process, I suggest that the extant literature in this area has not paid sufficient attention to the structural role that exogenous, social (i.e. non-market) forces play in shaping economic activities (Baron, 1995; Granovetter, 2005). The evidence in this paper also suggest that exogenous social forces, such as community ideologies, are particularly important in shaping knowledge-spillover based entrepreneurial activity (Agarwal *et al.*, 2007) when the regional knowledge base is less specialized in a relevant technological domain. This is an important insight highlighting the fact that the impacts of supportive institutions are not static, and that their relative relevance changes as the technological base within regional industries evolve (Nelson, 1994), and specific technologies get increasingly adopted within markets. It also indicates that the degree to which incumbents and startups firms share a symbiotic relationship (Baumol, 2002; Hockerts and Wüstenhagen, 2010) as emphasized by knowledge spillover theories of entrepreneurship (Agarwal *et al.*, 2007) varies across different regional contexts. My findings also have useful policy implications. For instance, they suggest that when the regional knowledge base is less specialized in a technological domain of interest to policy makers, they should more explicitly focus their attention on either directly or indirectly shaping community attitudes and rhetoric (Walker *et al.*, 2010) and aligning them with the benefits of the technological solutions that they seek to implement (Cooke, 2008). They might be able to achieve this, for example, by supporting the activities of grassroots organizations that influence public opinion such as technology-focused social movements (Pacheco, York, and Hargrave, 2014).

The second and third papers in this dissertation shift the level of analysis down from the region to the firm (i.e. startup) level. The focus among scholars on understanding regional differences in entrepreneurial entry has meant that we know comparatively less about the strategic implications of geography (Sorenson and Baum, 2003) for startup firms. For instance, there is only a limited body of research on how regional factors such as clusters and institutions impact dimensions of startup performance such as innovation (Folta, Cooper, and Baik, 2006a), growth (Gilbert, McDougall, and Audretsch, 2008), and failure (Folta *et al.*, 2006a; Pe'er and Keil, 2013). Furthermore, there is even less research on how regional factors shape the strategic decision making processes of startup firms, such as the markets that they choose to operate within (Aharonson, Baum, and Feldman, 2007; Kalnins and Chung, 2004) and the performance goals that they set (DeTienne, Shepherd, and De Castro, 2008; Gimeno *et al.*, 1997).

In paper 2, I examine the generalizability of the limited number of case-based (Saxenian, 1996) and single-industry studies (Gilbert *et al.*, 2008) that have examined the impact of regional factors on the performance of venture-backed startup firms. In particular, I examine the degree to which valuation differences in venture-backed startups can be explained by their geographic location choices. The principal purpose of this empirical study is therefore to quantify the *degree* *to which* regions matter for startup performance and generate useful stylized facts (Hambrick, 2007; Helfat, 2007). To do so, I follow and extend a growing body of research that has adapted the use of variance decomposition approaches (McGahan and Porter, 2002; Rumelt, 1991) to the entrepreneurship context (Castellaneta and Gottschalg, 2014; Fitza, Matusik, and Mosakowski, 2009; Short *et al.*, 2009).

I find some interesting and surprising results. For instance, *relative* to other potential sources of competitive heterogeneity (e.g. founding team capabilities, VC investors, industry of operation), regional factors do not seem to matter a great deal. In fact, across the entire sample, regions only explain 2.13 percent of the variance in startup performance. In contrast, startup-specific factors explain 30.34 percent, and VC investors explain 19.88 percent of the variance in performance. This is an important and surprising finding as it suggests that, at least in the context of venture-backed startups, regional common-pool resources such as agglomeration externalities (Delgado *et al.*, 2010), inventor networks (Fleming, King, and Juda, 2007), and institutions (Saxenian, 1996) are not the primary locus of competitive advantage. However, in subsample analyses I also investigated the impacts of regions on startups at the earliest stages of development (i.e. seed-stage ventures). For these startups, I find that regions play a relatively more important role explaining 14 percent of the variance in venture-backed startup performance variance. This finding is consistent with theoretical arguments that suggest that firms internalize capabilities and rely less on relational resources as they mature (Almeida, Dokko, and Rosenkopf, 2003; Bradley, Shepherd, and Wiklund, 2011; Sirmon *et al.*, 2011; Srivastava and Gnyawali, 2011; Vissa and Chacar, 2009). Lastly, I also find that regions are more important for venture-backed startups operating in emerging industry sectors where the knowledge base is more likely to be tacit and location-specific (Guillén, 1998; Suddaby and Greenwood, 2001), explaining approximately 7 percent of the variance in venture-backed startup performance.

These findings have some interesting and important practical implications. For instance, on aggregate, they suggest that the "secret sauce" of successful entrepreneurial ecosystems driven by technology entrepreneurship (Isenberg, 2010; Mason and Brown, 2013a; Venkataraman, 2004; Zacharakis, Shepherd, and Coombs, 2003) such as the Silicon Valley, Boulder, and Austin might really lie in the quality of the people at the center of the action, that is their entrepreneurial capital[1] (Audretsch and Keilbach, 2004a; Erikson, 2002; Feld, 2012). The differential success of particular regions might thus simply be a reflection of: a) their ability to attract talented individuals (Florida, 2002) and b) the ability for individuals with complementary skills to find each other and found effective startup teams (Kenney, 2014; Packalen, 2007). In essence successful regions are able to support positive assortative matching dynamics, where the best startups can attract and team up with the best workers (Mendes, Van Den Berg, and Lindeboom, 2010; Shimer and Smith, 2000), thus allowing startups to internalize resources in an efficient manner. In addition, the finding that the common-pool resources in regions matter significantly more for less developed venture-backed startups and for venture-backed startups in nascent industry sectors is also important in terms of its strategic implications. For instance, it suggests that entrepreneurs operating within these settings should be relatively more selective in terms of the urban areas where they choose to found their firms (Aharonson *et al.*, 2007; Pe'er, Vertinsky, and King, 2008).

Lastly, in paper 3, I focus more explicitly on how one particular regional factor, geographic clustering (Delgado *et al.*, 2010) impacts startup failure. Prior research indicates that

[1] Erikson (2002) defines entrepreneurial capital as a multiplicative function of entrepreneurial competence and entrepreneurial commitment, and argues that is the most important asset and source of competitive advantage for new ventures. Audretsch and Keilbach (2004a, 2004b, 2005) make similar arguments at the regional level.

startups are more likely to fail when they are exposed to higher levels of geographic clustering (Folta *et al.*, 2006a; Sorenson and Audia, 2000; Staber, 2001). From a theoretical perspective, scholars have largely attributed such failure in clusters to agglomeration (i.e. spatial concentration) diseconomies such as localized competition (Baum and Mezias, 1992; Folta *et al.*, 2006a; Greve, 2002; Lomi, 1995), congestion (Arnott, 2007), and higher operating costs (Prevezer, 1997). Using 7 years of data from a representative cohort of nascent startups across a range of industries in the Kauffman firm survey, I test the validity of two alternate *pre and post-entry* behavioral explanations for such observed failure; adverse selection (Flyer and Shaver, 2003; Shaver and Flyer, 2000) and performance premium (Czarnitzki, Rammer, and Toole, 2014) effects. The adverse selection explanation suggests that high quality firms will locate away from clustered regions in the face of asymmetric costs and benefits (Flyer and Shaver, 2003; Kalnins and Chung, 2004; Shaver and Flyer, 2000). The performance premium explanation suggests that startups will have higher performance expectations and require a premium level of performance to operate within clusters, in the absence of which they are more likely to *voluntarily exit*. This occurs because clusters typically have lower occupational switching costs (Folta, Johnson, and O'Brien, 2006b) and higher peer-pressure effects (Porter, 2000), both of which impact entrepreneurial decision making.

I do not find any evidence for adverse selection into geographic clusters by startups. Instead, I find that high quality startups might actually be positively selecting (i.e. self-selection instead of adverse selection) into the largest of geographic clusters. When I examine *post-entry* behavioral dynamics I find systematic evidence of geographic clustering leading to performance premium effects. For instance, I find that startups within geographic clusters are more likely to terminate operations than their peers in more isolated locations, given comparable levels of profitability. In addition when I estimate latent (i.e. unobservable) startup exit thresholds, that is the lowest level of performance below which entrepreneurs will voluntary choose to close the firm (DeTienne *et al.*, 2008; Gimeno *et al.*, 1997), I find that geographic clustering raises these survival goal levels (March and Shapira, 1992). And lastly, I also find that geographic clustering leads to lower self-evaluated ratings of performance among entrepreneurs.

The key implication from these findings is that startup failure within geographic clusters is best understood as a combination of both involuntary and voluntary exit behaviors. From a theoretical perspective this means that since exit decisions are, at least in part, due to voluntary reasons (DeTienne, 2010; DeTienne *et al.*, 2008; Gimeno *et al.*, 1997; Wennberg *et al.*, 2010) scholars should be careful in attributing the locus of startup failure within geographic clusters exclusively to agglomeration diseconomies as they have largely done to date. Furthermore, in terms of policy implications, proponents of cluster-based strategies for entrepreneurship and economic development (Rocha, 2004; Rocha and Sternberg, 2005) have largely emphasized the positive behavioral impacts that geographic clustering can have on managers, such as enhanced motivation (Porter, 2000). However my findings suggest that high levels of geographic clustering can also potentially lead to some less ideal behavioral dynamics, such that entrepreneurs tend to be less satisfied with achieved levels of performance and aspire to possibly unrealistic levels of performance (Ordóñez *et al.*, 2009). I thus suggest that policy makers interested in promoting geographic clusters of entrepreneurship (Chatterji, Glaeser, and Kerr, 2014) should be cognizant of these dynamics, so that they are able to develop a better understanding and more nuanced appreciation of how geographic clustering can impact startup behaviors and exit decisions.

To summarize, the three studies in this dissertation examine how geographic factors impact the processes of startup entry, performance, and failure (as a combination of both involuntary exits and voluntary exits). Each of these studies thus focuses on a unique aspect of the entrepreneurial life-cycle. They also study dynamics at both the regional (paper 1) and firm (papers 2 and 3) level of analysis. Put together, these three studies therefore offer a more complete picture of how geographic factors impact startup firms. Table 1 below provides a summary of the research questions, methodologies, data sources, key findings, and implications of this work.

Table 1. Summary of three dissertation papers

	Paper 1	Paper 2	Paper 3
Chapter	Chapter 2	Chapter 3	Chapter 4
Title	Knowledge Spillovers, Community Environmental Ideologies, and New Venture Creation in Clean Technology	Explaining performance differentials among venture-backed startup firms: How much does region matter?	A behavioral theory of startup failure within geographic clusters: Investigating adverse selection and performance premium effects
Research question (s)	How do incubment innovation and community sociopolitical ideologies interact to drive regional differences in cleantech entrepreneurship?	1) How much does region matter for VC-backed startup performance? 2) Does the region effect vary by startup stage and industry?	Can startup failure within geographic clusters be explained by behavioral dynamics such as adverse selection and voluntary exit due to higher peformance premiums?
Level of analysis	MSA (metropolitan statistical area) level	Startup (firm level)	Startup (firm level)
Theoretical perspectives	1) Knowledge spillover theory of entrepreneurship 2) Institutional theory (sociology)	1) Resource dependence 2) Location economics 3)Industry evolution	1) Location economics 2) Behavioral economics
Methods	Panel-based regression methods	Variance decomposition	1) Suvival analyses 2) Logit models 3) Latent threshold estimation
Data sources	1) I3 cleantech database 2) Cleantech edge patent database 3) Harvard patent dataverse 4) League of conservation voters scorecard 5) Census bureau 6) BLS quarterly census of economics and wages 7) National center for educational statistics 8) Database of state incentives for renewable energy (DSIRE) 9) Thomson Reuters VentureXpert	Thomson VentureXpert	1) Kauffman firm survey restricted data enclave 2) Harvard cluster mapping project 3) Census bureau 4) National center for educational statistics
Research context	Firms entering the cleantech sector over the period 1999 to 2010	3,893 startups located in 133 MSAs (metropolitation statistical areas) from 1980-2012, in 71 different industry sectors based on VenturExpert industry classification (VEIC) codes.	7 years of data from a nation and industry-wide representative cohort of startups founded in 2004
Key findings	1) Ideologies positively moderate the relationship between incumbent innovation and startup creation 2) This moderation effect is stronger when the regional knowledge base is less specialized in cleantech	1) On average, regions only explain 2-3% of the variance in startup performance. 2) This effect is larger for seed-stage startups (14%) and for startups operating in nascent industry sectors (7%).	1) Startup failure in clusters can be explained by agglomeration diseconomies and higher performance premiums 2) I find no evidence for adverse selection into clusters 3) I find evidence for positive self-selection into the largest clusters
Implications	1) Theories of industry evolution and entrepreneurship (Agarwal and Braguinsky, 2014)should consider the impacts of the exogenous social environment 2) Sociocultural institutions have dynamic effects such that they are more important at the nascent stages of technology adoption within a region 3) Policy makers should focus on either directly or indirecetly shaping community rhetoric when attempting to promote new technological solutions in regions, for instance by encouraging grassroots organizations such as technology SMOs (Pacheco et al., 2014).	1) Regional effects play a supporting role in explaining VC-backed startup performance relative to firm-specific factors 2) The most effective entrepreneurial regions are those that allow for the formation of strong entrepreneurial teams 3) Entrepreneurs forming startups in nascent industry sectors should be more strategic in terms of where they choose to locate their firms	1) Failure in clusters is likely due to both higher involuntary (agglomeration diseconomies) and voluntary (higher performance premiums) exit dynamics 2) Clustering might lead to both positive (e.g. enhanced motivation) and negative behavioral effects (e.g. unrealistic expectations, lower satistifaction ratings)

PAPER 1: KNOWLEDGE SPILLOVERS, COMMUNITY ENVIRONMENTAL IDEOLOGIES, AND NEW VENTURE CREATION IN CLEAN TECHNOLOGY

ABSTRACT

I model the relationship between the innovative activities of marketplace incumbents and new venture creation in the context of the emerging cleantech sector. I extend current versions of the knowledge spillover theory of entrepreneurship (KSTE) by considering the contingent effects of community environmental ideologies. My core hypothesis is that community ideologies regulate KSTE dynamics by shaping both opportunity recognition processes and entrepreneurial motivation. Consistent with this prediction, I find evidence that the positive relationship between the innovative activity of marketplace incumbents and new firm creation is enhanced by the strength of community environmental ideologies. Furthermore, I also find that the moderating impacts of environmental ideologies on KSTE dynamics weaken as the knowledge base in an MSA becomes more specialized in cleantech.

INTRODUCTION

Understanding how entrepreneurial opportunities come about and why they are differentially exploited by individuals are foundational questions in the field of entrepreneurship (Shane and Venkataraman, 2000; Venkataraman, 1997). According to the knowledge spillover theory of entrepreneurship (henceforth referred to as KSTE), *new knowledge* created by the inventive activities of incumbents is an important source of *latent* entrepreneurial opportunity (Acs *et al.*, 2009; Agarwal *et al.*, 2007; Audretsch and Keilbach, 2007; Ghio *et al.*, 2014). This is because the uncertainty associated with new knowledge leads to incomplete commercialization by the organizations that create them (Acs *et al.*, 2009; Agarwal *et al.*, 2007; Audretsch and

Keilbach, 2007). And yet, individuals that are willing to bear uncertainty can leverage such knowledge and bring it to market through the creation of new ventures (Dew, Velamuri, and Venkataraman, 2004). The KSTE therefore provides a powerful explanation for both regional and temporal differences in the prevalence of latent entrepreneurial opportunities.

However, given its focus on the origins of opportunities, KSTE says little about how potential entrepreneurs within a region actually recognize and pursue the latent opportunities that exist within regional knowledge bases. That is, by directly linking the prevalence of opportunity to rates of entrepreneurial action in a region, KSTE implicitly assumes that potential entrepreneurs in a region are: a) both alert to (Kirzner, 1979) and b) able to recognize the opportunities that incumbents are unable to pursue. Furthermore, since it makes a strong form assumption about the willingness of the entrepreneur to bear uncertainty relative to incumbents, it also does not account for heterogeneity among potential entrants in the motivational dimension of entrepreneurial action (McMullen and Shepherd, 2006); for instance potential entrepreneurs might recognize but not be willing to exploit an opportunity due to the perceived threat of incumbent competition (Plummer and Acs, 2014).

The purpose of this paper is therefore to augment existing KSTE models by investigating regional mechanisms that shape both opportunity recognition and entrepreneurial motivation. I do so by drawing upon research examining how sociocultural institutional forces impact entrepreneurial action in emerging industry sectors (Thornton, Ribeiro-Soriano, and Urbano, 2011; Tolbert *et al.*, 2011). More specifically, I focus on the impacts of community sociopolitical ideologies (Sidanius, 1985; Simons and Ingram, 2004; Weigel, 1977). In the context of emerging industry sectors, sociopolitical ideologies act as shared mental models or communication “codes” (Wenger, 2000), while also providing normative legitimacy for the industry (Aldrich and Fiol, 1994). I argue that these two properties of ideologies impact opportunity recognition and motivational processes by directing entrepreneurial attention (Gaglio and Katz, 2001), increasing information flows (Kaish and Gilad, 1991), and lowering perceived barriers of market entry (Camerer and Lovallo, 1999; Hayward, Shepherd, and Griffin, 2006; Sorensen and Sorenson, 2003). I also investigate the boundary conditions of this proposed moderation effect of sociopolitical ideologies on KSTE, by considering the evolutionary dynamics of the regional knowledge base. More specifically, I argue that the moderating impacts of sociopolitical ideologies should weaken as the regional knowledge base becomes more specialized in the relevant industry sector. This is because as particular kinds of knowledge creation and exploitation activities becomes more routinized and codified, the need for ideologies to function as a shared communication code among actors diminishes. Furthermore, from a legitimacy perspective, as particular industry sectors become more established and hence “taken-for-granted” within a region, they naturally becomes viewed as more legitimate (Sine and David, 2010), and the need for sociopolitical ideologies to act as an external source of normative legitimacy diminishes (Bitektine, 2011).

This work contributes to a number of research streams. First, from a theoretical perspective the KSTE literature has largely ignored the impacts of the regional institutional environment (Acs *et al.*, 2009). Furthermore the few studies that do consider institutional factors focus on the impacts of formal institutions (North, 1990; Ritzer and Ryan, 2010), such as intellectual property rights (Acs and Sanders, 2012; Marx, 2011). The focus of this study on sociopolitical ideologies is therefore novel within a KSTE framework as it emphasizes the impacts that more informal aspects of the regional institutional environment can have on entrepreneurial behavior (Meek, Pacheco, and York, 2010; York and Lenox, 2014). Second, this study also contributes to the growing body of research on the regional dynamics of environmental entrepreneurship (see Lenox and York, 2011 for a review) by testing theoretical models in the context of the emerging cleantech sector. The investigation of entrepreneurial entry and competitive dynamics at the metropolitan statistical area (MSA) level is also novel within this research context as prior studies have almost exclusively focused on state-level differences in entrepreneurial entry rates (Hiatt, Sine, and Tolbert, 2009; Meek *et al.*, 2010; Sine and Lee, 2009; York and Lenox, 2014). Lastly, the results from this study also have important practical implications. For instance prior research indicates that entrepreneurs focus on the availability of knowledge spillovers when strategically deciding between geographic locations (Aharonson *et al.*, 2007; Pe'er *et al.*, 2008). The results from this study condition these existing findings by emphasizing that spillovers are not necessarily automatic, and that sociopolitical community ideologies can influence how accessible spillovers are to new entrants.

The rest of this paper is organized as follows. I first provide an overview of the basic KSTE framework. In doing so, I highlight the fact that existing KSTE models are largely silent about the process by which potential entrepreneurs recognize and pursue opportunities within regions, as well as the institutional environment within which knowledge spillovers take place. Second, I briefly summarize the extant literature on sociopolitical community ideologies and discuss their impacts on opportunity recognition and entrepreneurial motivation. Following that, I present the data, measures, and empirical models that I used to test the hypotheses. I conclude with a discussion of the key findings, study limitations, and potential avenues for future research.

THEORY AND HYPOTHESES DEVELOPMENT

The Knowledge Spillover Theory of Entrepreneurship

At its core, KSTE seeks to provide an explanation for spatio-temporal variations in the origin of entrepreneurial opportunities (Hayek, 1945), particularly in relation to technology entrepreneurship (Garud and Karnøe, 2003). Its key assertion is that entrepreneurial opportunities arise endogenously from *prior* technological innovations, and more specifically from "*new knowledge* that is created in one organizational context, such as a firm or university research laboratory, but that is left uncommercialized by the incumbent organization as a result of the uncertainty inherent in new knowledge" (Audretsch and Keilbach, 2007: 1246). The value of new knowledge is highly uncertain since the organizations developing it have little ability to predict whether it will be ultimately viable (Heeley and Jacobson, 2008). This problem is particularly acute in the context of emerging industry sectors where firms often operate with new, unproven business models (Sanders and Boivie, 2004). Under these uncertain conditions incumbent firms encounter a number of filters that impede their ability to exploit the knowledge that they have created (Acs and Plummer, 2005). For instance, incumbents might choose to not pursue new technology trajectories due to cognitive or structural path dependencies (Moran and Ghoshal, 1999; Stuart and Podolny, 1996; Tripsas, 1997; Tripsas and Gavetti, 2000). Beyond such inertial factors, incumbents might also lack the complementary assets needed to bring technologies to the marketplace in a timely manner (Teece, 1986). Hence while incumbent organizations, particularly large firms, are often willing to explore extensively to create new knowledge they are far more reluctant to bring them to the marketplace (Cassiman and Ueda, 2006; Dew *et al.*, 2004). These inefficiencies in the marketplace create opportunities for alert entrepreneurs to leverage existing pools of underexploited knowledge. Furthermore given that knowledge creation activities vary significantly across regions (Audretsch and Feldman, 1996; Fischer, 2001), KSTE suggests that geographic differences in new venture creation are a manifestation of knowledge spillovers from spatially heterogeneous "repositories of existing knowledge" (Agarwal *et al.*, 2007: 266).

KSTE also asserts that such spillovers are not necessarily detrimental to marketplace incumbents (Agarwal *et al.*, 2007; Kotha, 2010). Instead it suggests that a symbiotic relationship exists between the knowledge creation activities of incumbents and the knowledge exploitation role of new entrants in the marketplace (Baumol, 2002; Hockerts and Wüstenhagen, 2010). This is because over time, this endogenous process of knowledge creation and exploitation leads to the establishment of positive feedback loops between incumbents and new entrants. For instance established incumbents can benefit from reverse spillovers also known as knowledge spill-ins, such that they are able to learn from and collaborate with more nimble new entrants who occupy complementary positions in the value chain (Kotha, 2010; Operti and Carnabuci, 2014; Somaya, Williamson, and Lorinkova, 2008; Yang, Phelps, and Steensma, 2010). This process of creative construction (Agarwal *et al.*, 2007) further accelerates the process of innovation and leads to a specialization of the regional knowledge base in particular technological domains. Incumbents can leverage these spatially localized knowledge bases to gain competitive advantages over competitors in other geographic locales (Jenkins and Tallman, 2010; Tallman *et al.*, 2004), and the creation of specialized knowledge pools also generates more entrepreneurial opportunities in particular regions (Audretsch, Keilbach, and Lehmann, 2006; Shane, 2008). With regards to observable differences across regions, I therefore expect that in locations where more new knowledge is created there should also be higher rates of entrepreneurial activity. As a baseline hypothesis, I therefore suggest that:

Hypothesis 1: Stocks of new knowledge in a region are positively related to new venture creation rates.

Gaps in the KSTE framework: Opportunity recognition and entrepreneurial motivation

The efficiency of knowledge spillovers is therefore a central concern of KSTE. In particular, mechanisms that lead to enhanced opportunity recognition and pursuit by potential entrepreneurs in a region are an important consideration. If potential entrepreneurs are able to identify and act upon available opportunities more easily, they should be able to exploit the regional knowledge base more efficiently. But how do potential entrepreneurs actually recognize opportunities generated by the innovation activities of incumbents? And do all entrepreneurs pursue these opportunities once they are recognized?

It is thus rather surprising that while KSTE primarily adopts a discovery perspective on entrepreneurial opportunity (Alvarez and Barney, 2007; Eckhardt and Shane, 2003) with a strong form of distributed agency between the roles played by incumbents and new entrants (Acs *et al.*, 2009; Garud and Karnøe, 2003), it does not specify the mechanisms that facilitate the recognition and exploitation of opportunities by potential entrepreneurs. Instead the majority of KSTE studies implicitly assume that: a) a pool of profit-seeking entrepreneurs exist within a region (Audretsch, Bönte, and Keilbach, 2008; Audretsch and Keilbach, 2004a, 2005; Baumol, 1996) and b) all new knowledge created in a region and not pursued by incumbents is potentially available for exploitation by alert entrepreneurs (Plummer and Acs, 2014: 125). Adopting an Austrian economics lens, this view therefore assumes that there is no “radical ignorance” in the marketplace (Kirzner, 1979); the fact that entrepreneurs are inclined to pursue opportunities that incumbents will not can be explained solely by their greater willingness to bear the high levels of uncertainty associated with new knowledge (Alvarez and Barney, 2007; Knight, 1921; Langlois and Cosgel, 1993; Roth, 2009).

However, this relatively frictionless view of knowledge spillovers among actors within regions has been increasingly criticized in both the strategic management and economic geography literature as a relatively rosy picture of knowledge exchange dynamics (Breschi and Lissoni, 2001a; Huber, 2012; Lambooy, 2010; Shaver and Flyer, 2000). For example at the firm level, entrants into a region are likely to be heterogeneous in their ability to both benefit from and contribute to knowledge spillovers due to differences in capabilities (Aharonson *et al.*, 2007; Pe'er *et al.*, 2008; Shaver and Flyer, 2000). Similarly at the regional level, locations might differ in the degree to which actors are willing to exchange information and collaborate extensively (Laursen, Masciarelli, and Prencipe, 2012; Saxenian, 1996). Thus Breschi and Lissoni (2001a: 255) argue against the automaticity of knowledge spillovers, and indicate that far more attention needs to be paid towards "how knowledge is actually transmitted, among whom, at what distance, and on the basis of which codebooks".

Hence more recently, there has been increasing interest in identifying the specific mechanisms that facilitate knowledge spillovers more generally with regions, as well as spillover-based entrepreneurship more specifically. Much of this research has focused on firm-level mechanisms that influence the dynamics between marketplace incumbents and new entrants such as entrepreneurial spinoffs (Agarwal *et al.*, 2007; Klepper, 2009; Klepper and Thompson, 2005; Lockett *et al.*, 2005). In comparison, our understanding of how regional level institutions (e.g. property rights, tax structures, incentives, cultural attributes) shape KSTE dynamics is far more limited[2]. However, there has been no research on how the more informal and decentralized sociocultural elements of the regional institutional environment such as

[2] The few studies that have begun to explore this issue have been limited in their focus on formal (e.g. regulatory) institutions. For instance recent research has looked at the impacts of regional differences in legal environments on incumbent innovation as well as new venture creation rates (Acs and Sanders, 2012; Conti, 2014; Marx, 2011; Marx, Strumsky, and Fleming, 2009).

behavioral norms and ideologies (Elster, 1989; Saxenian, 1996; Thornton *et al.*, 2011; Tolbert *et al.*, 2011; York and Lenox, 2014) impact KSTE dynamics.

I address this gap in the KSTE literature by drawing upon a body of work that studies the impacts of such sociocultural institutions on entrepreneurial behavior (see Jennings *et al.*, 2013; Tolbert *et al.*, 2011 for reviews of this literature). Regional sociocultural institutions broadly refer to cultural practices, patterns, or behaviors that are organized and reinforced by groups of individuals, such as the members of a local community. In contrast to more formal aspects of the institutional environment, these institutions are typically decentralized and emerge from the social exchange dynamics of the individuals within the community (North, 1990; Ritzer and Ryan, 2010). In the context of this study, I focus in particular on the impacts of one particular example of such an institution, sociopolitical community ideologies (Buttel and Flinn, 1978; Sidanius, 1985; Simons and Ingram, 2004; Weigel, 1977)[3]. More specifically I examine how such ideologies influence both opportunity recognition and entrepreneurial motivation, and hence impact entrepreneurial entry dynamics within a KSTE framework.

Sociopolitical community ideologies, opportunity recognition, and entrepreneurial motivation

Ideologies broadly refer to "a set of beliefs about the social world and how it operates, containing statements about the rightness of certain social arrangements" (Wilson, 1973: 91) . Although they are primarily simply a set of ideals and vary significantly in the degree to which they encompass different social, economic, and political spheres, they can and do significantly affect the behaviors of both organizations and individuals (Ingram and Simons, 2000; Simons and Ingram, 2004). This is because as a value system, they support a series of behavioral norms

[3] Note that while I develop general theoretical arguments with respect to how such ideologies should operate in *general*, I subsequently test my hypotheses by studying community-level *environmental ideologies* since these are most directly relevant to the cleantech sector.

(Elster, 1989). Once an ideology has been accepted by a community, individuals who do not confirm to the ideology's behavioral norms risk social sanctions such as stigmatization (Knack, 1992). The wide-scale engagement of individuals in normatively acceptable behavior also results in the propagation of rhetoric that can further reinforce the underlying ideology (Barley and Kunda, 1992). This communal nature of ideologies and the need for rhetoric to maintain them means they also have a strong underlying socio-cognitive basis. In effect, they act as shared group-level mental models or cognitive maps, allowing sets of actors (e.g. individuals, organizations) to both interpret and structure the environment that they interact in (Denzau and North, 1994; Johnson-Laird, 1983; Wilson, 1973). To the extent that they are community or region-specific, they can also be therefore conceptualized as localized decentralized sociocultural institutions, defining the "informal rules of the game", and facilitating coordination among individuals in a community. Such coordination can occur even when individuals have potentially conflicting preferences and different economic rationale (e.g., Richards, 2001; Richards, McKay, and Richards, 2002). More specifically with respect to regulating entrepreneurial action and KSTE dynamics, I focus on the impacts of sociopolitical community ideologies. As the label suggests, such ideologies combine dimensions of both social and political systems. For example, a sociopolitical ideology of environmentalism combines social attitudes such as being more sustainable and ecologically conscious as well as political elements such as an emphasis on grassroots democracy, health, and human rights (Hoffman, 1999; Stern *et al.*, 1999).

I expect that such ideologies should play an important role in shaping both opportunity recognition and entrepreneurial motivation respectively. With respect to opportunity recognition, ideologies can make the process of identifying opportunities more efficient for two principal reasons. First, since all forms of ideologies serve as mental model or cognitive schema (Denzau and North, 1994), they help direct the attention of individuals towards specific problem domains and certain aspects of available information (Shepherd, Williams, and Patzelt, 2015). This focus of attention also has the effect of switching cognitive modes from more passive styles of information processing to active thinking and problem solving (Louis and Sutton, 1991). This in turn increases the likelihood that that potential entrepreneurs can recognize meaningful patterns and “connect the dots” through purposeful and directed search activities (Baron, 2006; Baron and Ensley, 2006; Gaglio and Katz, 2001; Hsieh, Nickerson, and Zenger, 2007) . Second, such shared group-level mental models also act as a transmission code or “carrier wave” (Wenger, 2000). The notion of a carrier wave can best be understood as a metaphor from the field of telecommunications, where a carrier is a single high frequency wave that acts as a medium over which lower frequency signals can be transmitted. I expect that the existence of such a communication code should allow potential entrepreneurs to recognize a wider range of opportunities within the available opportunity set. For instance, some opportunities are more tacit and hence harder to recognize relative to others (Samuelsson and Davidsson, 2009; Smith, Matthews, and Schenkel, 2009). Ideologies, by acting as a shared code and hence providing a basis of communication, can facilitate the exchange of such hard to articulate information thereby increasing the flow of information that a potential entrepreneur is exposed to. In this manner, ideologies can allow individuals to both make sense of and interpret larger volumes of information from others in a timely manner (Kaish and Gilad, 1991; Vaghely and Julien, 2010). This is particularly important within the opportunity discovery framework emphasized in KSTE, since the temporal window within which an opportunity can be exploited is often limited (Eckhardt and Shane, 2003; Plummer and Acs, 2014).

In addition to enhancing opportunity recognition by impacting entrepreneurial cognition, supportive sociopolitical ideologies can also have important motivational impacts on entrepreneurs. In particular, since they are able to shape what is accepted by the collective as acceptable and considered legitimate (Ingram and Simons, 2000), they act as an important source of normative legitimacy for the industry sector (Aldrich and Fiol, 1994; Deeds, Mang, and Frandsen, 2004). This kind of normative support can have significant impacts on the motivational element of entrepreneurial action, that is the decision to pursue opportunities once identified (McMullen and Shepherd, 2006). Since goods and services that are deemed socially acceptable are more likely to be adopted in the marketplace (Jonsson and Regnér, 2009), normative legitimation of an industry sector reduces the degree of perceived uncertainty (Townsend and Hart, 2008). Furthermore the presence of such legitimacy also allows for complementary resources to be more easily mobilized (Aldrich and Stern, 1983; Lounsbury and Glynn, 2001; Sirmon *et al.*, 2011). For instance, new entrants might find it easier to convince outside investors of the viability of their business models and hence acquire capital to finance their ventures (Petkova *et al.*, 2014) . The motivational impacts of such resource availability on new entrants is a general reduction in the fear of failure, a shift towards a promotion vis-à-vis a prevention regulatory focus (i.e. a focus on positive outcomes rather than potential downsides), and a higher expected level of performance (Brockner, Higgins, and Low, 2004; Forlani and Mullins, 2000; March and Shapira, 1987; Naffziger, Hornsby, and Kuratko, 1994; Sitkin and Pablo, 1992). These dynamics increase the confidence level of new entrants, such that they are more susceptible to decision making biases and hence more likely to ignore competitive threats posed by marketplace incumbents (Busenitz and Barney, 1997; Camerer and Lovallo, 1999; Hayward *et al.*, 2006; Sorensen and Sorenson, 2003). Under such conditions I therefore expect that individuals are more likely to pursue opportunities that are aware of, either through their prior experiences within incumbent organizations or through their contacts within region social networks (Agarwal *et al.*, 2007; Davidsson and Honig, 2003; Dew *et al.*, 2004; Klepper and Sleeper, 2005; Shane, 2000).

To summarize, these arguments suggest that community-level ideologies in support of an emerging industry sector can have important positive effects on both opportunity recognition and entrepreneurial motivation. From a KSTE perspective, this therefore means that community-level ideologies can act as an important contingency moderating the efficiency with which potential entrepreneurs exploit latent opportunities in the regional knowledge base. More formally, I therefore hypothesize that:

Hypothesis 2: Supportive sociopolitical community ideologies will moderate the positive relationship between stocks of new knowledge in a region and new venture creation rates such that this relationship will be stronger in regions with stronger ideologies.

Specialization of the regional knowledge base and the impact of sociopolitical community ideologies

Having thus suggested that sociopolitical ideologies are likely to be particularly important in moderating KSTE dynamics, I explore an important boundary condition of this extended model. In particular, I expect that the moderating influence of ideologies should decline as the regional knowledge base becomes more specialized in the relevant technological domain. I discuss why this is likely to be the case with regards to effects on both opportunity recognition and entrepreneurial motivation respectively below.

With respect to opportunity recognition processes, I expect that the influence of ideologies should decline as the regional knowledge base becomes more specialized for two main reasons. First, as regions specialize in certain technologies, they often lock-in to specific innovation trajectories (Martin and Sunley, 2006). This constriction of the knowledge domain has the effect of focusing the attention of actors on specific problem domains, hence naturally channeling the search process (Hsieh *et al.*, 2007). I therefore expect that the need for ideologies to regulate the attentional processes of potential entrepreneurs should decline under such conditions. Second, increased levels of specialization in particular kinds of knowledge creation activities leads to the establishment and codification of routines (Arora and Gambardella, 1994). Once such routines are established, both the innovation process as well as the knowledge base underlying the set of latent opportunities becomes less tacit (Cowan, David, and Foray, 2000; Ter Wal, 2013). This in turn simplifies the process of knowledge exchange between individuals. Such conditions should reduce the need for community ideologies to function as a code (Breschi and Lissoni, 2001a; Wenger, 2000) that facilitates information exchange.

With respect to the impacts of sociopolitical ideologies on entrepreneurial motivation, I similarly expect that their influence should decline as the regional knowledge base becomes more specialized in a particular technological domain. This is because specialization leads to an increase in the cognitive legitimacy and "taken-for-grantedness" of the focal industrial activity (Bitektine, 2011). This leads to a significant reduction in both actual and perceived levels of technological and marketplace uncertainty. The need for the normative legitimation provided by supportive ideologies will thus be lower and less of a factor in shaping entrepreneurial behavior. Moreover, from a decision making standpoint, both the value and use of a deliberate approach to market entry decisions increases significantly in these more established markets (Brinckmann, Grichnik, and Kapsa, 2010; Castrogiovanni, 1996; Delmar and Shane, 2003; Gruber, 2007). Potential entrants are more likely to be deterred by the competitive threat posed by incumbents, and heuristics have lower impacts on entry decisions (Chwolka and Raith, 2012). I therefore expect that under such conditions, entrepreneurs will be more reluctant to exploit opportunities that they are aware of (Shane, 2001), thereby further negating the influence of ideologies on entrepreneurial action.

To summarize, these arguments suggest that the impacts of sociopolitical community ideologies on moderating KSTE dynamics should be strongest when the regional knowledge base is less specialized in the relevant knowledge domain. More formally, I therefore hypothesize that:

Hypothesis 3: The moderating effects of supportive community ideologies (hypothesis 2) will be more positive when a region is less specialized in the relevant industry sector.

METHODOLOGY

Study Context

I tested these hypotheses by examining regional variations in new venture creation rates in the cleantech sector. This is an industry sector that encompasses a broad array of technologies that seek to address issues of sustainability and/or environmental degradation. For example, firms might operate in industry segments such as solar, recycling and waste, water, energy efficiency, biofuels, transportation, agriculture, energy storage, and smart grid (Pernick and Wilder, 2007; Petkova *et al.*, 2014). This sector provides an ideal research setting for a number of reasons. First, it is a technology intensive sector where knowledge creation activities are commonplace. For instance in a comprehensive analysis of innovation activity in the renewable energy industry, a core aspect of clean technology, Nanda *et al.* (2013) indicate that patenting rates have increased consistently over the past two decades. Furthermore, by identifying inventors and patent assignees they also demonstrate that the majority of new knowledge is generated by large, established incumbents in both the private and public sector (e.g. large corporations, federal R&D labs, universities). Furthermore, consistent with other sectors of the economy, they find that substantial geographical variations exist in both the rates and types of innovations that are granted. Given that the knowledge spillovers from these "regional knowledge pools" are also spatially localized, one might also expect to find regional variance in new venture creation. Not surprisingly, prior studies focusing on subsectors of cleantech have documented such geographic differences (Hiatt, 2010; Kapoor and Furr, 2014; Meek *et al.*, 2010; Pacheco *et al.*, 2014; Russo, 2003; Sine and Lee, 2009; York and Lenox, 2014).

Second, the cleantech context allows for a direct study the impacts of community-level environmental ideologies. These sociopolitical ideologies vary significantly across regions and have been shown to be particularly important in shaping the behaviors of both individuals and organizations (Delmas, Russo, and Montes-Sancho, 2007; Kahn, 2007; Kahn and Vaughn, 2009). In contrast, in other settings that have been traditionally studied (e.g. software, biotechnology), such ideological viewpoints are harder to empirically operationalize. Hence, while I expect that sociopolitical ideologies should also similarly matter in other knowledge-based sectors of the economy, the use of this "value-laden" industry context allows me to quantify their impacts in a systematic fashion.

Lastly, extant frameworks suggest that the "environmental entrepreneurs" (Lenox and York, 2011) active in this sector seek to solve problems primarily through technological solutions, while exploiting opportunities in imperfect markets under conditions of uncertainty (Cohen and Winn, 2007; Dean and McMullen, 2007; York and Venkataraman, 2010). Cleantech is an emerging industry sector characterized by high levels of both market and technological uncertainty (Jacobsson and Johnson, 2000; Petkova *et al.*, 2014; Schilling and Esmundo, 2009). Furthermore, historical path-dependencies have meant that regions have differentially specialized in cleantech subsectors by leveraging existing knowledge bases in related industries (Cooke, 2008; Hayter, 2008); for instance much of the innovation in the solar photovoltaic industry occurs in the Silicon valley region as it draws on earlier innovations in the semiconductor industry. There is therefore a significant degree of heterogeneity in the degree to which regional knowledge bases are specialized in clean technologies (Cooke, 2008), allowing for a test of the third hypothesis in the model.

Data Sources and Variables

Dependent Variable

New venture creation: I utilized the *i3* cleantech database (www.i3connect.com) as the primary source of information on startup activity in the cleantech sector. This database uses a broad-based definition of cleantech so as to capture startup activity in a wide array of technologies that seek to address issues of sustainability and/or environmental degradation (e.g., Pernick and Wilder, 2007; Petkova *et al.*, 2014). Startups in the database are categorized in a number of custom-defined sub segments such as solar, recycling and waste, water, energy efficiency, biofuels, transportation, agriculture, energy storage, and smart grid. Firm founding years and location information at the zip code level are also available in this database. For ventures with missing founding date or location information, I conducted a variety of online searches on company websites, secretary of state websites, and online business directories to triangulate and backfill information. I then aggregated these entry events at the MSA (metropolitan statistical area) level on annual basis. The dependent variable in the model is therefore a non-negative count variable. In sum, I was able to identify 3,523 cleantech venture new entrants across 237 MSAs over the period 1999-2010.

Note that to match and aggregate zip code information to the MSA level, I used a correspondence file from the Master Area Block Equivalency (MABLE) database (http://mcdc2.missouri.edu/websas/geocorr2k.html). Hosted by the Missouri census data center, this tool allows users to query and generate concordance tables for spatial units of different scales such as zip codes, counties, MSAs, congressional districts, and states. Furthermore, it also provides population weighted allocation factors which can be used as weights in situations where a direct one-to-one match between spatial levels is not possible (e.g. zip codes cross MSA boundaries, MSAs can cross state boundaries, congressional districts fall partly within MSAs).

Independent Variables

Regional stocks of new knowledge (in cleantech): Following prior work that has used a KSTE framework (Acs *et al.*, 2009; Audretsch *et al.*, 2008; Audretsch and Keilbach, 2008; Plummer and Acs, 2014), I used a count of patents that were granted by region and year as a measure of incumbent innovation. For this purpose, I purchased data on patenting in the cleantech sector from the *Cleantech Patent Edge* database (www.cleantechpatentedge.com). This data was then geocoded by matching it to the disambiguated US patent inventor database (Li *et al.*, 2014) publicly available from the Institute for Quantitative Social Science (http://thedata.harvard.edu/dvn/dv/patent). This database uses inventor address information on all patents filed with the USPTO, and geocodes patents at the zip code level over the period 1975-2010. Once the zip codes associated with the subset of cleantech patents was identified, I aggregated this measure to the MSA level on a yearly basis again using concordance files from the MABLE database. Consistent with prior research, I log-transformed this variable prior to estimation since it was strongly right-skewed (Samila and Sorenson, 2010, 2011).

Community environmental ideologies: To construct a time-varying, MSA level measure of sociopolitical ideologies towards the cleantech sector environmental issues I used measures from the League of Conservation Voters scorecard (www.lcv.org). The League of Conservation Voters is a policy advocacy organization whose mission is to advocate for sound environmental policies and elect pro-environmental representatives at both the house and senate branches of the legislature. LCV scores have been previously used in the literature as a measure of environmental ideologies of communities (Delmas *et al.*, 2007; Kahn, 2007).

Crucially for the purpose of this study, the LCV maintains a record of the voting patterns of both house and senate representatives on environmental issues. This data is available on an annual basis at the congressional district level. Using concordance files from the MABLE database, I first identified spatial overlaps between congressional districts and MSAs. I did this in a time varying fashion for each congressional term (i.e. every two years) as both the number and spatial boundaries of congress districts do change over time. Using this matched information, I then created aggregated LCV scores for each MSA taking into account concordances and allocation factors between congressional districts and MSAs. For instance, the MSA of Akron, OH over the period 1999-2002 encompasses both the 13th and 14th congressional district of the state of Ohio. The population (from the decennial census in 2000) of the 13th district was 111,913 and the population of the 14th district was 583,047. LCV scores from the 13th district therefore receive a comparatively lower weight of 0.16, while LCV scores from 14th receive a weight of 0.84 in computing the average score for Akron during this time period. This method allowed me to compute a time-varying, weighted average LCV score for each MSA.

Specialization of the regional knowledge base in cleantech: To estimate the degree to which a regional knowledge base was specialized in cleantech I again used patent data. More specifically, starting with the first available geo-coded patent data in 1975 from the USPTO, I computed the cumulative number of cleantech and all patents granted in an MSA. I then calculated the ratio of these two cumulative measures. Higher values of this metric therefore refer to a region where a greater share of the cumulative innovation has been historically focused on cleantech.

Control Variables

I included a number of control variables in estimation models to account for other factors that might drive regional differences in new venture creation in cleantech. While I was able to estimate a number of these variables at the MSA level, I could only obtain other measures at the state level. For state level controls, I used the MABLE database to identify concordances between MSAs and states, and assigned scores accordingly. For MSAs that spanned state boundaries, I assigned MSAs to states with the largest allocation factor (i.e. the state within which the highest % of the MSA population resided within). I then weighted the state-level control variables by this allocation factor. In a separate set of robustness tests, I also used a mixed-methods approach to analyze the data, thus explicitly incorporating the fact that MSAs are geographically hierarchically clustered within states.

General MSA level controls

I first included a set of general MSA level controls, not specific to the cleantech context, but which could influence rates of new venture creation. First, using data from the Census bureau I accounted for the physical size of MSAs by controlling for the MSA *land area*. Second, I captured the general level of economic prosperity in a region by controlling for the *poverty rate* measured as the number of persons in poverty per capita. I also tried alternate measures such as per capita *GDP*, *real GDP*, and *personal income* in sensitivity tests. Third, I controlled for the residential zoning code of the MSA through the annual *number of housing permits* issued per capita. Presumably, MSAs that are dominated by residential zones are less likely to have office space and hence lower rates of new venture creation. Fourth, using data from the integrated postsecondary education data system (IPEDS) maintained by the National Center of Education Statistics (http://nces.ed.gov/ipeds/deltacostproject/) I computed per capita levels of *enrollment in postsecondary education.* This acted as a measure of general human capital. This data was available at the zip code level for institutions. I therefore aggregated it to the MSA level using concordance tables from the MABLE database. Note that I preferred to use this metric instead of the more sparsely sampled decennial census and American Community survey data on population education levels, given that I required time-varying measures on a yearly basis. Lastly, given that the extant literature that has identified strong co-location patterns between new ventures and universities (Audretsch, Lehmann, and Warning, 2005), I used the IPEDS database to control for the number of *doctoral granting and research institutes* per capita in an MSA. Within the IPEDS database, I used the Carnegie classification of institutions to identify the relevant institutions of higher education.

MSA level controls specific to cleantech

Investment of risk capital in cleantech: The availability of risk capital in a region influences startup entry rates (Samila and Sorenson, 2011). To capture regional differences in the availability of financing for startups in the cleantech sector, I utilized data on investments from the *i3* database as well as overall rates of venture capital investment from the Thomson *VentureXpert* database. The i3 data is available from 1999 onwards, and aggregates financing from both private (i.e. Venture capitalists) and public (i.e. government grants) sources. Using these two sources, I computed a metric of *cleantech risk capital investment* at the MSA level by taking the ratio of investments in cleantech relative to total venture investment.

Clustering and competition in cleantech: Industrial clusters might attract entrepreneurs and hence lead to higher rates of startup formation (Delgado *et al.*, 2010). At the same time, the increased levels of localized competition within clusters might also spur incumbent firms to innovate more rapidly (Audretsch and Feldman, 1996), as well as potentially impede new venture creation rates (Plummer and Acs, 2014). Hence, I used a location quotient measure (Alcácer and Chung, 2014) to account for the relative regional level of clustering and localized competition in the cleantech sector.

Given the broad-based definition of cleantech, regional establishment counts of "clean vs. non-clean" firms are not readily available over the time span of the study from 1999-2010, particularly at a fine-grained MSA level. I was however able to approximate a location quotient metric by combining data from two Bureau of Labor Statistics sources, the Quarterly Census of Employment and Wages (QCEW) (http://www.bls.gov/cew/datatoc.htm) and the green good and services (GGS) classification list (http://www.bls.gov/ggs/). The GGS created by the BLS in 2010 allows for an identification of "clean vs. non-clean" industry sectors based on 6-digit NAICS 2007 definitions. Given the time span of the study from 2000-2011, I also created a similar list based on 6-digit NAICS 2002 definitions using concordance files from the census to match between historical NAICS categorizations. I then extracted data on annual averages of monthly employment and establishments from the QCEW by 6-digit NAICS code, year, and geographic county. By joining this data file to the GGS NAICS classification list I was therefore able to estimate annual county-level averages of establishment and employment levels in both “cleantech and non cleantech industries”[4]. Using concordance files from the MABLE database I then aggregated this county level data to the MSA level. Lastly, I created location quotient metrics by normalizing the ratio of cleantech establishment and employment at the MSA to the same ratio at the national level. I used the following shown below to compute these location quotients for each MSA *i* in a focal year *t*:

$$\mathrm{LQ}_{\mathrm{i,t}} = \frac{\frac{\text{Annual monthly average of cleantech establishments in MSA}}{\text{Annual monthly average of all establishments in MSA}}}{\frac{\text{Annual monthly average of cleantech establishments in USA}}{\text{Annual monthly average of all stablishments in USA}}}$$

General state level controls

State enforcement of non-compete covenants: The mobility of knowledge workers can have a direct influence on spillover-based entrepreneurship (Agarwal *et al.*, 2007). Following prior work (Stuart and Sorenson, 2003), I therefore controlled for state-level variations in the enforcement of non-compete covenants (Garmaise, 2011; Marx, 2011). Presumably, in states such as California where these covenants are not enforced, individuals are more likely to job-hop between organizations and create spinoff firms (Fallick, Fleischman, and Rebitzer, 2006) hence increasing the likelihood of KSTE dynamics. I multiplied this measure by the MSA-state allocation factor provided by the MABLE database for MSAs that spanned state boundaries.

State level controls specific to cleantech

State renewable energy generation ratio: The usage and consumption of energy derived from renewable sources differs significantly from state to state. Given the high proportion of renewable energy goods and services within cleantech, I expected that this factor might have an impact on both cleantech innovation as well as entrepreneurial entry rates. To capture this state-

[4] As indicated by the GGS listing of industries, not *all jobs* within an entire NAICS-defined industry segment are necessarily “green jobs”. Hence the measures that I constructed based on NAICS categorizations, while indicative of industry segments where there is a higher or lower emphasis on cleantech, are necessarily approximations.

level differences, I used data from the State Energy Data System (SEDS) provided by the US Energy Information Administration (http://www.eia.gov/state/data). More specifically, I computed an annual ratio of energy generated by renewable sources relative to the total energy production in a state. I multiplied this measure by the MSA-state allocation factor provided by the MABLE database for MSAs that spanned state boundaries.

State implementation of renewable portfolio standards: US states have varied in the degree to which they have adopted and implemented renewable portfolio standards. Similar to the effects of state energy profiles, I expected that this variance could potentially impact market dynamics in the cleantech sector. I therefore controlled for this state wide difference in legislative action, by identifying the year when RPS standards were first implemented in a state. I assigned a dummy value of 1 to this year and all subsequent years, and a value of 0 for all years prior to adoption. I multiplied this measure by the MSA-state allocation factor provided by the MABLE database for MSAs that spanned state boundaries.

Policies and incentives for renewable energy: States vary significantly in terms of the incentives and policies in place to support renewable energy adoption. While primarily directed at fostering innovation and consumer adoption, these policies might also influence startup activity in cleantech. Hence, following prior research (Meek *et al.*, 2010; York and Lenox, 2014) I controlled for these regulatory factors using data from the database of state incentives for renewables and energy (DSIRE). I multiplied this measure by the MSA-state allocation factor provided by the MABLE database for MSAs that spanned state boundaries.

Estimation Model

I conducted all analyses at the MSA level, organizing the data in a panel format, and using temporally lagged covariates with clustered standard errors. As indicated earlier, the dependent variable is a non-negative count of entrants in a region-year. Furthermore residuals violated assumptions of normality and were overdispersed. Hence, following prior work that has estimated similar models (Rao, 2004; York and Lenox, 2014) I estimated random effects negative binomial models with standard errors clustered by group (i.e. MSA). Similar to these studies, I did not use a fixed effects specification given that the fixed effects estimator is biased and unreliable in a negative binomial framework, and does not control for unchanging covariates (Allison and Waterman, 2002; Cameron and Trivedi, 2013). I did include a full set of year dummies to account for cross-regional temporal variations in the data.

RESULTS

Table 2 provides descriptive statistics and bivariate correlations. I highlight the relationships that are statistically significant with respect to the dependent variable of interest, cleantech new venture creation.

Table 2 (paper 1). Descriptive statistics and correlation coefficients

Variables	Mean	S.D.	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1.Number of cleantech new ventures	1.04	3.38														
2.Stock of new knowledge (ln # of patents granted)	1.35	1.35	0.50***													
3.Supportive (enviromental) ideologies	0.44	0.33	0.20***	0.26***												
4.Land area (sq. miles) (X1000)	3.35	3.5	0.18***	0.24***	-0.15***											
5.Poverty rate (num persons in poverty per capita)	0.12	0.05	-0.17***	-0.13***	-0.17***	0										
6.Residential permits per capita (X0.01)	0.51	0.46	-0.03	0.03*	-0.21***	0.08***	-0.04*									
7.Post secondary (PS) institution enrollment per capita (X0.01)	0.23	0.83	0.04*	0.08***	0.03	0.11***	0.02	-0.03								
8.Number of doctoral & research universities per capita (X100,000)	1.01	2.07	0	0.11***	0.04*	-0.07***	0.05***	0.02	-0.03							
9.Cleantech cluster(location quotient)	1.03	0.18	0	0.03	0.04*	0.02	-0.09***	0.09***	0.01	0.18***						
10.Investment of risk capital in cleantech	0.05	0.18	0.01	0.03	0.03	0.01	0.04*	-0.06***	0.04*	0.02	0.04*					
11.State enforcement of non-compete covenants	4.10	2.16	-0.16***	-0.05***	-0.14***	-0.16***	-0.09***	0.18***	0	0.02	0.04*	-0.03				
12.State renewable energy production ratio	0.27	0.27	0.01	-0.04*	0.23***	-0.01	-0.13***	0.07***	-0.05*	0.03	0.19***	0.06***	0.24***			
13.State implementation of RPS standards	0.35	0.47	0.11***	-0.09***	0.22***	0.03*	-0.02	-0.17***	0.09***	-0.07***	-0.05***	0.1***	-0.31***	0		
14.State policies for renewable energy	8.04	10.99	0.16***	0.06***	0.1***	0.02	0.1***	-0.07***	0.01	-0.05*	0.05*	0.04*	-0.24***	0.12***	0.24***	
15. Specialization of the regional knowledge base in cleantech	0.05	0.02	--0.13***	-0.06***	-0.15***	0.04*	0.18***	-0.01	-0.03	-0.04*	0	0.02	0.03	-0.15***	-0.03	-0.06**

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

As expected the knowledge spillover variable has a strong positive correlation with rates of cleantech new venture creation (r=0.50, p<0.001), affirming the core mechanism of KSTE. Supportive ideologies are also positively correlated with rates of entrepreneurial activity (r=0.20, p<0.001). Among the MSA level control variables, I find that rates of cleantech new venture creation are significantly positively correlated with the size (land area: r=0.18, p<0.001), level of affluence (poverty rate: r=-0.13, p<0.001) and the general level of human capital of an MSA (post-secondary enrollment: r=0.04, p<0.05). Lastly, with respect to state level control variables I find that enforcement levels of non-compete covenants are negatively correlated to clean tech venture creation (r=-0.16, p<0.001), while a state legislature that is favorable to renewables is positively correlated with cleantech venture creation (implementation of RPS standards: r=0.11, p<0.001; policies for renewable energy: r=0.16, p<0.001). I also find that the specialization of the regional knowledge base in cleantech is negatively correlated to rates of new venture creation (r=-0.13, p<0.001).

Table 3 presents findings from negative binomial regression models.

Table 3 (paper 1). Random effects negative binomial models for cleantech entry into $MSA_{i,t}$

	Full sample analysis			Subsample analysis (median split on specialization of the regional knowledge base in cleantech)		
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
VARIABLES	Controls Only	Main Effects Included	Interaction Effects	Full sample with specialization of MSA knowledge base in cleantech included	Low specialization of MSA knowledge base in cleantech	High specialization of MSA knowledge base in cleantech
Stocks of new knowledge (ln # of patents granted)		0.92***	0.72***	0.72***	0.71***	0.80***
		(0.06)	(0.08)	(0.07)	(0.11)	(0.11)
Supportive (enviromental) ideologies		0.30*	-0.47	-0.50	-0.40	0.07
		(0.17)	(0.31)	(0.32)	(0.39)	(0.39)
Stocks of new knowledge*Supportive ideologies			0.33***	0.32***	0.36***[a]	-0.03[a]
			(0.12)	(0.12)	(0.13)	(0.14)
Residential permits (per capita)	-0.62***	0.28**	0.33**	0.31**	0.32*	0.20
	(0.24)	(0.13)	(0.13)	(0.12)	(0.18)	(0.15)
Poverty rate(num persons in poverty per capita)	-23.60***	-5.08***	-4.87***	-3.42**	-1.97	-5.69**
	(3.10)	(1.66)	(1.61)	(1.64)	(2.55)	(2.21)
Post secondary (PS) institution enrollment per capita (X0.01)	0.00	0.03	0.04	0.02	-0.02	0.15
	(0.10)	(0.04)	(0.04)	(0.04)	(0.03)	(0.21)
Number of research universities per capita (X100,000)	0.08*	0.04	0.04	0.03	0.03	0.04
	(0.04)	(0.03)	(0.03)	(0.03)	(0.05)	(0.03)
Land area (sq. miles)(X1000)	0.27***	0.04	0.04*	0.04*	0.07**	0.03**
	(0.05)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
Cleantech cluster(location quotient)	0.24	0.40	0.41	0.53	0.34	0.67
	(0.53)	(0.37)	(0.35)	(0.33)	(0.44)	(0.44)
Investment of risk capital in cleantech	-0.07	0.14	0.16	0.21	0.31	0.14
	(0.30)	(0.36)	(0.34)	(0.35)	(0.44)	(0.48)
State enforcement of non-compete covenants	-0.05	-0.07**	-0.07**	-0.06*	-0.07	-0.02
	(0.05)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)
State renewable energy production ratio	-0.25	0.27	0.20	0.06	-0.07	-0.16
	(0.30)	(0.27)	(0.25)	(0.28)	(0.29)	(0.54)
State implementation of renewable portfolio standards	0.16	0.09	0.10	0.13	0.18	0.15
	(0.18)	(0.11)	(0.11)	(0.11)	(0.12)	(0.16)
State policies for renewable energy	0.03***	0.01**	0.01**	0.01*	0.01***	0.00
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Specialization of the regional knowledge base in cleantech				-11.12***		
				(3.28)		
Constant	1.82**	-2.33***	-1.99***	-1.71***	-2.32***	-2.38***
	(0.85)	(0.58)	(0.55)	(0.51)	(0.80)	(0.56)
Observations	2,359	2,359	2,359	2,349	1,178	1,171
Year dummies included	Yes	Yes	Yes	Yes	Yes	Yes
χ^2	198.2	903.9	1096	1263	992.2	757.6
Number of MSAs	237	237	237	236	136	135

Robust standard errors in parentheses. All covariates were lagged by one year.
*** p<0.01, ** p<0.05, * p<0.1

In model 1, where only control variables were included, I find that rates of new venture creation in cleantech are higher in larger MSAs (*land area:* b=0.27, p<0.001), more affluent MSAs (*poverty rate:* b=-23.6, p<0.001), MSAs with fewer residential permits (*residential permits per capita:* b=-0.62, p=0.009), and MSAs with more research universities per capita (*research universities per capita:* b=0.08, p=0.05). At the state-level, I find that state *implementation of renewable portfolio standards* has a significant positive effect (b=0.03, p<0.001).

In model 2 I introduced the variable to capture potential regional knowledge spillovers (i.e. stocks of new knowledge). The model is a level-log regression, where changes in the dependent variable can be interpreted in terms of % changes in the independent variable (i.e. Δ new venture created=β*Δ% patents granted). I find that the coefficient associated with the knowledge stock variable is positive and statistically significant (b=0.92, p<0.000), thus providing strong support for hypothesis 1. In practical terms, this means that a doubling in the rate of cleantech innovation in a MSA (i.e. a 100% increase in the number of cleantech patents granted) leads to the creation of 0.92 new cleantech ventures. Note that while I did not focus on the main effects of community environmental ideologies, I find that they have a positive but marginally significant impact on cleantech new venture creation rates (b=0.30, p=0.08).

In model 3 I examined the interaction between stocks of new knowledge and environmental ideologies in an MSA. Consistent with hypothesis 2, I find evidence of a positive interaction effect between the knowledge stock and ideology variables (b=0.33, p=0.005). Since the ideology variable is scaled so as to range from 0-1, the coefficient can be directly interpreted as the difference in slope between regions on the two extremes of the ideology continuum (i.e. very strong vs. very weak). In practical terms this means that a doubling in the rate of cleantech innovation (i.e. a 100% increase in the number of cleantech patents granted)leads to the creation of 0.3 more ventures in MSAs with strong environmental ideologies relative to those with weak environmental ideologies. I illustrate this relationship graphically in figure 1 below.

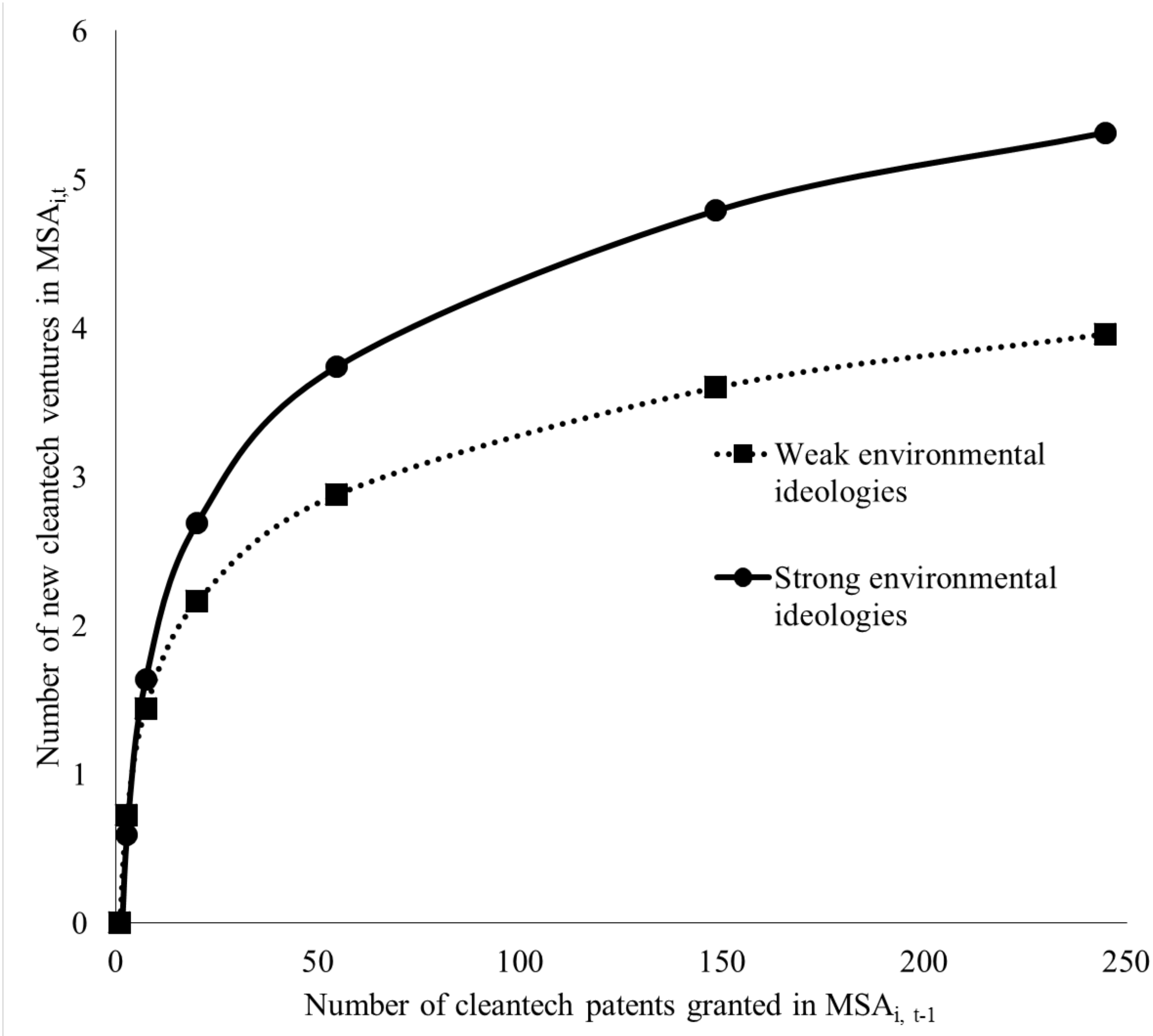

Figure 1 (paper 1). Positive moderating impacts of community environmental ideologies on the relationship between regional cleantech innovation rates and cleantech new venture creation

Lastly in models 4 to 6, I examined the impacts of the specialization of the regional knowledge base in cleantech as a boundary condition regulating the interaction between stocks of new knowledge and community environmental ideologies. In model 4, I introduced the regional knowledge base specialization variable to understand its marginal effect. I find that it has a negative and significant impact on cleantech new venture creation rates (b=-11.12, p=0.001). Since this variable is a ratio that ranges from 0 to 1, this means that each percent increase in the specialization of the regional knowledge base in cleantech decreases the number of new ventures created by 0.1. Next in models 5 and 6 I used a sub-sample approach with a median split to test hypothesis 3. That is, I studied the strength of the moderating effects of ideologies when the specialization of the regional knowledge base in cleantech was <50% (model 4) and greater than >50% (model 5) of the median level of cleantech specialization in the entire sample (the median value of this variable was 0.047(4.7%) and the mean value was 0.051 (5.1%) with a standard deviation of 0.02 (2%)). As hypothesized, I observe that both the strength and statistical significance of the interaction effect declines as a function of the specialization of the regional knowledge base in cleantech. When cleantech specialization levels are low the interaction between stocks of new knowledge and community environmental ideologies is positive and statistically significant (b=0.36, p=0.006). In contrast, when cleantech specialization levels are high the interaction between stocks of new knowledge and community ideologies become weak and statistically insignificant (b=-0.03, p=0.85). I confirmed these findings using a formal test of the difference in model coefficients with a seemingly unrelated estimation equation. This allowed me to statistically reject the null hypothesis that the coefficients of the interaction terms in models 5 and 6 were statistically equivalent (χ^2 (1) = 4.06, p=0.04). In sum, I therefore find strong statistical support for hypothesis 3. These relationships are illustrated graphically in figure 2 below.

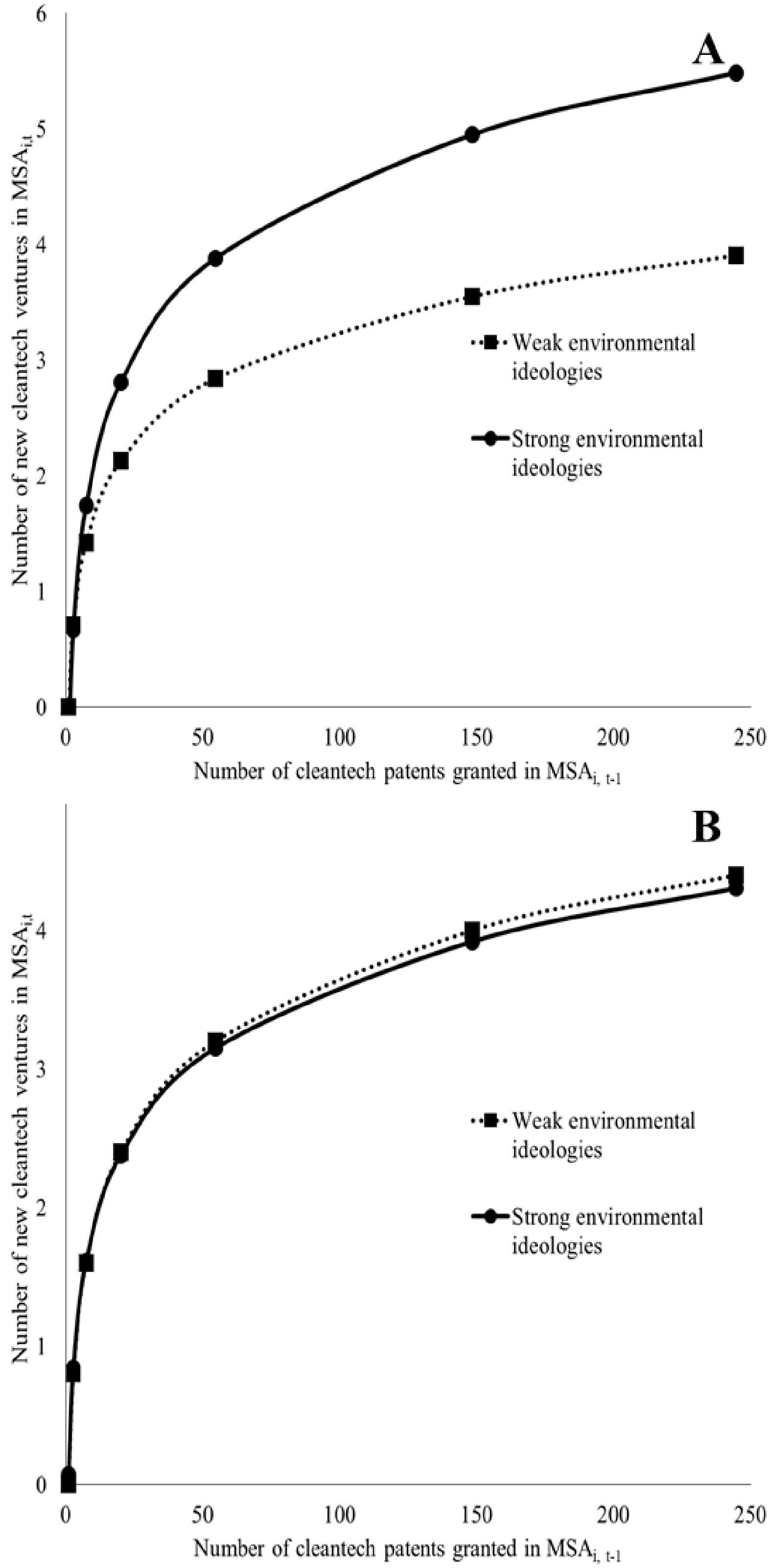

Figure 2 (paper 1). The interaction effect between community environmental ideologies and regional rates of innovation when the specialization of the regional knowledge base in cleantech is (A) low and (B) high. The positive interaction effect occurs when the specialization of the regional knowledge base in cleantech is low.

Sensitivity Analyses

I also carried out a series of sensitivity analyses to ensure that the results were robust to alternate variable and model specifications. I present the results from some of these tests in table 4. Note that I re-estimated all specifications in the context of a full model (i.e. model 3, table 3).

Table 4 (paper 1). Sensitivity analyses

	Alternate temporal lags			Alternate variables				Alternate model specifications	
	Model 1[a]	**Model 2[a]**	**Model 3[a]**	**Model 4** Depreciated Knowledge stocks	**Model 5** State dummies	**Model 6** real GDP per capita	**Model 7** Employment based cluster LQ	**Model 8[f]** Population averaged neg binomial	**Model 9[g]** Mixed hierarchial linear model
VARIABLES	2 year lag	3 year lag	4 year lag	stocks	dummies	capita	LQ	neg binomial	linear model
Stocks of new knowledge (ln # of patents granted)[b,g]	0.67***	0.68***	0.70***	0.80***	0.74***	0.68***	0.69***	0.68***	0.10***
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.07)	(0.02)
Supportive (enviromental) ideologies	-0.36	-0.37	-0.12	-0.47	-0.37	-0.28	-0.41	-0.34	-0.16**
	(0.34)	(0.34)	(0.37)	(0.33)	(0.29)	(0.29)	(0.32)	(0.29)	(0.07)
Stocks of new knowledge*Supportive ideologies[b]	0.36***	0.39***	0.35***	0.31**	0.31***	0.22*	0.30**	0.26**	0.15***
	(0.12)	(0.12)	(0.13)	(0.12)	(0.11)	(0.12)	(0.12)	(0.12)	(0.03)
Residential permits (per capita)	0.33**	0.38***	0.44***	0.39***	-0.13	0.26**	0.30**	0.10	-0.07**
	(0.13)	(0.13)	(0.13)	(0.12)	(0.14)	(0.13)	(0.13)	(0.11)	(0.03)
Poverty rate(num persons in poverty per capita)	-5.70***	-5.14***	-4.88***	-4.42***	-8.49***		-4.93***	-5.54***	-1.91***
	(1.76)	(1.72)	(1.77)	(1.66)	(1.98)		(1.65)	(1.43)	(0.52)
Post secondary (PS) institution enrollment per capita (X0.01)	0.04	0.08**	0.14**	0.04	0.00	0.05	0.04	0.03	0.01
	(0.04)	(0.04)	(0.06)	(0.04)	(0.03)	(0.03)	(0.04)	(0.02)	(0.01)
Number of research universities per capita (X100,000)	0.04	0.02	0.03	0.04	0.04	0.03	0.03	0.04	0.01
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.01)
Land area (sq. miles)(X1000)	0.05**	0.05**	0.05**	0.03	0.03	0.04**	0.04*	0.06***	0.04***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)
Cleantech cluster(location quotient)[e]	0.26	0.40	0.28	0.49	0.15	0.38	0.49**	0.37	-0.03
	(0.38)	(0.37)	(0.36)	(0.35)	(0.37)	(0.33)	(0.20)	(0.29)	(0.08)
Investment of risk capital in cleantech	-0.01	0.05	0.26	0.06	-0.01	0.37	0.13	0.11	-0.03
	(0.42)	(0.32)	(0.31)	(0.33)	(0.30)	(0.29)	(0.34)	(0.27)	(0.06)
State enforcement of non-compete covenants	-0.05	-0.06*	-0.07*	-0.07**	-0.00	-0.05	-0.05	-0.07*	-
	(0.03)	(0.04)	(0.04)	(0.03)	(0.08)	(0.03)	(0.03)	(0.04)	-
State policies for renewable energy	0.01***	0.01*	0.01	0.01**	-0.00	0.01**	0.01	0.00	-
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	-
State renewable energy production ratio	0.08	0.07	0.04	0.26	-2.55**	0.25	0.16	0.25	-
	(0.25)	(0.25)	(0.26)	(0.25)	(1.01)	(0.24)	(0.26)	(0.24)	-
State implementation of renewable portfolio standards	0.13	0.12	0.05	0.09	0.01	0.09	0.12	0.14	-
	(0.11)	(0.11)	(0.12)	(0.11)	(0.09)	(0.11)	(0.11)	(0.09)	-
Real GDP per capita[d]						28.90***			
						(6.07)			
Constant	-2.33***	-3.29***	-6.50***	-2.24***	-1.42**	-4.04***	-2.55***	-1.75***	0.38***
	(0.58)	(0.59)	(0.85)	(0.56)	(0.66)	(0.42)	(0.39)	(0.48)	(0.13)
Observations	2,359	2,359	2,359	2,349	2,359	2,359	2,359	2,359	2,359
Year dummies included	YES	YES	YES	YES	YES	YES	YES	YES	YES
Number of MSAs	237	237	237	237	237	237	237	237	237
State dummies included	NO	NO	NO	NO	YES[c]	NO	NO	NO	NO

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

[a]Models 1-3 use temporal lags of 2, 3 and 4 years respectively

[b]Model 4 uses a weighted average depreciated metric of knowledge stocks

[c]Model 5 included state dummies

[d]Model 6 uses real GDP per capita instead of poverty rate per capita

[e]Model 7 uses an employment based location quotient metric instead of an establishment based measure to compute cleantech cluster levels

[f]Model 8 uses a population-averaged negative binomial specification

[g]Model 9 uses a multilevel mixed model specification nesting MSAs with states. It uses a log-log specification instead of a linear-log specification.
Coefficients are not directly estimated for state level variables

Alternate variables

Longer temporal lags: Since I estimated the main models with a one year temporal lag, I re-estimated equations with longer lags (2-4 years). This accounts for the fact that there might be a longer window of opportunity for exploitation of latent opportunities within the regional knowledge base by new entrants. While the availability of some of the model control variables started in 1999 and ended at 2010 hence limiting main models to that window, I was able to explore these longer lags without loss of data by coding entry events till the end of 2013. I find that results using these longer lags are consistent with the main models (coefficients of interaction terms ranged from 0.35-0.39, p<0.001). I present results using these lags of 2, 3, and 4 years in models 1-3 of table 4.

Depreciated average of patent stock: I re-estimated models using a temporally lagged depreciated average stock of knowledge as the main independent variable. I computed a weighted average by depreciating knowledge stocks by 0.2 each year and then log-transformed this measure. More specifically, I used the equation shown below to compute the weighted depreciated stock of knowledge for an MSA *i* in a given year *t:*

$$Patents\ granted_{i,t} = (patents\ granted_{i,t} + 0.8 * patents\ granted_{i,t-1} + 0.6 * patents\ granted_{i,t-2} + 0.4 * patents\ granted_{i,t-3} + 0.2 * patents\ granted_{i,t-4})/3$$

Lastly, I log-transformed this weighted average since it was right-skewed. As shown in model 4 of table 4, results with this alternate specification of regional knowledge stocks were consistent with the main models (coefficient of interaction term is 0.31, p<0.01).

State dummies: I computed a set of models including state dummy variables. Given that the primary unit of analysis was the MSA, I did not include these controls in the main models. As shown in model 5 of table 4, results were consistent with the main models (coefficient of interaction term is 0.31, p<0.001).

Real GDP per capita as an alternate indicator of metro affluence: I replaced the poverty measure of MSA affluence with a metric based on average personal income per capita. As reported in model 6 of table 4, I find that the results are once again largely consistent with the main models although the interaction term is now slightly smaller ($b=0.22$, $p<0.05$).

Cleantech cluster location quotient calculated using employment: Since the control variable for cleantech clustering is necessarily an approximate estimate given how “clean vs. non-clean” NAICS sectors are classified (see footnote 4), I also estimated an alternate location quotient metric using employment instead of establishment data. As reported in model 7 of table 4, results are similar to the main models with the interaction term positive and statistically significant ($b=0.30$, $p<0.01$).

Alternate model specifications

Model specified using a GEE population-averaged estimator: As an alternate approach to a negative binomial random effects specification, I also tried a generalized estimating equation (GEE) population-averaged negative binomial model. These set of semi-parametric models estimate coefficients by averaging over random effects rather than holding them constant (Allison, 2009; Land, McCall, and Nagin, 1996). Conceptually, this therefore allows for an estimation of the impacts on the average “subject” (i.e. MSA in this study context) within the sample. As reported in model 8 of table 4, I find results consistent with the main model (coefficient of interaction term is 0.26, $p=0.027$).

Mixed model specification: Given that the main models ignore the fact that the data is hierarchically nested (i.e. MSAs are located within states), and that I included control variables from both levels, I also tried a multi-level model specification. I log transformed the dependent variable for this purpose so that I could execute a linear mixed model in my statistical package of choice (Stata 12.1)[5]. Unlike a random effects specification, this specification allows for both random intercepts and random slopes. It directly estimates the lower level (i.e. MSA) coefficients while estimating variances and co-variances of higher level (i.e. state) parameters. As reported in model 9 of table 4, I again find a positive and significant interaction effect between knowledge spillovers and community ideologies ($b=0.15$, $p<0.001$). Note that the coefficient in this model is different from the other models, since this is a log-log specification (i.e. Δ% new ventures created=$\beta * \Delta$% patents granted) as opposed to a linear-log specification.

DISCUSSION

The KSTE suggests that regional differences in the innovation activities of marketplace incumbents can have a significant impact on the spatio-temporal distribution of entrepreneurial opportunities, and result in regional differences in new venture creation rates (Acs *et al.*, 2009; Agarwal *et al.*, 2007; Audretsch and Keilbach, 2007). However, in doing so, it implicitly assumes a strong form of distributed agency (Garud and Karnøe, 2003) where risk-bearing entrepreneurs are entirely aware of and willing to exploit latent opportunities not pursued by incumbent organizations (Acs *et al.*, 2009; Kirzner, 1979). In this study, I relax this assumption and demonstrate that regional differences in sociopolitical ideologies can have a significant impact on the efficiency of KSTE dynamics. In the context of the cleantech industry sector, I find that the relationship between incumbent innovation and new venture creation is positively moderated by the strength of community environmental ideologies. Furthermore, I find that this moderating effect of ideologies is finite and restricted to locations where innovation activities in cleantech are still relatively nascent. That is, as regions specialize in cleantech such that the

[5] Stata routines to execute negative binomial mixed models (command: menbreg) are only available in the latest version of Stata (v13). The version of Stata that is available on the research server at CU Boulder (12.1) does not have that functionality. I intend to try such a model prior to journal submission.

regional knowledge base contains a higher proportion of cleantech patents, environmental ideologies become less important in facilitating KSTE dynamics within the cleantech sector.

These findings have implications for entrepreneurship theory and practice. From a theoretical perspective, I demonstrate that regional differences in new venture creation are best understood when the combined influences of both economic and sociocultural factors are taken into account. As McMullen and Shepherd (2006) indicate in discussing the distinction between third-party and first-party opportunities, entrepreneurial action requires *both* knowledge (i.e. awareness) and motivation. Thus while KSTE with its emphasis on the economics of knowledge and the localization of spillovers specifies how opportunities originate, it ignores the motivational component of entrepreneurial action. Conversely, institutional theories of entrepreneurial action take the existence of opportunities as a given (Baumol, 1996) or "manna from heaven" (Agarwal *et al.*, 2007: 265), and focus instead on how social forces can shape *ex-post* motivations to pursue opportunities (Tolbert *et al.*, 2011). By acknowledging that both knowledge and motivational factors are important aspects of entrepreneurial action, the model in this study effectively combines both these perspectives.

These findings also indicate that rates of new venture creation increase non-linearly as a function of incumbent innovation rates. In practical terms I found that the doubling of the innovation rate in a region creates approximately 0.9 new ventures (model 2, table 3) and 0.3 additional ventures in regions with strong environmental ideologies (model 3, table 3). While once could potentially argue that the practical importance of these effect sizes that are relatively small, I suggest that is not the case. The approach in this study essentially looks at first-order effects, which is the direct impact of innovation rates on new venture creation. However, an important outcome from the enhancement of KSTE dynamics is the acceleration of regional innovation trajectories and establishment of positive feedback cycles such as knowledge spill-ins (Agarwal *et al.*, 2007; Kotha, 2010). An increase in the rate of entrepreneurship can also lead to important demonstration effects spurring the growth of entrepreneurial activity in a non-linear manner (Samila and Sorenson, 2010, 2011). In effect, these small effects can get compounded over time leading to large regional differences over time. I thus suggest that these empirical findings necessarily represent a lower bound on the impacts that regional innovation activities and ideologies have on new venture creation rates in cleantech.

The finding that community ideologies positively moderate KSTE dynamics provide support for the often-stated but infrequently tested argument that localized institutional forces can shape location-based externalities. These findings are also therefore consistent with the literature in economic geography which suggests technological knowledge spillovers are not frictionless and automatic within regional economies (Huber, 2012; Lambooy, 2010). Moreover, it also raises some interesting possibilities for future research. For instance, prior research on the evolutionary dynamics of industrial clusters indicates that they are primarily created and sustained by the geographically localized entry of spinoff firms from local incumbent organizations (Klepper, 2009; Sorenson and Audia, 2000). However since founders of spinoff firms are already likely to be aware of latent entrepreneurial opportunities that they have co-created within incumbent organizations (Agarwal *et al.*, 2007; Lockett *et al.*, 2005; Shane, 2000), one might expect that community ideologies should be relatively less important for the opportunity recognition process for spinoffs relative to other kinds of entrants (e.g. de novo firms, entrepreneurs that relocate from other regions). One might therefore also expect that the relative importance of spinoff firms to the formation and maintenance of industrial clusters should be lower in locations where strong ideologies are prevalent, whereas other entrants that strategically choose locations (Aharonson *et al.*, 2007; Alcácer and Chung, 2007; Chung and Alcácer, 2002; Pe'er *et al.*, 2008) should play a larger role. Furthermore, while the focus in this study is necessarily on entry dynamics, community ideologies can also potentially have important implications on the post-entry performances of new ventures. For instance, while extant research suggests that spinoff firms typically inherit routines from incumbents thus providing them with survival advantages over other kinds of entrants (Agarwal *et al.*, 2004; Buenstorf and Klepper, 2009), strong community ideologies might dampen such competitive advantages by facilitating the transmission and exchange of information among actors. These are all interesting possibilities that could be examined in future research.

Perhaps the most interesting finding from this work is that the importance of environmental ideologies in shaping KSTE dynamics declines as the regional knowledge base becomes more specialized in cleantech. Theoretically this finding is important as it points to the evolving impacts of sociocultural institutions on entrepreneurial entry dynamics within industries. More broadly, this also suggests that institutional forces can potentially substitute or complement each other over time. This is an interesting and important implication, particularly in terms of public policy initiatives that seek to stimulate entrepreneurship and develop regional clusters focused on specific technologies (Cooke, 2008). For instance it suggests that when the regional knowledge base is relatively unspecialized in the relevant technological domain, policy makers should pay more attention to and if possible attempt to influence the socio-political attitudes in a region. They might be able to do so directly by shaping the rhetoric within communities (Davies, 2013), or indirectly by encouraging the formation of organizations that have the ability to influence patterns of social and political engagement such as specialized technology oriented social movements (Hess, 2005; Pacheco *et al.*, 2014). To the extent that regions can strongly identify with specific technological domains (Appold, 2005; Romanelli and Khessina, 2005), the positive processes of creative construction (Agarwal *et al.*, 2007) highlighted by KSTE are more likely to take hold. However, once the regional knowledge base is more specialized and ideologies play less of a role, policy makers would be better of investing resources more directly into knowledge creation activities (i.e. public and private R&D) and motivating entrepreneurs through economic incentives.

This study is also not without its limitations. First, the data does not allow me to identify different kinds of entrepreneurial entrants into the cleantech sector (e.g. spinoff firms, de novo firms, diversifying entrants). While this approach is consistent with prior work using a KSTE framework (Acs *et al.*, 2009; Audretsch and Keilbach, 2008; Plummer and Acs, 2014), it does mean that I cannot disentangle the impacts of community ideologies on these different entrant populations. Such an analysis would be particularly useful and interesting to conduct in future research. Second, the use of patents is a noisy measure of the latent opportunity set region. Not all innovations are patented and not all patents are necessarily economically useful (Arrow, 1962). Yet, despite the obvious limitations of patent data, they do allow for a historical examination of regional innovation dynamics and their influence on new venture creation rates (Acs *et al.*, 2009; Plummer and Acs, 2014). Third, the use of the cleantech sector might also raise concerns about the generalizability of the impacts of sociopolitical community ideologies. While environmental ideologies are necessarily specific to the cleantech sector, I developed the theoretical arguments on the impacts of sociopolitical ideologies in a more general fashion. I therefore expect that these results should generalize well to other technologically intensive and value-laden contexts where both knowledge spillovers and sociopolitical ideologies might play an important role.

Concluding remarks

This study highlights the complex interplay between new ventures, incumbent innovation, and community environmental ideologies in the emerging cleantech sector. I do so by combining perspectives from the KSTE and sociological institutional theory under a discovery-based framework on entrepreneurial opportunity. Given the increasing interest among academics, practitioners and regional policy makers in transitioning towards a “greener economy”, I hope that these findings are an initial step in explaining regional variations in entrepreneurial dynamics in this sector.

PAPER 2: EXPLAINING PERFORMANCE DIFFERENTIALS AMONG VENTURE-BACKED STARTUP FIRMS: HOW MUCH DOES REGION MATTER?

ABSTRACT

I use a variance decomposition approach to assess the extent to which regional factors explain performance differentials among venture-backed startups. Drawing on evolutionary arguments, I also examine the contingent effects of business development stage and industry sector maturity on the magnitude of the "region effect". My findings indicate that regional factors have significant performance impacts on venture-backed startups that are at the earliest stages of development (explaining 14% of performance variance), and for startups that operate in nascent industry sectors (explaining 7% of performance variance). However regional effects are relatively less important for venture-backed startups with more developed business models, and for those that operate in mature industry sectors (explaining 2-3% of performance variance). I offer implications for theory, practice, and policy.

INTRODUCTION

Venture-capital backed startups are a critical component of the U.S. economy. Despite comprising less than 1% of total national entrepreneurial activity, they have a large and disproportionate economic effect stimulating innovation, job creation, and overall wealth creation (Gompers and Lerner, 2001; Kortum and Lerner, 2000; Shane, 2009). Extant research has investigated a number of factors that lead to performance differentials among these firms. For example, some studies have focused on attributes of the founding team such as entrepreneurial experience and human capital (Colombo and Grilli, 2010; Hsu, 2007). Others have instead focused their attention on the contributions of venture capitalists (VCs) to these startups (Baum and Silverman, 2004; Brander, Amit, and Antweiler, 2002; Fitza *et al.*, 2009; Nisar, 2005). From a theoretical perspective much of the extant literature has by-and-large adopted a resource-based lens (Barney, 1991), and focused on factors internal to these startups to explain performance differentials.

In contrast, scholars have paid comparatively less attention to factors outside firm boundaries that might also matter. In particular, in this study my interest lies in understanding whether regional factors, that is those specific to the geographic location that a venture-backed startup is situated in, can act as a significant basis of competitive advantage for these firms. Extant theory suggests that this should indeed be the case. For example, the relational view on strategy indicates that a firms' locus of competitive advantage is at last partially embedded in idiosyncratic, cooperative inter-firm relationships and routines (Dyer, 1996; Dyer and Singh, 1998), often between trading partners (e.g. alliance partners, buyers and suppliers) located in close geographical proximity to each other (Uzzi, 1996). This view is also consistent with perspectives on inter-organizational learning in both strategy and economic geography that emphasize primacy of knowledge exchange between actors in a regional economy as a basis of competitive advantage (Almeida *et al.*, 2003; Almeida and Kogut, 1999; Capello, 1999; Gnyawali and Srivastava, 2013; Lawson and Lorenz, 1999; Morgan, 2004; Whittington, Owen-Smith, and Powell, 2009), particularly for firms such as venture-backed startups that operate in dynamic and technology intensive sectors of the economy (Davila, Foster, and Gupta, 2003; Mann and Sager, 2007). And yet, the evidence linking regional factors to the performance of venture-backed startups remains surprisingly limited.

More specifically, the literature in this area is particularly lacking in two main aspects. First, much of the existing evidence around regional advantages for VC-backed startups is largely either based on detailed qualitative analysis of a small number of cases (Saxenian, 1996) or anecdotal evidence (Feld, 2012). These findings, while informative on the mechanisms at work within specific entrepreneurial hubs (e.g. Silicon Valley, Austin, Boulder, Route 128), lack large-scale generalizability. Second, the existing empirical work studying the impacts of specific regional factors (e.g. knowledge stocks, geographic clustering, social networks) on the performance of venture-backed startup firms has almost exclusively focused on the analysis of a single industry, biotechnology, and/or used samples of established (i.e. post IPO) VC-backed firms (e.g., Decarolis and Deeds, 1999; Folta *et al.*, 2006a; Gilbert *et al.*, 2008; McCann and Folta, 2011). Hence, as a whole, scholars have a limited understanding of both firm and industry-specific contingencies that might make regions less or more important as a source of competitive advantage for venture-backed startups.

To address these shortcomings in the literature, I use a variance decomposition methodology originally developed in the corporate strategy literature (e.g., McGahan and Porter, 2002; Rumelt, 1991) but increasingly used in entrepreneurship (Fitza *et al.*, 2009; Short *et al.*, 2009) to assess the degree to which regional factors can explain performance differentials among venture-backed startups. To be clear, my focus is not on whether regions should matter in this context; extant theory and evidence both suggest that they should to some extent. Rather I am principally interested in quantifying the degree to which regions matter, as well as their importance relative to other sources of competitive advantage (i.e. firm, VC investors, industry affiliation). I carried out analyses using a, comprehensive sample of U.S. venture-backed startups operating in the U.S. economy over the period 1980-2012. Furthermore, drawing on prior literature that has suggested that the importance of regional factors might also be contingent on evolutionary firm and industry-level dynamics, I also examined whether the identified “region effect" varies as a function of: a) the stage of business development of the startup (Ruhnka and Young, 1987) and b) the maturity of the industry segment within which the startup operated.

This study makes three main contributions to the literature. First, since the purpose of the variance decomposition approach is primarily to test existing theory (Hambrick, 2007; Pfeffer, 2007), it is well suited for studies that seek to replicate and extend existing research in new contexts (Short *et al.*, 2009). By studying regional effects in the context of venture-backed startup firms I therefore add to the very limited body of work that has looked at the impacts of regional factors in the variance decomposition literature, using samples of more established companies in other sectors of the economy and in non U.S. contexts (Chan, Makino, and Isobe, 2010; Ma, Tong, and Fitza, 2013). My examination of regional effects at different stages of business development and industry maturity is also novel in this literature stream. Second, this study also adds to the literature on the growth of venture-backed startup firms, by examining the role of geography in this context. While a number of studies have examined the impacts of *receiving* venture-capital financing on startup growth (Bertoni, Colombo, and Grilli, 2011; Colombo and Grilli, 2010; Davila *et al.*, 2003; Hsu, 2007; Tian, 2011), there have been surprisingly few studies that have specifically focused on the performance impacts of geography within a comprehensive sample of VC-backed startups. Furthermore as indicated previously, due to the specifics of prior research designs, the few studies that do exist are limited in their generalizability either in terms of the sample of startups examined or the number of regions studied. For instance, while Gilbert *et al.* (2008) look at a number of different regions, they do so in the context of mature (i.e. post IPO) biotechnology firms. Similarly, Anna-Lee Saxenian's (1996) now classic study on regional advantages is based on a detailed qualitative comparative analysis of two regions, Silicon Valley and Route 128 in the Boston area. Hence, while the use of the variance decomposition methodology does not allow me to identify the specific causal mechanisms that these prior studies explore (Fitza *et al.*, 2009), I am able to research a comprehensive sample of venture-backed startups operating in a variety of industry segments. In doing so, I am able to add to and complement existing work by generating generalizable and useful stylized facts (Helfat, 2007), while also qualifying my results by including firm and industry specific contingencies. Finally, this research also has important policy and practice implications, as it is among the first to provide large-scale empirical validation for efforts that seek to promote regional economic development through promoting venture-backed entrepreneurship (Gilson, 2003; Isenberg, 2010; Mason and Brown, 2013a; Venkataraman, 2004; Vogel, 2013). The findings from this study should help qualify and justify these efforts, as in the absence of a rigorous demonstration of the link between regions and the growth dynamics of startup firms much of the rationale for such approaches has been either theoretical (Venkataraman, 2004) or based on supportive but largely anecdotal evidence (Feld, 2012). And lastly, from a practical standpoint, these results should also be of interest to both venture capitalists and VC-backed startups, as they provide guidance on the conditions under which regional factors are likely to add the most value to these startups.

THEORY AND HYPOTHESES DEVELOPMENT

Why should regions matter for the growth of firms more generally, and for VC-backed startups more specifically? To answer this question, I primarily draw on resource-based perspectives on competitive heterogeneity in both strategy and economic geography, which suggest that firms can derive competitive advantages from the geographic regions that they are located in. This can occur when: a) locational resources are “sticky” in that they are imperfectly mobile (Hoopes, Madsen, and Walker, 2003; Krugman, 1991) and hence not accessible to firms *across* regions b) and even *within* a given region, if firms differ in their need, motivation, and ability to leverage and bundle such external resources with firm-specific assets, particularly as they mature and develop internal firm-specific capabilities (Almeida *et al.*, 2003; Bradley *et al.*, 2011; McCann and Folta, 2011; Sirmon, Hitt, and Ireland, 2007; Srivastava and Gnyawali, 2011).

The role of "sticky" resources in explaining region effects

Drawing on insights from the seminal work by Penrose (1959) on the growth of the firm, scholars in the resource-based tradition of strategic management have emphasized that the competitive heterogeneity between firms might, at least in part, be explained by distortions or imperfections in underlying factor markets, such as resource "stickiness" (Mishina, Pollock, and Porac, 2004). Furthermore, to the extent that these imperfections persist over time they can lead to sustained performance differences between firms (Peteraf, 1993). Location-based resources are considered to be sticky when they are imperfectly mobile across geographic space, and hence not accessible to firms across regions (Krugman, 1991). In the context of VC-backed startups that are often involved in the commercialization of novel technologies in technology focused industries, one important resource with sticky characteristics is knowledge (Krugman, 1991). I therefore use it as an example to illustrate how resource stickiness can contribute to regional differences in startup performance.

For VC-backed startups, the key "raw material" is typically specialized technological and managerial knowledge. Such knowledge is however regionally sticky for a number of reasons. First, with respect to technological knowledge, much of the components are largely embodied within individuals with specialized skills (e.g. inventors, scientists) who while mobile within regions, are typically relatively immobile across them (Almeida and Kogut, 1997, 1999; Jaffe, Trajtenberg, and Henderson, 1993). This means that in order to access these competencies and benefit from localized knowledge spillovers, firms need to necessarily locate in specific regions where such knowledge is accessible (Audretsch *et al.*, 2005). Furthermore the interactions between these startups and other actors within the regional economy means that the knowledge that is generated by these ventures is typically also socially created and constructed in a spatially localized manner (Amin and Cohendet, 2005). The interpretation and implementation of such co-created technical knowledge also requires a relational form of understanding that is typically location specific (Storper, 1995; Tallman *et al.*, 2004). For example, an inventor working within a venture-backed startup is likely to either draw upon and/or collaborate both formally and informally with colleagues outside the firm but within regional inventor networks. Effective communication and exchange of ideas among these inventors requires a high degree of spatial proximity (Fleming and Frenken, 2007; Fleming *et al.*, 2007). It is no coincidence that private and public institutions that create knowledge such as universities and R&D labs are often at the heart of successful clusters of VC-backed startup firms (Audretsch *et al.*, 2005; Feldman, 2001).

Second, in addition to technological knowledge, these startups also need specialized managerial skills, such as guidance on how to operate and navigate in highly uncertain and dynamic environments. While experienced VC investors do add value in that aspect (De Clercq *et al.*, 2006; Gorman and Sahlman, 1989; Matusik and Fitza, 2012; Sapienza, 1992), access to these investors typically requires co-location within regional clusters (Powell *et al.*, 2002; Zook, 2002). This is because due to the substantial uncertainty and risks associated with these investments investors largely prefer to invest in startups in close geographic proximity that they can monitor closely (Cumming and Dai, 2010). Furthermore, beyond such formal investor relationships, the level of "know-how" (Garud, 1997) in the region also matters. For example, in a detailed case study across a number of startup clusters, Bresnahan and colleagues (Bresnahan, Gambardella, and Saxenian, 2001) find that while traditional agglomeration economies often arise naturally as clusters grow, "old economy knowledge" such as generalized firm-building capabilities and managerial skills, is not only crucial for startup success but a relatively rare resource across regions. Furthermore such knowledge is highly tacit, and largely embedded within the informal relational social networks among actors within in a region (Bresnahan *et al.*, 2001), making it both regionally sticky, causally ambiguous and hence hard to imitate across regions (Lawson, 1999). Synthesizing these arguments, I therefore hypothesize, that in the context of venture-backed startups:

Hypothesis 1: A significant portion of variation in venture-backed startup performance is attributable to regional effects.

Firm and industry contingencies that impact regional effects: An evolutionary perspective

While I expect to isolate regional effects that should impact the performance of venture-backed startups, I also anticipate that these effects should be contingent on both firm- and industry specific conditions (McCann and Folta, 2011; Shaver and Flyer, 2000). I discuss each of these in turn below.

As firms' mature, their basis of competence, value creation, and the process by which they orchestrate external (e.g., locational) resources changes (Sirmon *et al.*, 2011, 2007). For instance, at the nascent stages of formation firms internal resources (e.g., employees) are generally limited, and organizational routines are not well developed. With respect to venture-backed startups in particular, such "seed stage" firms typically also do not have a well-defined business model or compete in product markets, and are still willing to still explore multiple opportunities (De Clercq *et al.*, 2006; Dimov and Murray, 2008; Ruhnka and Young, 1987).

Under such conditions, the need to access resources such as knowledge from other actors in the regional environment is particularly high (Lawson and Lorenz, 1999). For instance, startups might be able to initially gain access to the resources in the local community, by leveraging the pre-existing social ties of founding team members (Beckman, 2006). This ability can also serve as a key basis of differentiation for firms at this stage of development. Access to regional resources might allow these nascent startups to identify opportunities and applications for the technology being pursued, beyond those originally conceived (Shane, 2000). Beyond such technical knowledge, such nascent firms are also likely to benefit significantly from the general managerial know- how in the region (Bresnahan *et al.*, 2001), for instance by seeking advice from mentors in the local entrepreneurial community (Lafuente, Vaillant, and Rialp, 2007) or attending local networking events (Feld, 2012).

However, as these venture-backed startups become more established, I expect that both their relative dependence and their motivation to draw on regional knowledge should decline. I discuss this with regards to their use of both technological knowledge and managerial “know-how” in the regional knowledge base. First, from the perspective of technological knowledge, mature startups should have stronger internal R&D capabilities relative to more nascent firms (Pisano, 1994), as well as more firm-specific technological knowledge (Grant, 1996) that can serve as a basis of competitive advantage. They are therefore more likely to innovate by recombining internal knowledge in novel ways (Katila, 2002), and be less motivated to rely on the regional knowledge base (Almeida *et al.*, 2003; Srivastava and Gnyawali, 2011). At the same time, such firms are more likely to be able to source knowledge from geographically distant locations (Phene, Fladmoe-Lindquist, and Marsh, 2006), for instance by using formalized relational mechanisms such as cross-regional alliances or hiring migrant employees (Almeida, Song, and Grant, 2002; Rosenkopf and Almeida, 2003; Song, Almeida, and Wu, 2003). Hence by broadening the scope of their knowledge acquisition activities through both internal development as well as contractual mechanisms that allow them to overcome regional knowledge barriers, their reliance on the regional knowledge base as the dominant source of technical knowledge and a basis of competitive advantage is decreased (Coombs, Deeds, and Ireland, 2009).

Second, with respect to managerial know-how, more mature startups should be less reliant on the distributed know-how within the regional knowledge base, due to the development of firm-specific routines and processes as well as increased levels of formalization within the firm. For instance as venture-backed startups mature they implement formalized control systems and rely less on ad-hoc advice (Davila and Foster, 2007; Wijbenga, Postma, and Stratling, 2007). Furthermore, these firms also become more "professionalized", with VCs typically replacing founders with professional managers over time (Hellmann and Puri, 2002). For these more sophisticated and experienced managers, the role of community mentorship and guidance is less important, as they rely more on their own prior experiences when making decisions (Dew *et al.*, 2009; Isenberg, 1986). Integrating these sets of arguments, I therefore hypothesize that, in the context of venture-backed startup firms:

Hypothesis 2: The impact of regions on venture-backed startup performance will be moderated by the startups' developmental stage such that it will be larger for startups' at early stages of development than for startups' at later stages of development.

Much like firms, industries also evolve and progress through different stages of maturity (Benner and Tripsas, 2012; Rindova, Petkova, and Kotha, 2007; Santos and Eisenhardt, 2009). I next discuss why this process is also likely to change the dependence of VC-backed startups on regional resources, once again relating my arguments back to their reliance on technological knowledge and managerial know-how.

First, with respect to technological knowledge, a number of studies have indicated that as industries mature, there is often a shift in the technological regime (Nelson and Winter, 1982; Winter, 1984). Particularly, with respect to the underlying technological knowledge base, there is a shift from more abstract tacit knowledge in the initial stages of industry creation towards more broad, based generalized knowledge at later stages (Arora and Gambardella, 1994). However, as the knowledge base builds over time it also becomes more codified as more details about inventive processes are revealed (Cowan *et al.*, 2000; Johnson, Lorenz, and Lundvall, 2002; Nesta and Saviotti, 2006; Ter Wal, 2013). This shift in the knowledge base also has implications with respect to the importance of regions for VC-backed startups. In particular, since tacit knowledge needs face-to-face contact for transmission, it is regionally sticky (Polanyi, 1966). In comparison more codified and general knowledge, such as that found in basic science (Fleming and Sorenson, 2004), is far more easy to transmit and interpret across geographic boundaries outside the context that it was created in (Gittelman, 2007). This logic therefore suggest that VC-backed startups that require access to technological knowledge should benefit most from regional knowledge exchange dynamics when industries are relatively nascent (Audretsch and Feldman, 1996; Neffke *et al.*, 2011).

Similarly, the importance of specialized region-specific managerial know-how is also likely to decline as industries mature. With a longer history of startup formation and failure, both the ambiguity and uncertainty about the process of “firm building” (Bresnahan *et al.*, 2001; Rindova, Ferrier, and Wiltbank, 2010; Santos and Eisenhardt, 2009) is reduced, since startups are able to both draw upon and imitate previously established business models. Suddaby and

Greenwood (2001) refer to this evolutionary process as the "commodification of management knowledge" (also see Abrahamson, 1996), as practices become legitimized, taken for granted, and widely diffused. Since this process of standardization also makes managerial knowledge more codified and portable, it also diffuses more easily across both organizational (O'Mahoney, Heusinkveld, and Wright, 2013) and geographic boundaries (Guillén, 1998; Tempel and Walgenbach, 2007), and is therefore less likely to be confined to any particular location. Integrating these two sets of arguments, I therefore hypothesize that:

Hypothesis 3: The impact of regions on venture-backed startup performance will be moderated by the maturity of the industry segment that the startup is operating in, such that it will be larger for startups operating in more nascent industries.

METHODS

Study sample

To obtain a nation-wide, comprehensive sample of venture-backed startup firms, I used data from the VentureXpert database provided by Thomson Financial. This database has been used extensively for research on both venture capital firms and venture-backed startups operating in knowledge intensive sectors of the economy (see Dushnitsky and Lenox, 2005 for a review). To construct the data sample, I followed prior research that has used VentureXpert data and used a variance decomposition analytical approach (see Fitza *et al.*, 2009 for a detailed description of the sample and analytical approach). After excluding missing data, the final sample consisted of 7,813 observations (i.e. inter round periods (see measures section below)) from 3,893 venture-backed startups over the period 1980-2012[6]. These single-location ventures operated in 133 Metropolitan Statistical Areas (MSAs) across 71 industry segments (defined using Venture Economic Industry Classification (VEIC) codes). My choice of the MSA as a regional unit of analysis was primarily driven by the theoretical emphasis on spatially localized knowledge as a driver of regional effects (Breschi and Lissoni, 2001b) and empirical findings from the economic geography literature, which suggest that most regional knowledge exchange typically occurs within small physical distances below 50 miles (Powell *et al.*, 2002; Rosenthal and Strange, 2003). Furthermore, from a practical standpoint, venture-backed startups typically locate in metropolitan centers and not in isolated rural areas (Chen *et al.*, 2010). This choice of spatial scale is also consistent with prior empirical research that has looked at the impacts of specific regional factors (e.g. networks, knowledge stocks) on startup performance in the context of venture-backed startups (Decarolis and Deeds, 1999; McCann and Folta, 2011).

Analytical approach

I utilized a simultaneous analysis of variance (ANOVA) estimation technique to estimate the importance of factors that lead to competitive heterogeneity among these startup firms. Using this method, I was able to attribute portions of the variance in the dependent variable to individual effect classes. Closely following the work of Fitza *et al.* (2009), I first assessed the importance of firm (i.e. the new venture), development stage, owner (i.e. VC investor), industry segment, and year effects. In addition to these previously established effects, I then also examined the impact of region of primary interest to this study. Furthermore, to estimate region interaction effects (i.e. whether the region effect is in turn contingent on other variables), I

[6] I chose 1980 as a start-year for the sample both due to data limitations in the years prior to that, as well as the 1979 amendment in the Employee Retirement Income Security Act (ERISA) that led to a dramatic increase in commitments by institutional investors to venture capital (Gompers and Lerner, 2001).

estimated interactions between the region and development stage, and region and industry respectively. Specifically, I estimated the following model:

$$Performance_{i,r} = \mu + \lambda_{rg,r} + \alpha_{y,r} + \gamma_{s,r} + \xi_{rg,s,r} + \nu_{c,r} + \varkappa_{rg,c,r} + \text{ŧ}_{o,r} + \tau_{f,r} + \varepsilon_{f,r}$$

where μ is a constant equal to the grand mean. The term $\lambda_{rg,r}$ captures the *region* effect, $\alpha_{y,r}$ captures the *year* effect (a measure of macro-economic conditions), $\gamma_{s,r}$ the *development stage* effect, $\xi_{rg,s}$ the *region –development stage* interaction effect, $\nu_{c,r}$ the *industry segment effect*, $\varkappa_{rg,c,r}$ the *region – industry segment* interaction effect $\text{ŧ}_{o,r}$ the *owner* effect, $\tau_{f,r}$ the *firm* effect, and $\varepsilon_{f,r}$ the residual. Note that the region interaction effects estimate whether the region effect is contingent on other variables (the development stage, and the industry respectively). Each effect in the model is represented by a set of dummy variables. Following Fitza *et al.* (2009) I compared estimates of the model above with estimates of a model that omitted a specific effect (Judd and McClelland, 1989). For example the start-up effect is determined by comparing the explained variance (R^2) of an equation that contains all effects with the explained variance of an equation that omits the dummies representing the start-up. This approach allowed me to measure all effects in the context of a model that allows for the maximum possible covariance between the effects of interest. As a result the order in which the effects are introduced into the model is not important (as long as the startup effect is removed first as it is nested within other effects). In addition to applying this model across the entire dataset, I also constructed data subsamples, based on firm and industry segment maturity level, to test the directional interaction effects proposed in hypothesis 2 and 3.

Measures

Startup performance: Following prior literature (e.g., Fitza *et al.*, 2009; Gompers and Lerner, 2000), I investigated performance differentials between venture-backed startups by measuring differences in growth rates. Growth is a particularly important performance metric to investigate in the startup context (Zimmerman and Zeitz, 2002) as it enables firms to overcome the liabilities of adolescence and smallness that can lead to premature failure. Following prior research in the context of venture-backed startups, I used changes in valuation as a measure of startup growth. I computed this metric by using the percent increase (per month) in the valuation of a portfolio company *i* in the period between two investment rounds (the difference between the post money value of round n and the pre money valuation of round n+1, divided by the number of months between the two rounds) known as the inter-round period *r* (Fitza *et al.*, 2009; Gompers and Lerner, 2000; Hsu, 2007).

Startup development stage: For each investment round, the VentureXpert database identifies the development stage (Ruhnka and Young, 1987) of financed startup firms. Following prior research, I was therefore able to classify the observations in the sample into seed, early, expansion, late, and pre-IPO/acquisition stages (Dimov and Murray, 2008; Fitza *et al.*, 2009). I also used these groups for a sub-sample analysis to examine the impact of region for startup growth at different stages of development.

Industry segment maturity: Following prior research, I identified the maturity of the industry segment within which the startup operated using a measure of cumulative business activity (e.g., Baum and Haveman, 1997). More specifically, I used the historical cumulative number of venture investments rounds within venture investment market segments from the onset of the VentureXpert database in 1964. I identified market segments using VEIC codes at the sub-sector level as defined by the VentureXpert database (Dushnitsky and Shaver, 2009; Wadhwa and Kotha, 2006). I chose to uses these codes to delineate market segments as they reflect the targeted lines of business that VC-backed startups operate within (Dushnitsky and

Shaver, 2009), and are hence well suited to identify both nascent and mature markets that VCs might choose to invest in (Dimov, de Holan, and Milanov, 2012). I also checked the robustness of these results to using alternate specification such as Standard Industry Classification (SIC) and North American Industrial Classification (NAICS) systems, using a concordance table to map between categorizations (Wadhwa and Basu, 2013).

RESULTS

An initial examination of the data indicated that, not surprisingly, regional differences exist in startup performance. In figure 3 below, I demonstrate this visually by plotting the increase in valuation aggregated by MSA over the time period of the study. For ease of illustration, I show data for the top 25 MSAs in the sample. I also plotted the median increase in valuation to minimize the effects of outliers in this descriptive graph. Interestingly, I note that while the regions in the Silicon Valley area (i.e. San Francisco, San Jose, and Oakland) which arguably possess the strong entrepreneurial ecosystems in the nation are among the top performers, they are not among the top three locations. These are instead occupied by Los Angeles, Lowell, and Boulder; regions that have traditionally been viewed as satellite venture-capital markets (Chen *et al.*, 2010). This provides some tentative but suggestive evidence that regional factors, while contributing to VC-backed startup growth, might not be the most important determinants.

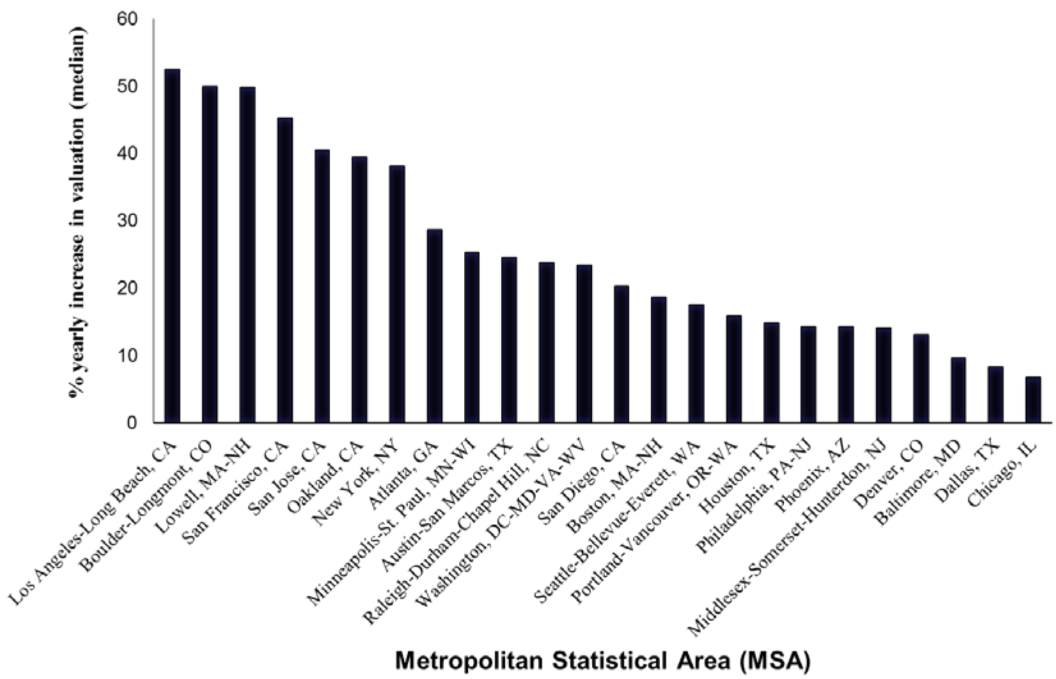

Figure 3 (paper 2). Regional differences in the valuation increases of venture-backed startup firms

Table 5 below summarizes the main findings from the variance decomposition analysis. Effect sizes in this study are comparable to the findings reported by Fitza *et al.* (2009) taking into account differences in the time span and industry sectors examined.

Table 5 (paper 2). Results of variance decomposition analyses[a]

	Studies using VentureXpert data		Previous various decomposition studies on regional effects		
	This study	*Fitza et al., 2009*	*Chan et al. (2010)*		Ma *et al.* (2013)
Effect Class	% of variance in inter-round valuation increase	% of variance in inter-round valuation increase	% of variance in ROS explained (U.S. affiliates)	% of variance in ROS explained (Chinese affiliates)	% of variance in ROA explained
Calendar Year	1.72	3.68	0.3	2.6	0.04 n.s
VC Investor[b]	19.88	11.15	18.8	19.5	5.28
Startup[c]	30.34	26.32	16.8	14.4	8.87
Industry Segment	2.34	-0.55 n.s.	12.8	5.9	5.65
Startup Development Stage	0.49 n.s.[e]	0.22 n.s.	-	-	-
Region[d]	2.13	-	2.1	1.4	3.09
Region X Industry Segment	8.85	-	-	-	10.08
Region X Startup Development Stage	2.34	-	-	-	-
Number of firms	3,893	3,756	4, 931 (in total)		1,625
Number of Observations	7,813	6,490	16,277	13,051	8,043
Time period covered by data	1980-2012	1980-2005	1996-2005		1998-2005

[a]All effect sizes are significant at p<0.05 or higher, except for those marked n.s. Effects marked with a - were not examined in the study.

[b] VC investor effects in our study and Fitza *et al.* (2009) are analagous to ownership effects in the traditional variance decomposition literature.

[c] Startup effects in our study and Fitza *et al.* (2009) are analagous to firm/buisness unit effects in the traditonal variance decomposition literature. In Chan *et al.* (2010) and Ma *et al.* (2013) they refer to foreign subsidiaries and affiliates respectively.

[d] We identify region effects across 147 U.S. MSAs. Chan *et al.* (2010) identify effects across 34 U.S. states, and 21 regions in China. Ma *et al.* (2013)

[e]A small negative effect is usually interpreted to be zero (McGahan & Porter, 2002).

Resources internal to the firm explain most of the variance between startups with VC investor and startup effects accounting for 19.88 percent and 30.34 percent of the variance in the valuation of venture-backed startups respectively. With respect to the region effect, I find that it is statistically significant and explains 2.13 percent of the variance in venture-backed startup valuation increase. It is important to note that while the magnitude of the region effect is relatively small, it is an important contextual factor. It explains a slightly greater portion of the variance in valuation than macro-economic conditions (i.e. year effects at 1.72%), and is comparable to industry sector effects (2.34%). In addition I find that the region-development stage interaction explains 2.34 percent of the variance, and the region-industry sector interaction explains 8.85 percent of the variance in the valuation of venture-backed startups.

As the variance decomposition approach cannot test the directionality of interaction effects, I also carried out a set of subsample analyses partitioning the data by startup development stage and industry sector maturity. As illustrated in figure 4 (solid bars), regions matter most for startups at the seed-stage of development, explaining 14.01 percent of the variance in venture valuation. However, as startups mature, regional influences decline explaining 5.31 percent for early-stage firms, 2.67 percent for expansion-stage firms, 2.33 percent for late-stage firms, and 0.14 percent for firms at the pre-acquisition/IPO stage. I find similar patterns when considering the impacts of industry sector maturity (figure one, shaded bars). Across the time period of the study, the average number of cumulative investments within a VEIC industry segment is 1952, with a standard deviation of 2682. I observe that the magnitude of the region effect declines as industries mature, explaining approximately 5 percent of the variance in startup valuation in nascent industries (2000 or fewer cumulative investments) and only 2–3 percent in more mature industry sectors.

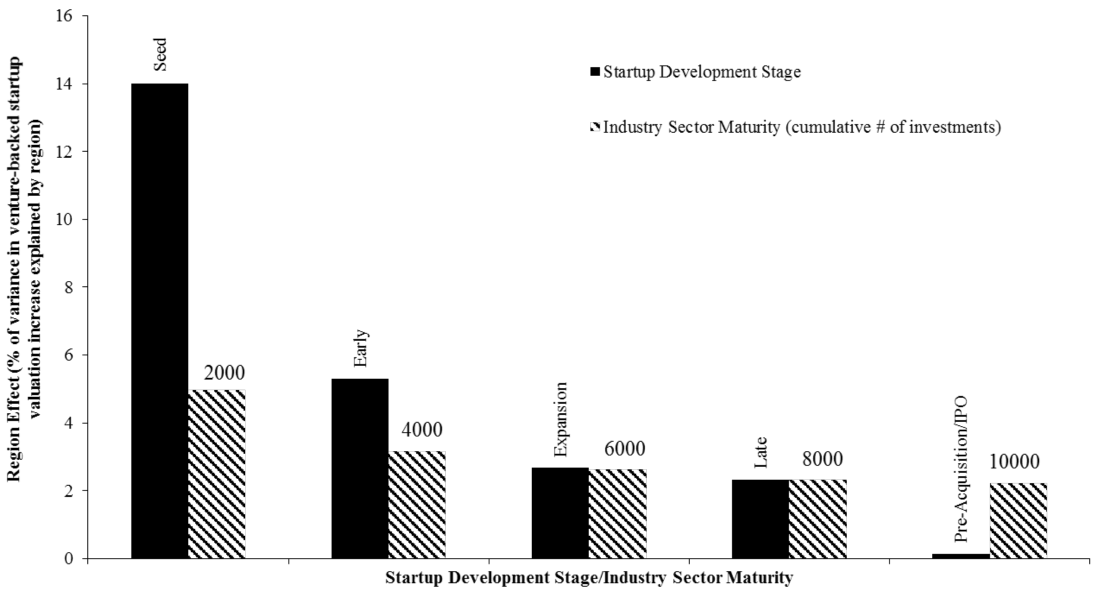

Figure 4 (paper 2). The region effect as a function of startup development stage (solid bars) and the maturity of the industry sector that the startup operates in (shaded bars)

Sensitivity Analyses

I also conducted a series of sensitivity tests. First, I defined industry segments using SIC and NAICS codes instead of VEIC codes, to ensure that the results were not driven by a specific definition of industry segments. In testing the third hypothesis assessing the conditional impacts of industry segment maturity, I also used a cumulative count of IPOs in each industry segment instead of the cumulative number of investments as an alternate indicator of industry maturity (Dimov *et al.*, 2012). Second, while I did account for yearly macro-economic effects in the main analysis so as to capture the periodic booms and busts and market volatility in the VC industry as a whole (Gompers and Lerner, 2003), I followed prior work (Matusik and Fitza, 2012) and repeated analyses after deleting data from IPO market bubble years (1998-2000). Given that a large number of VC-backed startups were both formed and grew at unusual rates during this time period, this allowed me to verify that the inclusion of this time period in the study did not systematically bias my results. Third, I also used alternate, larger regional spatial specifications, such as economic areas (Alcácer and Chung, 2007) and U.S. states instead of MSAs. While extant evidence and theory suggests that regional effects should be most salient in smaller localized spatial scales (Rosenthal and Strange, 2003), I nevertheless repeated the analyses to assess the sensitivity of the methodology to the choice of spatial scale. Results from these specifications were qualitatively similar to the main results and are illustrated in Table 6 below.

Table 6 (paper 2). Sensitivity analyses[a]

Effect Class	**Model 1** (Main Model from Table 5)	**Model 2** (VEIC[b] industry sub-sectors)	**Model 3** (dotcom bubble removed)	**Model 4** (regions defined at the state level)
Calendar Year	1.72	2.15	0.59 n.s.	2.13
VC Investor	19.88	18.67	30.75	19.5
Startup	30.34	38.64	24.34	36.38
Industry Segment	2.34	0.15	6.76	2.35
Startup Development Stage	0.49 n.s.	0.41 n.s.	0.67 n.s.	0.61 n.s.
Region	2.13	1.73	6.54	0.57 n.s.
Region X Industry Segment	8.85	7.81	7.26	5.82
Region X Startup Development Stage	2.34	2.45	2.75	2.56

[a]All effect sizes are significant at $p<0.05$ or higher, except for those marked n.s.
[b]The VEIC codes are industry classifications specific to VentureXpert. They reflect the primary lines of business that startups within the database operate in. As such, they are well suited to identify both nascent and more mature industry segments that VCs invest in (e.g., Dimov, de Holan, and Milanov, 2012).

DISCUSSION

In recent years, the variance decomposition literature has moved beyond its origins in corporate strategy, and begun to examine the importance of factors beyond firm and industry effects. In particular, a growing body of work has begun to study populations of startup firms (Fitza *et al.*, 2009; Short *et al.*, 2009) and quantify regional effects (Chan *et al.*, 2010; Ma *et al.*, 2013). This study is the first to examine regional effects in the context of venture-backed startups, high-growth firms which when successful can have significant economic impacts (Shane, 2009).

The two prior variance decomposition studies examining region effects (Chan *et al.*, 2010; Ma *et al.*, 2013) were carried out in samples of subsidiaries of mature firms, across different institutional contexts (e.g. China), and spatial scales (e.g. U.S. states). Given that extant theory suggests that regions should be important in the startup context and within small spatial scales, I expected that isolated region effects would be larger in this study in comparison to prior work. This was not the case however, as regions explain 2–3 percent of the variance in performance. However, these findings should not be interpreted to mean that regions do not matter. The subsample results clearly highlight that the importance of region is contingent on firm and industry evolutionary stages. Furthermore, even the relatively small main region effect is comparable in magnitude to industry sector effects and larger than macro-economic effects. Hence while these other external factors have received attention in the prior literature; for instance entrepreneurs are often encouraged to exploit opportunities in industry sectors with growth potential, the findings from this study suggest that it might be just as important to choose the firm's location strategically[7].

Furthermore, the interaction models (subsample analysis) also provide guidance on the firm and industry specific conditions under which regions are particularly important. In particular, the finding that regions matter significantly for nascent, seed-stage ventures while mattering much less for later-stage startups should be particularly interesting to policy makers. A direct implication is that ecosystem level efforts to stimulate regional growth through technology oriented, venture-backed entrepreneurship (Mason and Brown, 2013a) should be targeted towards early-stage ventures. Similarly, the research on startup location choice indicates that founders start firms where they live within existing regional clusters (Dahl and Sorenson, 2012).

[7] The geographic location of firms in this study is endogenous; entrepreneurs do not randomly decide where to start the firm. However such selection effects are common in variance decomposition research. Firms choose the regions in which they locate and corporations choose the industry in which they operate. However given that the goal of the variance decomposition method is to estimate the performance implications of these strategic choices, this is not problematic.

These results suggest that the utility of this approach towards location choice declines with the maturity of the industry sector within which the startup operates.

Like all studies, this study too is not without its limitations. For instance, the sample is restricted to new ventures that have received at least two rounds of venture capital financing and that have valuation data available. While this is potentially problematic, I ran a series of sensitivity tests similar to Fitza *et al.* (2009) to ensure that the sample was representative of the population of firms in the VentureXpert dataset. Furthermore I recognize that the use of venture-backed startups potentially limits the applicability of the findings to new ventures that rely on alternate forms of financing. However, since my interest is primarily in studying the impacts of regions in the context of high-growth startups, venture-backed firms are a good population to examine. In addition, as with all variance decomposition research, I am unable to examine the precise causal mechanisms that impact effect sizes, but only study the importance of regions as a whole. Hence, this study is but an initial step in assessing the importance of regions. Taken together, these results indicate that there is value in examining which exact factors and which resources in particular are the sources of the regional effect that I measure.

Concluding remarks

This paper uses the context of venture-backed startups to quantify the degree to which regions impact increases in new venture valuation. While extant theory suggests that regions should act as a source of competitive advantage for these firms, their impacts have been understudied relative to other internal (e.g. founding teams, VCs) and external (industry) factors. I find conclusive evidence that regions are a particularly important resource for startups at the earliest stages of development, and for those operating in nascent industry sectors.

PAPER 3: A BEHAVIORAL THEORY OF STARTUP FAILURE WITHIN GEOGRAPHIC CLUSTERS: INVESTIGATING ADVERSE SELECTION AND PERFORMANCE PREMIUM EFFECTS

ABSTRACT

Startup firms typically locate within dense spatial concentrations of industrial activity, known as geographic clusters. While geographic clustering is typically associated with favorable outcomes such as innovation and growth, it can also increase the likelihood of startup failure, particularly as clusters grow larger. The literature has primarily attributed such failure to cluster-based agglomeration diseconomies, such as higher competition for resources and increased levels of congestion. In this study I investigate whether higher failure rates within geographic clusters can instead be explained by *pre and post-entry* behavioral dynamics, namely adverse selection and performance premium effects. I do not find evidence for adverse selection into clusters, but do find that startups within geographic clusters require a performance premium, that is a comparatively higher level of performance relative to isolated peers, to persist with operations. I offer implications for theory, practice, and public policy.

INTRODUCTION

Startup firms often locate within dense spatial concentrations of existing industrial activity. This phenomenon is typically referred to as geographic clustering (Agarwal and Braguinsky, 2014; Chatterji *et al.*, 2014; Delgado *et al.*, 2010; Feldman, Francis, and Bercovitz, 2005; Glaeser, Kerr, and Ponzetto, 2010; Klepper, 2010). Geographic clusters are typically argued to be key drivers of competiveness at the firm (Bell, 2005; Jenkins and Tallman, 2010; Porter, 1998; Tallman *et al.*, 2004), industry (Delgado, Porter, and Stern, 2014a; Spencer *et al.*, 2010), regional (Boschma and Iammarino, 2009; Porter, 2003), and national (Delgado *et al.*, 2012; Porter, 1990) levels. Furthermore, in terms of their subsequent impacts on startup performance, much of the research has positively linked geographic clustering to outcomes such as startup innovation (Folta *et al.*, 2006a; Gilbert *et al.*, 2008; McCann and Folta, 2011), growth (Gilbert *et al.*, 2008; Wennberg and Lindqvist, 2010), early internationalization (AlLaham and Souitaris, 2008; Fernhaber, Gilbert, and McDougall, 2008), and successful initial public offerings (Decarolis and Deeds, 1999).

And yet the literature also suggests that there is a potential downside to geographic clustering. In particular, higher levels of geographic clustering also lead to more startup failure (Audia and Rider, 2010; Baum and Mezias, 1992; Folta *et al.*, 2006a; Globerman, Shapiro, and Vining, 2005; Greve, 2002; Sorenson and Audia, 2000; Staber, 2001). The traditional explanation for this finding in the literature is that as geographic clusters grow larger, positive agglomeration (i.e. spatial concentration) externalities such as knowledge spillovers, pooled labor, and supplier linkages (Alcácer and Chung, 2007, 2014; Marshall, 1890) increasingly get replaced by negative agglomeration diseconomies. These cluster-based agglomeration diseconomies primarily occur due to higher levels of localized competition among firms for resources (Baum and Mezias, 1992; Folta *et al.*, 2006a; Lomi, 1995; Sorenson and Audia, 2000), but can also be a result of factors such as increased congestion, higher operating costs, and more turnover within larger geographic clusters (Arnott, 2007; Combes and Duranton, 2006; Pouder and John, 1996).

In this study I argue that this traditional explanation, while valid, might not be comprehensive. In particular I consider the relative explanatory power of two alternate behavioral explanations, *adverse selection* and *performance premiums.* The adverse selection explanation focuses on *pre-entry* behavioral dynamics. It highlights the fact that firms both contribute to and benefit from the location-based externalities that occur in geographic clusters. Thus under conditions of firm heterogeneity and asymmetric benefits from geographic clustering, larger geographic clusters might deter high quality firms and attract a disproportionate number of low quality entrants (Cantwell and Santangelo, 2003; Flyer and Shaver, 2003; Oakey and Cooper, 1989; Shaver and Flyer, 2000). Such negative selection effects are also accentuated by the fact that entry barriers are also typically lower within geographic clusters (Porter, 2000), thus facilitating the entry of more low quality startups who have little to contribute to in the way of positive agglomeration externalities (Flyer and Shaver, 2003; Kalnins and Chung, 2004; Shaver and Flyer, 2000). The performance premium explanation focuses on *post-entry* behavioral dynamics. It suggests that higher levels of geographic clustering could impact entrepreneurial decision making processes, and in particular trigger more *voluntary exits* by entrepreneurs (McCann and Folta, 2008). This is because larger geographic clusters typically have lower occupational switching costs (Folta *et al.*, 2006b) and higher levels of peer-pressure (Porter, 2000). This in turn leads to a general increase in expected levels of performance (Cassar, 2006; Landier and Thesmar, 2009) as well as an upward shift in specific goal levels such as survival targets (March and Shapira, 1992), more typically referred to as exit thresholds[8] (Gimeno *et al.*, 1997). Thus, the higher rates of failure observed in clusters might simply be a reflection of entrepreneurs requiring a *premium level of performance* (Czarnitzki *et al.*, 2014) to persist with operations.

Given that these behavioral dynamics have *prima facie* validity, the purpose of this study is to empirically assess their utility in characterizing startup failure within geographic clusters. This is important to do as these alternate explanations have been theoretically discussed in the

[8] The exit threshold refers to the target level of performance below which an entrepreneur will voluntarily choose to discontinue operations (DeTienne, 2010; DeTienne, Shepherd, and De Castro, 2008; Gimeno et al., 1997).

literature (Folta *et al.*, 2006a; McCann and Folta, 2008; Shaver and Flyer, 2000), but not empirically examined. By combing data from the Kauffman Firm Survey (KFS) and the Harvard Cluster Mapping Project (HCMP), I examined startup failure dynamics in a nation-wide representative cohort of startup firms over the period 2004-2011. I do not find any evidence for adverse selection into geographic clusters by startups based on data from the initial year of the KFS (year 2004). Instead, I find that high quality startups might actually be positively selecting (i.e. self-selection instead of adverse selection) into the largest of geographic clusters. When I examine *post-entry* behavioral dynamics I also find systematic evidence of geographic clustering leading to performance premium effects. For instance, results from survival analyses indicate that startups within geographic clusters are more likely to terminate operations than their peers in more isolated locations, for comparable objective levels of performance (objective performance levels were measured as risk-adjusted profitability). In addition, when I estimate latent (i.e. unobservable) startup exit thresholds using censored regression estimation models (Folta and O'Brien, 2008; Gimeno *et al.*, 1997; McCann and Folta, 2012; Nelson, 1977), I find that geographic clustering raises these goal levels. And lastly, I also find that geographic clustering leads to lower self-evaluated ratings of performance among startups that were in operation during the fourth follow-up survey of the KFS (year 2008).

These findings have important implications. From a theoretical perspective, they highlight the important role that *post-entry* behavioral dynamics, specifically performance premium effects, play in contributing to startup failure within geographic clusters. This challenges the dominant theoretical explanation in the literature which attributes startup failure within clusters exclusively to agglomeration diseconomies (McCann and Folta, 2008). I therefore suggest that startup failure within geographic clusters can be more accurately characterized as a function of both voluntary and involuntary firm exit (DeTienne, 2010), due to a heightening of both real competitive pressures (Baum and Mezias, 1992; Sorenson and Audia, 2000) and higher performance expectations respectively. These results also have important policy implications. For instance, cluster-based policy instruments have been increasingly used to stimulate entrepreneurial activity, innovation, and economic growth within regions (Asheim, Boschma, and Cooke, 2011; Cooke, 2008; Gilbert, Audretsch, and McDougall, 2004; Rocha, 2004; Rocha and Sternberg, 2005). Evidence of adverse selection dynamics into geographic clusters would have therefore pointed to a wastefulness of resources, and the unintended creation of perverse incentives from a public policy standpoint (Shane, 2009). Encouragingly, my findings indicate that that is not the case. However they do suggest that geographic clustering can contribute to higher, but possibly unrealistic and exaggerated expectations among entrepreneurs (Cassar, 2014; Hayward *et al.*, 2006; Ordóñez *et al.*, 2009), which in turn can increase rates of voluntary exit and startup failure. Thus while proponents of cluster-based strategies have largely emphasized the positive behavioral impacts that clustering can have on managers, such as enhanced motivation (Porter, 2000), I suggest however that policy makers should also be cognizant of the possible downside in terms of the behavioral effects of such competitive rivalry.

The rest of this paper is organized as follows. First, I briefly summarize the existing empirical evidence on the strategic implications of geographic clustering for startups. In doing so I highlight the fact that geographic clustering can act as a double-edged sword, positively impacting some aspects of startup performance while also increasing the likelihood of failure. Next, I briefly summarize the traditional theoretical explanation in the literature, where startup failure is attributed to agglomeration diseconomies within clusters. In doing so, I emphasize that this explanation largely ignores the voluntary (i.e. behavioral) aspects of firm exit (DeTienne, 2010; Gimeno *et al.*, 1997). I then introduce the alternate *pre* and *post-entry* behavioral explanations linking geographic clustering to startup failure, namely adverse selection and performance premium effects. Given that these alternate explanations are mutually non-exclusive, I develop relevant testable hypotheses. Next, I introduce the study context and provide empirical results. I conclude with the implications of my findings, as well as some potential avenues for future research.

THEORY AND HYPOTHESES

The intuition that dense industrial agglomerations can lead to location-based externalities for firms is not novel. In fact most contemporary scholars credit Alfred Marshalls' treatise on the *Principles of Economics* at the turn of the 19^{th} century for the seminal ideas of agglomeration theory, often citing the "holy trinity of agglomeration economies" (knowledge spillovers, pools of skilled labor, and local supplier linkages) that he uncovered from his field research in the Sheffield metal industry (Marshall, 1890). Over time, scholars have periodically revisited and refined the applicability of Marshall's key ideas most notably with Paul Krugman's work on "new economic geographies" of knowledge-based industries (Krugman, 1991), as well as Michael Porter's cluster-based policies for economic development (Porter, 1998).

Recently there has been increasing scholarly interest in understanding the interrelationship between geographic clusters and entrepreneurial dynamics (Agarwal and Braguinsky, 2014; Chatterji *et al.*, 2014; Delgado *et al.*, 2010; Glaeser *et al.*, 2010). In particular, a number of studies have focused on the strategic implications of cluster-location choices for startup firms. Consistent with the benefits espoused by agglomeration theory, some studies have shown that startups can indeed benefit from the positive externalities available in geographic clusters. For instance, geographic clustering has been shown to positively impact startup innovation (Folta *et al.*, 2006a; Gilbert *et al.*, 2008; McCann and Folta, 2011), new venture growth (Wennberg and Lindqvist, 2010), early internationalization rates (AlLaham and Souitaris, 2008; Coombs, Mudambi, and Deeds, 2006; Fernhaber *et al.*, 2008), and initial public offering valuations (Decarolis and Deeds, 1999). And yet, the evidence is not unequivocally positive. In particular a number of studies have shown that geographic clustering can also increase startup failure rates, particularly once clusters become very large (Audia and Rider, 2010; Folta *et al.*, 2006a; Sorenson and Audia, 2000; Staber, 2001).

The conventional explanation: Agglomeration diseconomies and startup failure

The traditional explanation for high observed rates of startup failure within geographic clusters is that in addition to positive agglomeration externalities (e.g. knowledge spillovers, supplier linkages, labor pooling), firms also suffer from location-based diseconomies in geographic clusters. For instance, as the density of firms in a location increases, firms are exposed to higher levels of direct localized competition (e.g., Baum and Mezias, 1992; Greve, 2002; Lomi, 1995; Plummer and Acs, 2014). These dynamics are particularly likely to occur in declining and/or mature industries where the scope of positive externalities have diminished and resources are scarce (Neffke *et al.*, 2011; Potter and Watts, 2011; Sorenson and Audia, 2000; Staber, 2001). Beyond such direct competitive effects, high levels of geographic clustering can also lead to indirect competition for resources in a range of economic sectors. For example, the ability to attract highly skilled star scientists and inventors is a key source of competitive differentiation in technology intensive industry sectors such as biotechnology (Zucker, Darby, and Brewer, 1998). However geographic clusters are comprised of multiple, closely related industry sectors (Delgado *et al.*, 2014a), and employee skills are typically transferable across careers that require similar skillsets (Barnett and Miner, 1992). Thus as clusters grow larger, labor market boundaries can get increasingly blurred (Sullivan and Arthur, 2006), such that startups find it increasingly difficult to both recruit and retain personnel due to dynamics such as employee poaching (Combes and Duranton, 2006), and job-hopping (Fallick *et al.*, 2006; Marx, 2011). In addition to these deleterious effects of both direct and indirect competition within factor markets, startups in large geographic clusters might also incur other agglomeration diseconomies, such as higher operating costs. For instance, employee wages typically increase in the presence of an educated and pooled workforce (Hanson, 2001), and both commercial and residential real estate costs are also higher within larger geographic clusters (Prevezer, 1997). In aggregate, these sets of arguments suggest that as geographic clusters grow larger, agglomeration diseconomies can outweigh positive agglomeration effects (Pouder and John, 1996: 1206). In particular, with respect to startup failure, this also means that any marginal benefits derived from geographic clustering also decline as a geographic clustering levels increase (Folta *et al.*, 2006a).

While the arguments listed above have been well accepted and studied in the literature, they are somewhat narrow in their focus in that they exclusively privilege location-based factors to the exclusion of other drivers of startup failure. In particular, they suggest that entrepreneurs have limited agency over the exit process and that startups that fail within geographical clusters largely succumb to the heightened competitive and resource scarcity pressures that they face in such environments (Baum and Mezias, 1992). And yet, research on organizational behavior and decision making indicates that firm-specific *perceptions* of environmental factors might be just as important in explaining their behaviors (Short and Palmer, 2003; Simon, Houghton, and Aquino, 2000). In particular, with regards to startup failure, a growing body of research has emphasized that exit behaviors need to be conceptualized, at least in part, as an explicit managerial choice driven by entrepreneurial expectations (DeTienne, 2010; DeTienne *et al.*, 2008; Gimeno *et al.*, 1997; Wennberg *et al.*, 2010). In this study, I therefore draw on this research and link it to the impacts of geographic clustering on startup failure. In particular, I discuss how both the *pre- and post-entry expectations* of entrepreneurs can potentially lead to both adverse selection and performance premium effects within geographic clusters, and thus lead to higher observed rates of startup failure.

Behavioral explanation # 1: Adverse selection and startup failure in clusters

Adverse selection arguments indicate that as geographic clusters grow in size, they might increasingly attract weak firms while deterring stronger entrants (Flyer and Shaver, 2003; Shaver and Flyer, 2000). These arguments center on the fact that firms are not passive actors within clusters. That is, they both contribute to and benefit from agglomeration externalities (Kotha, 2010; Operti and Carnabuci, 2014; Yang *et al.*, 2010). Furthermore to the extent that firms are heterogeneous in their capabilities, proponents of adverse selection indicate that high quality firms have relatively less to gain and more to lose (i.e. lower marginal benefits) from co-location relative to low quality firms. Given the potential for such free-riding behavior in geographic clusters, some scholars have therefore suggested that high quality firms might seek to strategically avoid locating within clusters, particularly when they grow very large (Alcácer and Chung, 2007; Chung and Alcácer, 2002; Kalnins and Chung, 2004; Shaver and Flyer, 2000).

Adverse selection arguments therefore effectively reverse the direction of causality between geographic location and firm performance (Folta *et al.*, 2006a). While these arguments are not specific to startups per se they are worth considering given that entry barriers are usually lower in larger clusters, and low quality startups can and do enter on a regular basis (Delgado *et al.*, 2010; Glaeser *et al.*, 2010; McCann and Folta, 2008; Porter, 2000; Shane, 2009). Thus to the extent that such selection processes are prevalent within geographic clusters, the higher levels of observed failure might be a reflection of the diminishing quality of the entrant pool. Furthermore, while scholars exploring entrepreneurial location choices have traditionally indicated that entrepreneurs usually start firms in the regions where they live and do not choose strategically between locations (Dahl and Sorenson, 2012; Klepper, 2010), recent empirical studies do find evidence of such selective sorting dynamics. For example, studies have found that the possession of high-levels of resources and capabilities (Pe'er *et al.*, 2008), as well as the presence of low quality peers (Kalnins and Chung, 2004) can deter startups from locating within geographic clusters. Similarly, while startups that need to access technological spillovers preferentially locate within geographic clusters (Aharonson *et al.*, 2007), others that are less focused on accessing such benefits might actually move away from established industrial agglomerations (Berchicci, King, and Tucci, 2011). In sum, these arguments suggest that the positive relationship between geographic clustering and startup failure can be attributed to adverse selection effects into geographic clusters. I test this relationship by examining initial locational choice dynamics and accordingly hypothesize that:

Hypothesis 1: Startup firm quality is negatively related to the likelihood of locating in a cluster.

Behavioral explanation # 2: Cluster-based performance premiums and startup failure

In addition to such selective sorting, geographic clustering can also potentially influence the *voluntary* exit decisions of startup firms. In particular, some scholars have argued that the observed patterns of higher failure within larger geographic clusters could simply be a reflection of higher expected levels of performance and exit thresholds (Cooper and Folta, 2000; McCann and Folta, 2008), where the exit threshold refers to the minimum level of performance below which entrepreneurs will voluntarily choose to close startups (DeTienne, 2010; DeTienne *et al.*, 2008; Gimeno *et al.*, 1997). I suggest that this could occur for two main reasons; lower occupational switching costs (McCann and Folta, 2008) and higher peer-pressure (Nanda and Sørensen, 2010; Porter, 2000) within geographic clusters.

Occupational switching costs are likely to be lower for startups operating within geographic clusters relative to their peers in more isolated locations for a number of reasons. First, due to the spatial agglomeration of firms in related industries within clusters (Delgado, Porter, and Stern, 2014b), entrepreneurs are likely to have more opportunities to use their specific human capital outside the firm (Rosa, 1998). Second, clusters also result in the formation of thick labor markets (Puga, 2010), such that job search costs that are likely to be incurred in the event of firm termination are reduced (DiAddario, 2011; Wheeler, 2001). Much like their employees, entrepreneurs within clusters are also more likely to be able to job hop (Fallick *et al.*, 2006; Freedman, 2008; Wheeler, 2008), starting and closing firms with more regularity than in isolated locations. Third, beyond reducing such personal occupational switching costs, the co-location of specialized buyers and suppliers within geographic clusters also means that there is a more active secondary market for the startups' assets in the event of closure (Folta *et al.*, 2006b). This in turn increases the potential salvage value of the firm hence lowering termination costs (Folta *et al.*, 2006b). Put together, these set of factors generally lower the irreversibility of human and capital investments committed to the startup (O'Brien, Folta, and Johnson, 2003; Sandri *et al.*, 2010), and hence reduce switching costs for entrepreneurs located within geographic clusters.

Larger geographic clusters are also characterized by higher levels of peer-pressure (Porter, 2000). For instance, Porter (2000: 23) in discussing the behavioral incentives induced by geographic clustering suggests that the "competitive pressure in a cluster is amplified by peer-pressure, even among indirect or noncompeting firms. Pride and the desire to look good in the local community motivate firms to outdo each other". According to this perspective, social comparison processes within clusters are therefore viewed as relatively healthy and a positive motivating factor. However, in the extreme, such social comparison dynamics can also potentially lead to some deleterious dynamics such as Red Queen Effects, where firm aspirations continually escalate beyond realistically achievable targets (Barnett and Pontikes, 2004). Such goal escalation is particularly likely to be problematic in the startup context, where firms have limited slack resources (George, 2005), and a lower ability to overcome strategic errors caused by unrealistic goal-setting. For example, high levels of peer-pressure might accentuate the degree to which entrepreneurs are willing to undertake undue risks (Gardner and Steinberg, 2005), so as to meet their more ambitious performance expectations. They might also contribute to startups pursuing unsustainable strategies such as excessive growth (Churchill and Mullins, 2001; Markman and Gartner, 2002; Pierce and Aguinis, 2013) and engaging in head-to-head competition with dominant incumbents (Fan, 2010). And lastly, they can also lead to hubris among entrepreneurs (Hayward *et al.*, 2006), such that entrepreneurs are less able to set performance goals that are reflective of their actual capabilities (Cassar, 2014).

Integrating these arguments, I expect that the lower occupational switching costs as well as social comparison effects within larger clusters can in turn impact startup failure rates. To the extent that firm closures are at least in-part due to voluntary reasons (Bates, 2005; DeTienne *et al.*, 2008; Gimeno *et al.*, 1997; Headd, 2003), these clustering dynamics can impact what entrepreneurs view as acceptable levels of performance. In essence, I expect that startups operating within larger clusters must achieve a performance premium (Czarnitzki *et al.*, 2014) relative to more isolated firms so as to continue operations and not voluntarily close the firm. To the extent that startups are unable to consistently achieve these premiums, the positive relationship between geographic clustering and startup failure can be attributed to these higher expectations within geographic clusters. I test this relationship by examining post-entry behavioral dynamics and accordingly hypothesize that:

Hypothesis 2: Geographic clustering will be positively related to performance premium effects.

METHODS

Study Context and Data Sources

My primary source of information is the confidential data from the restricted access enclave of the Kauffman Firm Survey (KFS). The KFS is the only large, nationally representative dataset providing longitudinal information of startup firm financing and performance. It uses a complex stratified sample design with 6 groups (high-tech women owned, medium-tech women owned, low-tech women owned, high-tech male owned, medium-tech male owned, low-tech male owned) to identify representative startups based on the population of new businesses in the Dun and Bradstreet (D&B) database. In particular, it collects longitudinal data on a cohort of 4,928 startups founded in 2004, and surveys them annually, with data currently available up to 2011. While the startup and founder identities are anonymized, detailed information is available on variables such as the firm's industry (6 digit NAICS), physical location (zip code), demographics of up to ten owners, employment, patenting, financing structure, and measures of performance such as profitability. Remote access to this data was obtained upon request through the National Opinion Research Center (NORC). More details on this dataset and the sampling procedures using to identify representative startups are available at www.kauffman.org/kfs (also see Coleman and Robb, 2009; Robb and Watson, 2012).

The KFS dataset is well suited to test the hypotheses in this study for a number of reasons. First, given that it uses a longitudinal cohort design, it provides information on entrant characteristics across a representative sample of firms that began operations during the same calendar year. This information allows me to check for adverse selection dynamics unlike prior agglomeration research which has lacked such granular, historical data (Folta *et al.*, 2006a: 120). Second, the information also allows me to distinguish between both successful and unsuccessful exits, such as an exit by a sale or merger versus an exit by closure (Coleman, Cotei, and Farhat, 2013; Wennberg *et al.*, 2010). I am thus able to more confidently isolate voluntary failure events to test threshold models, rather than potentially confounding survival and failure dynamics (Cooper and Folta, 2000; DeTienne *et al.*, 2008; McCann and Folta, 2008). Third, the availability of data on firm finances also allows me to empirically estimate indicators of performance premium effects such as higher exit thresholds (Gimeno *et al.*, 1997; McCann and Folta, 2012). And finally, the broad geographic coverage of the KFS ensures that there is sufficient variance in terms of cluster location choices among firms.

I supplemented the data from the KFS with regional measures of industrial clustering using data from the Harvard Cluster Mapping Project (http://clustermapping.us). I used these cluster-based (i.e. sets of related industries) measures rather than individual NAICS based specifications of agglomeration, since they account for the fact that labor markets and location-based externalities span related industry sectors (Delgado *et al.*, 2014a). This publicly available dataset groups all 6-digit U.S. North American Industrial Classification System (NAICS) industries into 67 mutually exclusive traded (51 clusters) and local (16 clusters) industrial clusters[9] (Delgado *et al.*, 2014b). Establishment, employment, and wage levels along with

[9] While firms within *traded clusters* serve markets beyond the regions they are located in and are exposed to competition globally, firms within *local clusters* primarily sell locally (Delgado, Porter, and Stern, 2014a, 2014b).

related metrics (e.g. growth rates) are available at the county, core based statistical area (CBSA), and economic area (EA)[10] spatial levels. Using information on startup NAICS codes within the KFS, and the industry-cluster concordance tables on the Cluster Mapping website (http://www.clustermapping.us/content/cluster-mapping-methodology) I was able to assign firms to relevant clusters. Since the Harvard cluster mapping project uses NAICS codes based on 2007 definitions and the KFS started following firms in 2004, I also created mappings using concordance files that matched 2007 NAICS code definitions to 2002 NAICS code definitions. Next, to geo-locate startups and assign clusters to particular regions, I matched startups to specific Core Based Statistical Areas (CBSAs)[11]. I did so by using the physical location (zip code) of startups and zip-CBSA concordance tables from the MABLE database hosted by the Missouri Census data center (http://mcdc2.missouri.edu/websas/geocorr2k.html). Through this process, I was able to assign 4,620 of the 4,928 startups (94%) in the KFS to a specific CBSA-cluster combination. The 308 startups that did not match to a specific CBSA were located in rural counties, and were thus excluded from CBSA-level analyses. However, I did include these firms in sensitivity tests where I defined clusters at the EA level.

In addition to these cluster metrics I also compiled a set of time varying CBSA-level statistics from sources as the Census Bureau, USPTO, and National Center for Educational Statistics. These were used as control variables in estimation models.

---

[10] The Bureau of Economic analysis divides the U.S. into 171 Economic areas based on regional economic activity. See Alcacer and Chung (2007), or Delgado *et al*. (2014a) for more details on how these are defined.

[11] There are 929 such CBSAs in the U.S., including 388 Metropolitan Statistical Areas (MSAs) which are defined by urbanized areas in excess of 50,000 people and 541 Micropolitan Statistical Areas (µMSAs) that have populations in between 10,000-50,000. CBSA definitions were introduced in 2003 by the Office of Management and Budget (OMB) and are updated periodically.

Analyses

I carried out a range of statistical analyses to test the two hypotheses in this study. Each of these tests uses different dependent variables and covariates. I thus organize the analyses such that the relevant measures and statistical models are described for each hypothesis in turn.

Tests for adverse selection effects (hypothesis 1)

Measures

Geographic clustering: Testing for adverse selection effects required me to first identify geographic clusters vs. more isolated locations. I created such a classification by using a normalized measure of regional industrial density. In particular, following prior work (Alcácer and Chung, 2014; Delgado *et al.*, 2014b), I computed a location quotient (LQ) measure. The LQ measures the specialization or concentration of a cluster in a particular region relative to the national level. Regional (i.e. CBSA, EA, county) LQ measures for each of the 67 clusters are directly available from the Harvard cluster mapping project website for the years 1998-2012. LQ measures can also be computed from underlying census employment data using the following formulae for a given cluster *i* in year *t:*

$$\mathrm{LQ}_{\mathrm{i,t}} = \frac{\frac{\text{Employment in cluster in region (e.g.CBSA)}}{\text{Total employment in all clusters in the region (e.g.CBSA)}}}{\frac{\text{Employment in cluster in the U.S.}}{\text{Total employment in all clusters in the U.S.}}}$$

Since LQ values are calculated on a cluster-specific (i.e. sets of related industries) basis this method offers a robust way of identifying the impacts of both direct and indirect agglomeration effects for firms operation in different industry sectors. To test for adverse selection effects into clusters, I dichotomized this measure using data from the first KFS survey (year 2004). Given that the LQ is a normalized metric, I used a LQ score greater than 1 to identify geographic clusters vs. more isolated locations. I used this cutoff since a LQ value greater than 1 indicates that the degree of clustering in a region is higher than the national average. This LQ value also happened to be the sample median in effect allowing me to use a median split of my sample. In analytical models (see below) I also tried alternate cut-offs, for instance by defining truly large clusters as regions where the location quotient was above the 90th percentile in the sample (LQ>1.46 in the sample).

Models

Once I had thus split the sample of initial KFS entrants (year 2004) into two groups, that is those that entered geographic clusters vs. more isolated locations, I tested for adverse selection into clusters in two ways. First, I conducted a pairwise comparison (i.e. t-tests) of startups that entered into geographic clusters vs. more isolated locations. This descriptive analysis was used to examine whether observable entrant characteristics differed systematically for these two groups of startups. Second, again using the data from the initial KFS survey in 2004, I used multivariate logit based estimations to predict whether the likelihood of locating within a geographic cluster vs. a more isolated location was a function of startup-specific covariates, while controlling for other regional factors.

Tests for performance premium effects within clusters (hypothesis 2)

I conducted three different tests to study whether geographic clustering contributed to performance premium effects. In the first test, I used survival analyses methods to look at the combined effects of objective startup performance levels (risk-adjusted profitability) and geographic clustering on the likelihood (hazard rate) of failure. This analysis allowed me to examine whether, for comparable levels of economic performance, startups exposed to higher levels of geographic clustering were more likely to fail by closing down operations. In the second test, I used censored regression models to estimate latent exit thresholds (Gimeno *et al.*, 1997). This analysis allowed me to identify whether higher levels of geographic clustering led to startups raising their survival goals, in terms of minimal acceptable levels of performance. Lastly, in the third test, I looked at the combined effects of objective startup performance levels (risk-adjusted profitability) and geographic clustering levels on subjective self-rated performance ratings by startups that were in operation during the fourth follow up survey of the KFS (year 2008). This analysis allowed me to identify whether startups exposed to higher levels of geographic clustering where likely to be subjectively less "happy" with a given objective level of economic performance. I describe the measures and models used in each of these three tests in more detail below.

Test # 1 for performance premium effects: Survival analyses regressing startup failure on objective performance and geographic clustering levels

Dependent variable

Startup failure: Since the KFS is an annual survey of a cohort of startups all founded in 2004, it is able to track their exit behaviors over time. In particular, startups in the KFS are coded on an annual basis and their operational status is recorded till 2011. At a particular year, I am therefore able to observe whether a startup was still in operation, successfully sold or merged, or permanently closed operations. Given that sales and mergers are generally considered as successful exit events, I followed prior research and used the incidence of permanent closures to track startup failure events in my primary analyses (Coleman *et al.*, 2013; Robb and Watson, 2012). This variable thus took three values for a given sample year (i.e. 2005-2011), with startups still in operation coded as 0, closures coded as 1, and successful exits coded as 2 (set as a competing risk-refer to models section below). The variable was set to missing for startups that had already exited the sample in a prior year.

Independent variables

Startup performance (risk-adjusted profitability): Following prior literature that has examined startup performance using KFS data (Robb and Watson, 2012) I computed a time varying, risk-adjusted measure of profitability for each startup. In particular, I calculated a Sharpe rewards-to-variability ratio (Sharpe, 1975), by taking the ratio of the average profitability of a firm to the standard deviation in profitability over a four year temporal window. As Robb and Watson (2012) indicate, this risk-adjusted profitability measure is a particularly useful *objective* indicator of performance in the context of startup firms who typically have limited assets, and whose owners have a less than fully diversified investment portfolio. Furthermore, in comparison with other financial metrics that I computed from the data (e.g. ROA, ROE) this measure also exhibited the strongest correlation (r=0.14) with entrepreneurs' self-evaluations of performance assessed in the fourth KFS follow up survey in 2008[12].

Geographic clustering: To measure the level of geographic clustering that each startup was exposed to, I used a LQ measure as described earlier (see formula 1 on page 17). Given the longitudinal nature of the survival analyses, I computed a time-varying LQ measure (i.e. by cluster-year pair). I then assigned the appropriate LQ measure to all startups located within a particular geographic cluster (based on their industry and CBSA location).

Model

[12] For instance, the correlation between entrepreneurial self-assessments of performance with ROA and ROE was 0.02 and 0.04 respectively.

To study the likelihood of startup failure with survival analyses, I used hazard rate models. Following prior work that has used KFS data (Robb and Watson, 2012), I modelled the impacts of geographic clustering on the likelihood of startup failure using a cox-hazard specification with competing risks[13] (Lunn and McNeil, 1995). I also tried a variant of these conventional survival models, known as the accelerated failure time model (Wei, 1992). In contrast to the cox-hazard model which examines the impact of covariates (i.e. clustering levels) on the likelihood that particular events (i.e. startup failure) occur, the accelerated failure time model regresses the log of survival time directly over covariates. Thus while cox-hazard models can be used to understand whether startups in clusters are *more likely to fail* over the study period, accelerated failure time models can also indicate whether they also *fail more quickly*. In specifying these two sets of hazard models, I also accounted for the fact that the data in the KFS is based on a stratified random sample with 6 groups (high-tech women owned, medium-tech women owned, low-tech women owned, high-tech male owned, medium-tech male owned, low-tech male owned). This design allows the baseline hazard function to be unique to each stratum, while providing interpretable coefficient estimates across all strata.

Test # 2 for performance premium effects: Examining how clustering impacts firm-specific latent exit thresholds

Dependent variables

Startup performance (risk-adjusted profitability) and firm failure: While the threshold levels below which startups might choose to terminate businesses are latent (i.e. unobservable) and firm-specific, they can be estimated from two observable data points. In particular, I am able

[13] Competing risks here refer to the fact that startups might also exit the sample due to positive exit events such as sales or mergers. The competing risk specification thus models the likelihood of the closure event occurring while acknowledging that startups might also exit the sample due to this alternate reason.

to observe objective levels of startup performance (risk-adjusted performance) for startups that survive as well as specific failure events (observed closures). These two variables are the dependent variables in the estimation model, which uses a joint maximum likelihood specification to estimate coefficients (see model section below).

Independent variables

Geographic clustering: To measure the level of geographic clustering that each startup was exposed to I used a LQ measure as described earlier (see formula 1 on page 17).

Model

Using the information on the performance levels (risk-adjusted profitability) of surviving firms as well as observed firm closures, exit thresholds can be modeled using censored regression models (Gimeno *et al.*, 1997; Nelson, 1977). In particular, following prior literature (Folta and O'Brien, 2008; Gimeno *et al.*, 1997; McCann and Folta, 2012), I used a joint maximum likelihood specification to simultaneously estimate equations for objective firm performance (risk-adjusted profitability) as well as the latent exit threshold (for an example see Table 4 of McCann and Folta, 2012). Closely following the method used by Gimeno *et al.* (1997) who modeled thresholds for *startup exit decisions* I used cross-sectional data at two time points, t=0 (first KFS survey in 2004) for all firms and t=4 (fourth KFS follow up survey in 2008) for startups that survived till 2008. The objective economic measure of performance (risk-adjusted profitability) is therefore observed for startups that did not exit the sample during this time period, and censored for the other startups. I excluded any positive exit events (e.g. sales and mergers) prior to this analysis.

More formally, the performance level of a given firm in the sample (P_i) and the latent exit threshold (T_i) can be respectively modelled as:

$$P_i = B_1 X + v$$

$$T_i = B_2 X + u$$

where *X* refers to the common set of covariates across both equations observed at t=0 (first KFS survey in 2004). Note that for a given firm *i* in the sample, P_i is only observed when $P_{i>=}$ $T_{i\,,}$ and F_i =0 (where *F* is a binary variable representing where a startup has failed or not). The relevant model parameters are β_1, the coefficients of the covariates on risk-adjusted profitability *P*, $\beta_{2,}$ the coefficients of the covariates on the unobservable exit threshold *T*, s1 the standard deviation of the error term of the risk-adjusted profitability equation (*v)*, and s2 the standard deviation of the error term of the threshold equation (*u)*. Consistent estimates of the parameters of the performance and threshold equations can then be estimated if either a covariate is not shared across the two equations, or the covariance between the error terms *v* and *u* is set to 0. Given that I use the same set of covariates to predict both profitability *P* and the latent exit threshold *T,* I follow prior research (Folta and O'Brien, 2008; Gimeno *et al.*, 1997; McCann and Folta, 2012) and impose the second restriction (i.e. covariance (*v*,*u*)=0).

The probability of observing a startup failure event (i.e. *Prob* (F_i =1)) is:

$$\phi\ \left(\frac{\sum(\beta_{2i}-\beta_{1i})X_i}{s1^2+s2^2}\right)$$

where ϕ represents the normal cumulative distribution function, while the probability of observing a survival event (i.e. *Prob* (F_i =0)), that is a startup persisting with operations is:

$$\frac{1}{s1} Z\ \left(\frac{P_i - \sum B_{1i} X_i}{s1}\right) \phi \left(\frac{P_i - \sum B_{2i} X_i}{s2}\right)$$

where Z is the unit normal density function. The likelihood function aggregates these probabilities by multiplying them over all the sample observations. By taking the log-transformation of this function, I am able to arrive at a log-likelihood model in equation 6 below that can be maximized so as to estimate the parameters of the covariates within vector *X*:

$$\sum_{F_i=1} ln\left(\frac{\sum(\beta_{2i}-\beta_{1i})X_i}{s1^2+s2^2}\right) + \sum_{F_i=0} ln\left(\frac{1}{s1} Z\ \left(\frac{P_i-\sum B_{1i}X_i}{s1}\right)\phi\left(\frac{P_i-\sum B_{2i}X_i}{s2}\right)\right)$$

Test # 3 for performance premium effects: Multinomial logit models of subjective self-ratings of performance

Dependent variables

Subjective self-ratings of performance: For this final test, I used a specific survey question that was posed to startups that were in operation during the fourth-follow up survey of the KFS (year 2008). In particular, firms at this time were asked whether their performance over the first four years of the survey had exceeded, met, or fell below what they had anticipated at the time of founding. This particular survey item has been used in prior research that has looked at the impacts of industry and startup experience on the ability of entrepreneurs to accurately forecast performance into the future (Cassar, 2014). I used these survey responses as a subjective self-rated measure of satisfaction. This variable is therefore an ordinal variable that takes three values.

Independent variables

Startup performance (risk-adjusted profitability) & geographic clustering: Similar to the approach followed in the survival analyses described above (test#1 for adverse selection effects), I used the objective level of startup performance (risk-adjusted profitability) as well as the degree of geographic clustering that a firm was exposed to as the main model covariates.

Model

I modelled self-rated satisfaction ratings as the dependent variable in a multinomial logit model with objective performance (risk-adjusted profitability) and geographic clustering variables as the main predictors. The multinomial logit model requires a base category of the dependent variable to be specified, and computes the likelihood of the other values occurring relative to this base value. Given my interest in studying subjective positive and negative reactions, I set the base level to the intermediate value of the dependent variable (i.e. performance just met expectations).

Control Variables (for both adverse selection and performance premium effects)

I included a set of both firm and regional level controls in all estimation models, testing for both adverse selection and performance premium effects. At the startup level, following prior literature that has used KFS data (Cassar, 2014; Coleman *et al.*, 2013; Robb and Watson, 2012; Shah and Smith, 2010), I accounted for factors such as the *total number of* employees, founding team *entrepreneurial experience* (average number of ventures founded), *work experience* (average number of years worked), *age, gender, gender similarity, racial similarity,* and *average education level.* I also controlled for additional variables such as whether startups possessed any *intellectual property*, whether they were a *sole proprietorship*, and the industry sector within which they operated (using 2-digit NAICS codes). At the region level, I included controls for *affluence*, *racial similarity*, *innovation*, *education* as well as a dummy code for whether the region was an MSA or an MISA to capture urbanization effects (see footnote 11).

RESULTS

Descriptive sample statistics

Table 7 below shows descriptive characteristics of the 4,620 startups in the KFS located within CBSAs at the time of the first survey in 2004.

Table 7 (paper 3). Demographics of startups (n=4,620) and regional environment in the initial KFS survey (year 2004)

Variable	Mean	Std
Startup (Founding Team) Demographics		
Total Employees	2.52	6.01
Entrepreneurial experience (number of firms started)	0.78	1.17
Work experience (years)	11.49	9.84
Age	44.22	10.25
Gender	0.68	0.42
Gender similarity (herfindahl index)	0.92	0.18
Racial similarity (herfindahl index)	0.98	0.09
Education of founding team (0 to 10)	6.10	2.04
Possesses intellectual property (dummy variable)	0.19	0.40
Sole propietorship (dummy variable)	0.35	0.48
Regional (CBSA) Characteristics		
Cluster-CBSA geographical clustering level (location quotient)	1.03	1.03
Metropolitan vs. Micropolitan Area (dummy variable)	0.90	0.29
Innovation (Number of patents per capita) (X 1,000)	0.32	0.41
Affluence level (Number of people in poverty per capita)	0.12	0.03
Racial similarity (CBSA) (herfindahl index)	0.31	0.13
Education levels (number of research universities per capita) (X 100,000)	0.88	1.71
Industry Breakdown		
Professional, Scientific, &Technical Services	16.57%	
Retail	14.19%	
Construction	11.38%	
Admin, support, waste management, and remediation services	10.10%	
Other Services (except public administration)	9.34%	
Manufacturing	6.36%	
Wholesale	5.49%	
Finance and insurance	5.34%	
Real estate, rental, and leasing	5.11%	
Health care and social assistance	3.22%	
Information	3.21%	
Transportation and warehousing	2.79%	
Accomodation and food services	2.58%	
Art, entertainment and recreation	2.43%	
Educational services	0.69%	
Agriculture, forestry, fishing, hunting	0.67%	
Management of companies & enterprises	0.23%	
Mining, quarrying, oil & gas extraction	0.13%	
Utilities	0.10%	
Public administration	0.07%	

The average startup in the sample had 2-3 employees (2.52 ± 6.01). The typical founding team had relatively little prior entrepreneurial experience in terms of the number of ventures founded (0.78±1.17), but had approximately 11 years of work experience (11.49±9.84). Founding teams were typically middle-aged (44.22 ±10.25), and were predominantly composed of males (0.68±0.42). Herfindahl indices also indicated that founding teams were largely homogenous in terms of gender (0.92 ± 0.18) and race (0.98 ± 0.09). In terms of formal education, the average founding team had an associate's level degree (6.10±2.04). 35% of the entrants were sole proprietorships, while 19% of the firms possessed some form of intellectual property (e.g. patents, copyrights, trademarks). In terms of the three major industries represented in the sample, 16.57% are in professional, scientific and technical services (2-digit NAICS code 54), 14.19% in retail (2-digit NAICS code 44 and 45), and 11.38% in construction (2-digit NAICS code 23).

In terms of CBSA characteristics, the average level of geographic clustering (based on paired cluster-CBSA location quotients) is 1.03 ± 1.03. A closer examination of this variable indicated that 90% of startups are located in CBSAs with a LQ value of <1.46, and 99% of startups are located in CBSAs with a location quotient <4.69. 90% of startups are located within an MSA (i.e. urban areas with population larger than 50,000) while 10% are in MISAs (urban areas with population between 10,000 and 50,000). A herfindahl index of metro-region racial diversity (0: highly diverse, 1: not diverse) indicated that startups were in locations with moderate to high levels of racial diversity (0.31 ± 0.13). The average poverty rate was 12 ± 3%. Lastly indicators of education and innovation showed that on average, there were 0.88 (± 1.71) research universities per 100,000 people, as well as 0.32 (±0.41) patents granted per 1,000 people.

Test # 1 for adverse selection effects (t-tests for group mean differences)

Table 8 shows results from pairwise comparisons of means (t-tests) of entrants into clustered vs. more isolated locations. I first used a LQ of 1, which was also the sample median, as a cutoff to define locations with higher than average levels of clustering. As shown by the third column of table 8, I find no marked evidence of adverse selection dynamics. The only variable that is significantly different between the two groups of startups is the founding team racial similarity (mean difference of -0.1, significant at p=0.1). However, in terms of the industry breakdown I do find some statistically significant differences. Firms in professional and scientific services, waste remediation, and other service sectors were more likely to locate in regions with higher than average industrial concentration (LQ>1), while firms in information technology, health care, and manufacturing were less likely to do so.

Table 8 (paper 3). Test # 1 for adverse selection: Mean-differences of entrants (t-test) into clusters vs. more isolated locations

	Sample split at LQ=1 (also the sample median)			Sample split at LQ=1.46 (sample 90th percentile)		
	Model 1: Isolated (n=2,342)	Model 2: Clustered (n=2,278)	Difference in group means	Model 3: Isolated (n=4,155)	Model 4: Clustered (n=465)	Difference in group means
Startup (founding team) demographics						
Total Employees	2.44	2.60	0.16	2.45	3.44	0.09*
	(8.01)	(9.07)		(5.83)	(36.74)	
Entrepreneurial experience (number of firms started)	0.76	0.80	0.04	0.76	1.01	0.25***
	(1.65)	(1.65)		(1.19)	(5.15)	
Work experience (years)	11.59	11.38	-0.21	11.36	13.17	1.81***
	(14.12)	(13.73)		(10.18)	(37.74)	
Age	44.50	43.94	-0.56	44.15	45.10	-0.05
	(14.36)	(14.65)		(10.64)	(38.36)	
Gender	0.68	0.68	0	0.68	0.74	0.06**
	(0.60)	(0.60)		(0.44)	(1.48)	
Gender similarity (herfindahl index)	0.93	0.92	-0.01	0.92	0.94	0.02
	(0.25)	(0.25)		(0.18)	(0.61)	
Racial similarity (herfindahl index)	0.99	0.98	-0.01*	0.98	0.98	0
	(0.11)	(0.13)		(0.09)	(0.38)	
Education of founding team (0 to 10)	6.06	6.14	0.08	6.07	6.52	0.45***
	(2.97)	(2.79)		(2.12)	(7.00)	
Possesses intellectual property (dummy variable)	0.19	0.19	0	0.19	0.26	0.07**
	(0.56)	(0.56)		(0.41)	(1.62)	
Sole propietorship (dummy variable)	0.36	0.34	-0.02	0.36	0.29	-0.07**
	(0.68)	(0.67)		(0.50)	(1.68)	
Regional (CBSA) characteristics						
Cluster-CBSA geographical clustering level (location quotient)	0.69	1.37	0.68***	0.91	2.62	1.71***
	(0.44)	(1.90)		(0.35)	(12.06)	
Metropolitan vs. Micropolitan Area (dummy variable)	0.87	0.94	0.07***	0.91	0.88	-0.03
	(0.48)	(0.33)		(0.30)	(1.19)	
Innovation (Number of patents per capita) (X 1,000)	0.30	0.34	0.04**	0.32	0.34	0.02
	(0.58)	(0.60)		(0.43)	(1.76)	
Affluence level (Number of people in poverty per capita)	0.12	0.12	0	0.12	0.12	0
	(0.04)	(0.04)		(0.03)	(0.13)	
Racial similarity (CBSA) (herfindahl index)	0.30	0.33	0.03***	0.31	0.34	0.03***
	(0.20)	(0.19)		(0.14)	(0.57)	
Education levels (number of research universities per capita) (X 100,000)	0.88	0.88	0	0.86	1.07	0.19
	(2.75)	(2.15)		(1.68)	(9.66)	
Industry Breakdown						
Mining, quarrying, oil & gas extraction	0.00	0.00	0	0.00	0.01	0.01
	(0.05)	(0.06)		(0.03)	(0.30)	
Utilities	0.00	0.00	0	0.00	0.00	0
	(0.05)	(0.03)		(0.03)	(0.22)	
Construction	0.11	0.12	0.01	0.12	0.05	-0.07***
	(0.44)	(0.45)		(0.34)	(0.81)	
Wholesale	0.05	0.06	0.01	0.06	0.03	0.03**
	(0.31)	(0.33)		(0.24)	(0.65)	
Information	0.04	0.02	-0.02***	0.03	0.07	0.04**
	(0.29)	(0.21)		(0.18)	(0.92)	
Finance and insurance	0.05	0.05	0	0.05	0.12	0.07***
	(0.32)	(0.32)		(0.22)	(1.22)	
Real estate, rental, and leasing	0.05	0.06	0.01	0.05	0.01	-0.04***
	(0.30)	(0.33)		(0.23)	(0.43)	
Professional, Scientific, &Technical Services	0.15	0.18	0.03***	0.16	0.28	0.12***
	(0.50)	(0.55)		(0.38)	(1.66)	
Management of companies & enterprises	0.00	0.00	0	0.00	0.00	0
	(0.07)	(0.06)		(0.05)	(0.00)	
Admin, support, waste management, and remediation services	0.09	0.11	0.02**	0.11	0.05	-0.06***
	(0.40)	(0.45)		(0.32)	(0.79)	
Educational services	0.01	0.01	0	0.01	0.01	0
	(0.12)	(0.11)		(0.09)	(0.26)	
Health care and social assistance	0.04	0.02	-0.02**	0.03	0.01	-0.02***
	(0.28)	(0.22)		(0.19)	(0.28)	
Art, entertainment and recreation	0.03	0.02	-0.01	0.02	0.04	0.02*
	(0.23)	(0.20)		(0.16)	(0.74)	
Accomodation and food services	0.03	0.02	-0.01	0.03	0.02	-0.01
	(0.24)	(0.21)		(0.17)	(0.53)	
Other Services (except public administration)	0.08	0.10	0.02**	0.10	0.02	-0.08***
	(0.39)	(0.43)		(0.31)	(0.45)	
Public administration	0.00	0.00	0	0.00	0.00	0
	(0.05)	(0.00)		(0.03)	(0.00)	
Manufacturing	0.08	0.05	-0.03***	0.06	0.17	0.11***
	(0.38)	(0.30)		(0.24)	(1.40)	
Retail	0.15	0.14	-0.01	0.15	0.08	-0.07***
	(0.50)	(0.49)		(0.37)	(1.01)	
Transportation and warehousing	0.03	0.03	0	0.03	0.04	0.01
	(0.24)	(0.23)		(0.17)	(0.69)	

Standard deviations in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

To verify the robustness of these results to an alternate specification of clustered vs. isolated locations, I re-ran the analysis but this time split it on the 90th percentile of the sample LQ variable (LQ=1.46). This specification thus tests for difference in entrant profiles, into very large geographic clusters vs. other locales (e.g. smaller geographic clusters and isolated locations). With this alternate specification, I do find significant and consistent group differences. However the evidence points towards positive (i.e. self-selection) rather than negative (i.e. adverse-selection) dynamics into the largest of clusters. For instance, I find that startups in the largest clusters (defined as having a LQ>=1.46, in the 90th percentile or above) tend to have more employees (mean difference of 0.09, significant at p=0.1), have more entrepreneurial experience (mean difference of 0.25, significant at p=0.01), have more work experience (mean difference of 1.81, significant at p=0.01), have higher levels of education (mean difference of 0.45, significant at p=0.01), and are more likely to possess intellectual property (mean difference of 7%, significant at p=0.05) relative to peers outside large clusters. I also find that they are more likely to be male owned (mean difference of 6%, significant at p=0.05), and less likely to be a sole proprietorship (mean difference of -7%, significant at p=0.05). I also find consistent industry differences. For instance, the difference among scientific and technical services is 12% (significant at p=0.01), and manufacturing is 11% (significant at p=0.01) indicating that firms in these industries might actually be selecting into very large clusters.

Test # 2 for adverse selection effects (logit models)

Table 9 shows results from logit models modeling entry choice into clustered vs. more isolated locations.

Table 9 (paper 3). Test # 2 for adverse selection into clusters: Logit and multinomial logit models of startup location choice

	Baseline is "isolated" regions (LQ<=1)	Baseline level is "small"	
	Model 1: logit	Model 2: mlogit	Model 3: mlogit
VARIABLES	LQ >1	LQ<=1	LQ>1.46
Startup Variables			
Total Employees	0.01	-0.00	0.01
	(0.01)	(0.01)	(0.01)
Entrepreneurial Experience	0.07**	-0.03	0.13**
	(0.03)	(0.03)	(0.06)
Work experience	-0.00	0.00	0.02**
	(0.00)	(0.00)	(0.01)
Age	-0.00	0.01	-0.00
	(0.00)	(0.00)	(0.01)
Gender similarity	-0.02	0.09	0.43
	(0.23)	(0.23)	(0.48)
Racial similarity	-0.65	0.65	0.29
	(0.42)	(0.44)	(0.81)
Education of founding team	0.00	0.02	0.08**
	(0.02)	(0.02)	(0.03)
Has intellectural property	0.08	0.09	0.24
	(0.10)	(0.10)	(0.16)
Sole Propietorship	-0.01	-0.00	-0.14
	(0.09)	(0.09)	(0.16)
Regional Variables			
MSA location (versus MISA location)	0.39***	-0.65***	-1.25***
	(0.14)	(0.15)	(0.27)
Number of patents per capita (X1,000)	0.11	-0.10	-0.02
	(0.10)	(0.11)	(0.13)
Poverty rate	-2.75*	1.47	-8.18**
	(1.48)	(1.46)	(3.78)
Racial similarity (CBSA)	1.59***	-1.21***	1.87***
	(0.30)	(0.30)	(0.62)
Education levels (number of research universities per capita) (X100,000)	0.01	0.02	0.07**
	(0.02)	(0.02)	(0.04)
Constant	-2.81**	-0.20	-1.82*
	(1.15)	(0.54)	(0.95)
	(1.19)	(0.63)	(1.21)
Industry dummies included	Yes	No	No
Observations	4180	4183	4183

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Model 1 of table 9 shows findings from a specification using a LQ cutoff value of 1 to define more clustered locations. Among the startup demographic variables, I find that startups with more entrepreneurial experience (b=0.07, p=0.01) are more likely to locate in clustered locations. In models 2 and 3, I tried an alternate multinomial logit specification based on a two level cutoff for the location quotient variable. Specifically, I examined the relative likelihood of firms locating in isolated (LQ<=1), small (1<LQ<=1.46), and the largest (LQ>1.46) of geographic clusters. I used the intermediate level (i.e. small clusters with LQ values in the range of 1-1.46) as the base level for this specification. In model 2, I do not find any evidence indicative of observable differences in characteristics between startups in isolated locations and small clusters. In model 3, I find evidence that startups with more entrepreneurial experience (b=0.13, p<0.05), more work experience (b=0.02, p<0.05), and higher levels of education (b=0.08, p<0.05) are more likely to locate in large vis-à-vis small clusters.

In aggregate these tests find little evidence for adverse selection dynamics into clusters. Thus I do not find empirical support for hypothesis 1. On the other hand, given that attributes such as entrepreneurial experience, work experience, and educational levels are typically used as indicators of the level of general and specific human capital of founding teams (Colombo and Grilli, 2010), the results from model 3 suggest that there might actually be positive (i.e. self-selection) selection dynamics into very large clusters. I return to the implications of these findings in the discussion section.

Test # 1 for performance premium effects (survival analyses)

I examined the combined impact of objective startup performance (risk-adjusted profitability) and geographic clustering on startup failure using both cox-hazard and accelerated failure time survival models. Table 10 shows bivariate correlations of variables used in these analyses, while table 11 shows results from the hazard rate analyses.

Table 10 (paper 3). Correlation matrix and descriptive statistics for data used in survival analyses (test # 1 for performance premium effects)

Variables	Mean	Std	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1.Firm Closure	0.3	1.03																
2.Total Employees	3.85	10.23	-0.03***															
3.Entrepreneurial Experience	0.8	1.17	0.00	0.07***														
4.Work experience	13.16	10.24	-0.05***	0.05***	0.11***													
5.Age	47.83	10.54	-0.01†	-0.02*	0.17***	0.39***												
6.Gender similarity	0.92	0.18	0.01†	-0.05***	0.01†	0.13***	-0.01											
7.Racial similarity	0.98	0.09	-0.00	-0.07***	-0.02*	0.04***	0.04***	0.22***										
8.Education of founding team	6.52	1.98	-0.04***	0.02*	0.06***	0.05***	0.12***	0.03***	-0.03***									
9.Has intellectual property	0.22	0.41	-0.02*	0.09***	0.09***	0.01	0.01†	-0.03***	-0.05***	0.14***								
10.Sole proprietorship	0.31	0.46	-0.02*	-0.17***	-0.08***	-0.04***	0.04***	0.30***	0.13***	-0.14***	-0.10***							
11.Metropolitan/Micropolitan area	1.9	0.31	0.00	0.00	0.02*	0.03***	-0.02***	-0.00	-0.03***	0.12***	0.06***	-0.08***						
12.Number of patents per capita (X1,000)	0.33	0.46	-0.01	0.04***	-0.01	0.03***	-0.01	0.01	-0.04***	0.10***	0.07***	-0.01†	0.24***					
13.Poverty rate	0.12	0.03	-0.02***	0.06***	0.03***	-0.03***	0.01	-0.01	-0.01†	-0.08***	-0.02*	0.05***	-0.12***	-0.23***				
14.Racial similarity (CBSA)	0.31	0.14	0.00	0.03***	-0.02***	0.02*	-0.02†	0.05***	-0.05***	0.15***	-0.01	-0.05***	0.38***	0.13***	0.03***			
15.Education levels (number of research universities per capita) (X100,000)	0.83	1.86	0.01	-0.02*	-0.03***	0.01	0.02***	-0.03***	0.01	0.06***	0.02*	-0.03***	-0.02*	0.07***	0.04***	-0.04***		
16. Startup performance (risk-adjusted profitability)	0.54	5.69	-0.03**	-0.00	-0.04***	0.03**	-0.01†	0.03***	0.02*	-0.01	-0.04***	0.02**	-0.00	0.00	0.00	0.00	-0.01	
17.Cluster location quotient	1.05	1.14	0.00	0.09***	0.04***	0.06***	0.02*	0.01†	-0.03***	0.01†	0.06***	-0.04***	0.02***	0.07***	0.00	0.04***	0.00	0.00

†p<0.1, *p,0.05, **p<0.01, ***p<0.001

With respect to firm closure and statistically significant effects in table 10, I find that startup closure is negatively correlated to the total number of employees (r=-0.03), founding team work experience (r=-0.05), age (r=-0.01), founding team education (r=-0.04), having IP (r=-0.02), being a sole proprietor (r=-0.02), being located in regions with a higher poverty rate (r=-0.02), and having a higher level of objective economic performance in terms of profitability (r=-0.03). It is also positively correlated with founding team gender homogeneity (r=0.01).

Table 11 (paper 3). Test # 1 for clustering and performance premium effects: Survival analyses

	Only Closures					**All Exits (Sales, Mergers, Closures)**	
VARIABLES	Model 1: **Cox Hazard model** with controls	Model 2: **Cox Hazard model** with main effects	Model 3: **Cox Hazard model** with interaction effects	Model 4: **Accelerated failure time model** with main effects	Model 5: **Accelerated failure time model** with interaction	Model 6: **Cox Hazard model** with interaction effects	Model 7: **Accelerated failure time model** with interaction
Startup Controls							
Total Employees	-0.02*	-0.03**	-0.03**	0.02*	0.02*	-0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Entrepreneurial Experience	-0.01	0.01	0.01	-0.02	-0.02	0.04	-0.03
	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Work experience	-0.02***	-0.02***	-0.02***	0.02***	0.02***	-0.02***	0.02***
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
Age	0.01	0.01	0.01	-0.00	-0.00	0.01	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Gender similarity	0.33	0.23	0.23	-0.22	-0.22	0.32	-0.12
	(0.24)	(0.29)	(0.29)	(0.30)	(0.30)	(0.26)	(0.25)
Racial similarity	-0.24	-0.41	-0.42	0.43	0.43	-0.43	0.31
	(0.41)	(0.48)	(0.48)	(0.51)	(0.51)	(0.43)	(0.41)
Education of founding team	-0.05**	-0.05**	-0.05**	0.08***	0.08***	-0.05**	0.06***
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)
Has intellectual property	-0.18*	-0.14	-0.14	0.10	0.10	-0.11	-0.01
	(0.11)	(0.13)	(0.13)	(0.13)	(0.13)	(0.11)	(0.11)
Sole Propietorship	-0.18**	-0.19*	-0.19*	0.11	0.11	-0.24**	0.22**
	(0.09)	(0.11)	(0.11)	(0.12)	(0.12)	(0.10)	(0.10)
Region (CBSA) controls							
MSA location (versus MISA location)	0.20	0.26	0.27	-0.34*	-0.34*	0.13	-0.16
	(0.15)	(0.18)	(0.18)	(0.20)	(0.20)	(0.16)	(0.16)
Number of patents per capita (X1,000)	-0.17	-0.13	-0.13	0.32	0.32	-0.05	0.16
	(0.13)	(0.16)	(0.16)	(0.21)	(0.21)	(0.12)	(0.15)
Poverty rate	-0.32	0.54	0.55	2.98*	3.00*	0.24	2.98*
	(1.28)	(1.60)	(1.59)	(1.77)	(1.77)	(1.44)	(1.52)
Racial similarity (CBSA)	0.23	0.40	0.39	-0.26	-0.25	0.52*	-0.26
	(0.28)	(0.34)	(0.34)	(0.37)	(0.37)	(0.31)	(0.31)
Education levels (number of research universities per capita) (X100,000)	0.01	0.02	0.02	-0.05***	-0.05***	0.02	-0.04***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)
Firm profitability and Clustering							
Risk-adjusted profitability		-0.02**	-0.03**	0.03***	0.05***	-0.02	0.02
		(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Cluster location quotient		0.09***	0.09***	-0.08***	-0.07***	0.07**	-0.05***
		(0.04)	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)
Cluster location quotient*Risk-adjusted profitability			0.01*		-0.014*	0.00	-0.00
			(0.006)		(0.008)	(0.01)	(0.01)
Constant				2.03***	2.02***		1.48***
				(0.64)	(0.64)		(0.52)
Number of startups	2782	2388	2388	2388	2388	2388	2388
Industry dummies included	Yes	Yes	Yes	No[a]	No[a]	Yes	No[a]
Number of CBSAs	528	528	528	528	528	528	528
Number of industrial clusters	63	63	63	63	63	63	63
Total Observations (startup-year)	15,469	11,950	11,950	11,950	11,950	11,950	10,663

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

[a]I could not include industry dummies in the accerated failure time models (non-symmetric variance matrix error)

Models 1-5 of table 11 shows results from hazard rate specifications where firm closures are used to capture firm failure. Models 1-3 shows data from cox hazard specifications. Positive coefficients in these equations correspond to an increased likelihood of failure. Models 4-5 show data from accelerated failure time specifications. Negative coefficients in these equations correspond to a lowering of the survival time (i.e. faster failure). In models 6-7, I also present results from sensitivity tests where all forms of exit (e.g. sales, mergers, closures) are treated as failures.

In model 1, where only control variables are included, a number of variables are statistically significant. I find that startups with more employees ($b=-0.02$, $p<0.10$), higher average work experience ($b=-0.02$, $p<0.01$), more educated founding teams ($b=-0.05$, $p<0.05$), those that have some form of intellectual property ($b=-0.18$, $p<0.10$), and those that are sole proprietorships are less likely to fail ($b=-0.18$, $p<0.05$). In model 2 of table 11, I add the risk-adjusted profitability and geographic clustering variables to this model to examine their main effects. I find that firms that have higher levels of risk-adjusted profitability are less likely to fail ($b=-0.02$, $p<0.05$). On the other hand, the degree of geographic clustering has a positive effect on the likelihood of failure ($b=0.09$, $p<0.01$). This coefficient corresponds to a hazard rate of 1.09 which means that there is a 2.2% increase in the baseline probability of failure for a 1 unit increase in the cluster LQ. I note here that these results are as expected and in accordance with conventional theories of failure caused by agglomeration-based diseconomies in larger clusters.

However, when I interact the firm profitability and clustering variables, I find some interesting effects indicating that failure dynamics might be more complex and driven by behavioral factors. I present these results in model 3-5 of table 11. In model 3, I observe that the interaction term has a positive effect ($b=0.01$, $p=0.06$). This means that a 1 unit increase in the cluster LQ attenuates the negative relationship between startup profitability and the likelihood of failure. To confirm these interaction models, I ran alternate accelerated failure-time specifications in models 4 and 5. Note that negative coefficients in these models correspond to a lowering of the survival time (i.e. faster failure). As indicated in model 5, I find that the interaction term between geographic clustering and startup profitability is significant (b=-0.014, p<0.1). I graphically illustrate these effects in Figure 5 below.

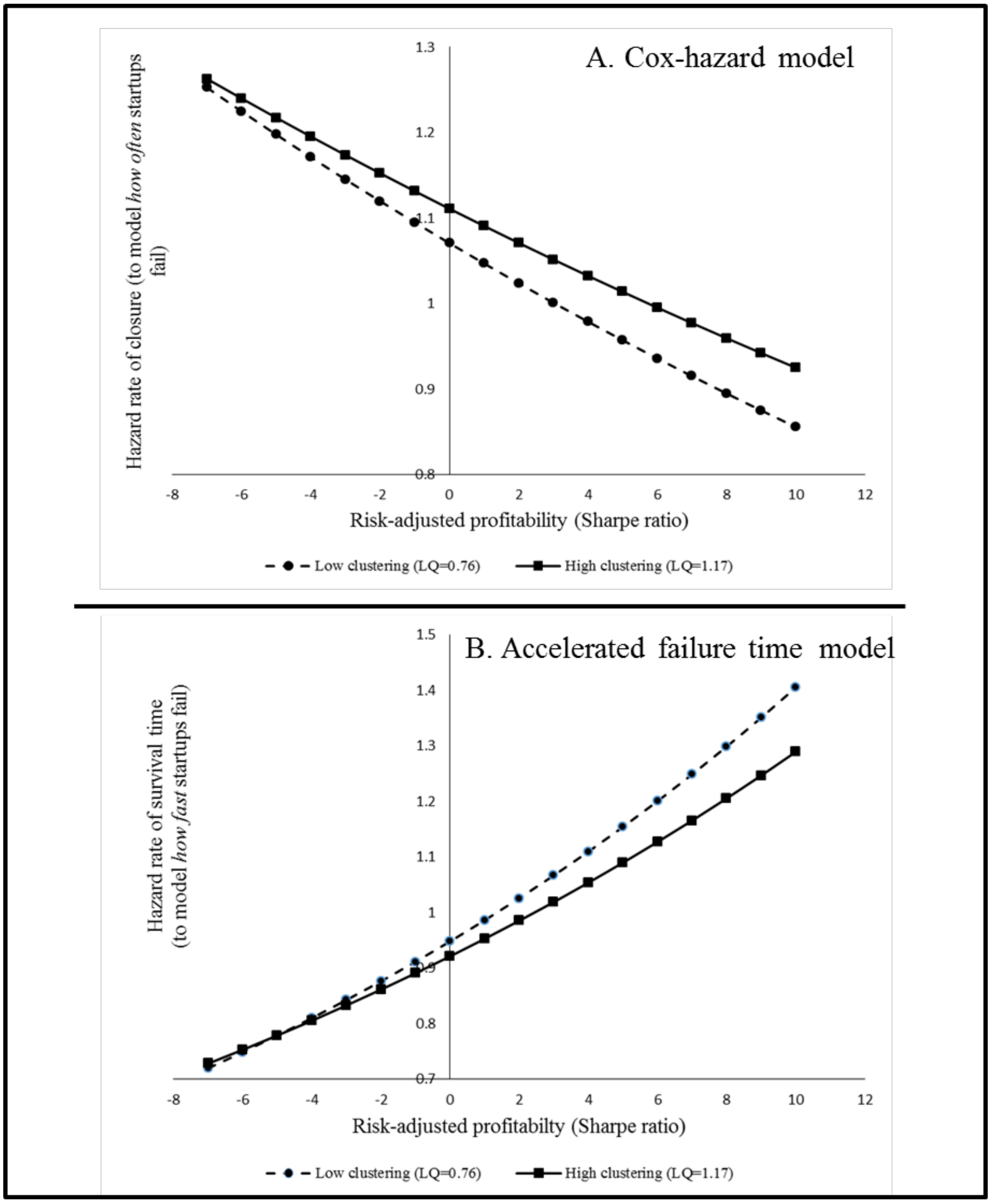

Figure 5 (paper 3). Survival analyses interaction plots. (A) Cox-hazard and (B) accelerated failure time models of startup failure as function of risk-adjusted profitability. The degree of clustering that the startup is exposed to both increases the likelihood of failure as well as lowering the time that elapses before failure occurs (i.e. leads to *quicker* failure)

In panel A I observe that, as expected, startups that are more profitable have a lower hazard of failure (i.e. a negative relationship). However, this effect is weaker for startups that are exposed to higher levels of clustering (solid line) compared to peers exposed to lower levels of clustering (dotted line). In panel B I observe that, as expected, startups that are more profitable are likely to survive for longer periods (i.e. a higher hazard rate). However, this effect is again weaker for startups exposed to higher levels of clustering (solid line) indicating that they are likely to *fail faster*¸ for comparable levels of performance.

Lastly, in models 6 and 7 of table 11 I treated both sales and mergers as possible failures (Bates, 2005; Cooper and Folta, 2000; McCann and Folta, 2008; Wennberg *et al.*, 2010), and ran cox-hazard and accelerated failure time models respectively to test the sensitivity of choosing only closures as a measure of failure. In effect, these models look at all survivors, and are agnostic about whether firms are no longer in operation due to positive or negative exits. While the main effects of the clustering variable remain significant in these models ($b=0.07$, $p<0.05$, and $b=-0.05$, $p<0.01$ respectively), I find that neither the profitability variable nor the interaction effects are significant. This finding suggests that while profitability has a strong impact on closure decisions it has less of an impact on positive exit decisions. It is also consistent with prior research that indicates that researchers should be careful to not conflate survival and failure dynamics (Cooper and Folta, 2000; Folta *et al.*, 2006a; McCann and Folta, 2008).

In aggregate, these results provide support for the argument that for a given level of performance, startups exposed to higher levels of geographic clustering are more likely to voluntarily terminate operations.

Test # 2 for performance premium effects (positive impacts of clustering on exit thresholds)

Table 12 shows results from joint maximum likelihood models estimating the impacts of clustering on both startup performance (risk-adjusted profitability) and latent exit thresholds.

Table 12 (paper 3). Test # 2 for clustering and performance premium effects: Joint maximum likelihood estimation of risk-adjusted profitability (model 1) and exit thresholds (model 2) in KFS survey 4 (year 2008) based on initial conditions (KFS survey 1, 2004)

VARIABLES	Model 1 Risk-adjusted profitability (year 2008)	Model 2 Exit threshold (year 2008)
Startup controls (year 2004)		
Total Employees	0.31	-0.24
	(0.39)	(0.20)
Entrepreneurial Experience	1.00**	-0.32*
	(0.42)	(0.17)
Work experience	-0.03	0.01
	(0.03)	(0.01)
Age	-0.32	-0.09
	(0.20)	(0.08)
Gender similarity	0.06*	0.02
	(0.03)	(0.01)
Racial similarity	0.05	-0.03***
	(0.03)	(0.01)
Education of founding team	-0.33	0.73
	(1.02)	(0.68)
Has intellectual property	1.40	-0.59
	(2.00)	(0.64)
Sole Propietorship	0.35*	-0.10*
	(0.21)	(0.06)
MSA controls (year 2004)		
MSA location (versus MISA location)	-0.65	-0.39*
	(0.58)	(0.22)
Number of patents per capita (X1,000)	0.03	0.02
	(0.40)	(0.18)
Poverty rate	-1.65	0.14
	(5.95)	(3.40)
Racial similarity (CBSA)	-2.27	0.74
	(1.40)	(0.84)
Education levels (number of research universities per capita) (X100,000)	-0.13	-0.29
	(0.16)	(0.29)
Geographic clustering effect (year 2004)		
Cluster location quotient	0.02	0.15***
	(0.13)	(0.05)
Constant	15.50***	24.25***
	(5.48)	(0.98)
Observations	4,026	4,026

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note that given the two-period cross-sectional comparison approach used to estimate thresholds (Gimeno *et al.*, 1997; McCann and Folta, 2012), covariates in these models are treated as time-invariant (i.e. I use initial conditions at year 2004), with the risk-adjusted profitability values only observed for firms that do not fail before the second observation time period (year 2008). In model 1, I find that higher levels of initial entrepreneurial experience ($b=1.00$, $p<0.05$), higher gender similarity ($b=0.06$, $p<0.1$) and being a sole proprietorship ($b=0.35$, $p<0.1$) has a positive impact on risk-adjusted profitability for firms that were in operation during the fourth follow up survey (year 2008). Interestingly, I find that the degree of clustering in itself has an insignificant impact on risk-adjusted profitability ($b=0.02$).

Turning to the exit thresholds in model 2, I find that having more entrepreneurial experience ($b=-0.32$, $p<0.1$), founding team racial similarity ($b=-0.03$, $p<0.01$), and being a sole proprietorship ($b=-0.10$, $p<0.1$) have a negative impact on exit thresholds. However, the degree of clustering has a positive and significant impact on exit threshold levels ($b=0.15$, $p<0.01$). I thus find support for the argument that startups exposed to higher degrees of clustering will have higher performance expectations, reflected in the manner in which they set exit threshold levels.

Test # 3 for performance premium effects (joint effects of objective performance and clustering on subjective self-rated levels of performance)

I also ran a final test to investigate the impact of geographic clustering on the extent to which startups in operation in the fourth follow up of the KFS (year 2008) were subjectively satisfied with the performance they had achieved over the first four years of operation. Models 1-4 of table 13 present the results from these analyses.

Table 13 (paper 3). Test #3 for clustering and performance premiums effects: Impact of risk-adjusted profitability on self-rated satisfaction levels. Data from cross-sectional analysis of startups still operating in the fourth follow up survey (year 2008)

	Impact of profitability on satisfication ratings		Interaction with degree of clustering	
	Model 1	Model 2	Model 3	Model 4
VARIABLES	High satisfaction: Exceeded growth expectations	Low satisfaction: Below growth expectations	High satisfaction: Exceeded growth expectations	Low satisfaction: Below growth expectations
Startup controls				
Total Employees	0.02*	-0.01	0.02	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Entrepreneurial Experience	0.14*	0.12**	0.13	0.09
	(0.08)	(0.06)	(0.08)	(0.06)
Work experience	0.00	-0.01*	0.00	-0.02**
	(0.01)	(0.01)	(0.01)	(0.01)
Age	-0.01	0.02**	0.00	0.02***
	(0.01)	(0.01)	(0.01)	(0.01)
Gender similarity	0.53	0.74*	0.67	0.80**
	(0.55)	(0.40)	(0.58)	(0.41)
Racial similarity	-0.61	-0.68	-1.17	-0.82
	(1.08)	(0.86)	(1.10)	(0.90)
Education of founding team	0.02	0.04	0.04	0.04
	(0.05)	(0.04)	(0.05)	(0.04)
Has intellectural property	0.25	-0.09	0.19	-0.15
	(0.23)	(0.17)	(0.24)	(0.18)
Sole Propietorship	-0.36	-0.09	-0.28	-0.25
	(0.23)	(0.16)	(0.25)	(0.17)
CBSA controls				
MSA location (versus MISA location)	0.20	0.12	0.20	0.08
	(0.33)	(0.25)	(0.35)	(0.25)
Number of patents per capita (X1,000)	0.13	-0.12	0.19	-0.15
	(0.24)	(0.21)	(0.25)	(0.22)
Poverty rate	8.16***	2.84	8.20***	3.20
	(2.98)	(2.37)	(3.18)	(2.47)
Racial similarity (CBSA)	-0.47	0.65	-0.45	0.75
	(0.76)	(0.53)	(0.79)	(0.54)
Education levels (number of research universities per capita) (X100,000)	0.01	-0.04	0.01	-0.03
	(0.04)	(0.04)	(0.04)	(0.04)
Impact of profitability on satisfaction ratings				
Risk-adjusted profitability	0.00	-0.22***	0.22***	0.07
	(0.01)	(0.06)	(0.07)	(0.10)
Cluster location quotient	-0.00	-0.12	0.31**	0.18
	(0.06)	(0.07)	(0.13)	(0.14)
Risk-adjusted profitability*Cluster location quotient			**-0.17*** **	**-0.27*** **
			(0.06)	(0.10)
Constant	-1.60	-0.75	-2.60**	-0.80
	(1.26)	(1.03)	(1.30)	(1.09)
Startups	1,692	1,692	1,557	1,557

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In models 1 and 2 I look at the main effects of both startup profitability and geographic clustering on subjective self-rated levels of satisfaction. I find that while being more profitable lowers the likelihood of a negative satisfaction rating (b=-0.22, p<0.01 in model 2), the other main effects are not significant. However, when I look at the interaction effects in models 3 and 4, I find that the degree of clustering negatively moderates the impact of profitability on both positive (b=-0.17, p<0.01 in model 3) and negative (b=-0.27, p<0.01) self-ratings of satisfaction. I plot these dynamics in figure 6 below.

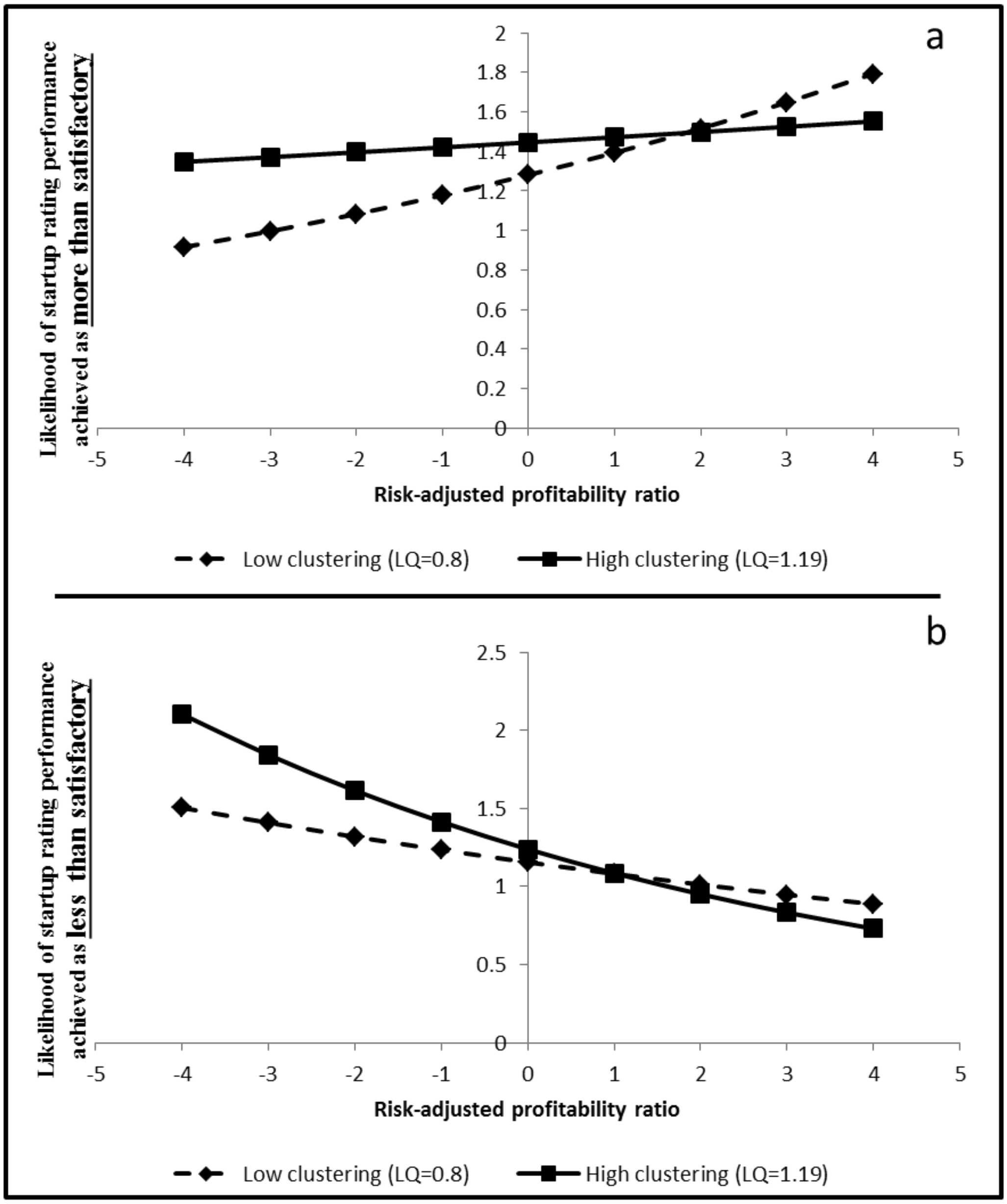

Figure 6 (paper 3). Multinomial logit interaction plot: The impact of risk-adjusted profitability on startups self-evaluations of performance over the first four years of KFS as (a) more and (b) less than satisfactory. Clustering *lowers* both positive and negative self-evaluations.

As is observed in both panels A and B, an increase in risk-adjusted profitability increases the likelihood that startups will be more than satisfied and lowers the likelihood that startups will be less than satisfied with the performance level that they have achieved. However, the degree of clustering that startups are exposed to dampens these subjective ratings. In the gain domain (i.e. firms that have positive profitability), startups that are exposed to higher levels of clustering are *less likely to be subjectively satisfied* for comparable levels of performance. Similarly, in the loss domain (i.e. firms that have negative profitability), startups that are exposed to higher levels of clustering are *more likely to be subjectively dissatisfied* for comparable levels of performance. In sum, I find strong evidence that startups exposed to higher levels of clustering are likely to be subjectively *"unhappier"* for a given objective level of performance. This finding is again consistent with startups experiencing a performance premium within geographic clusters.

DISCUSSION

While geographic clustering is typically associated with positive outcomes for startup firms such as innovation and growth, a number of studies have shown that it can also lead to higher rates of startup failure (Folta *et al.*, 2006a; McCann and Folta, 2008; Sorenson and Audia, 2000) . In this study, I examine the underlying behavioral mechanisms that lead to such failure, exploring the validity of two alternate explanations, *adverse selection effects* and *performance premium effects.* While I do not find support for adverse selection into clusters, I do find consistent evidence for performance premium effects. In aggregate these findings suggest that startup failure in clusters is best explained by a combination of both increased competition for resources (i.e. cluster-based agglomerations diseconomies) as well as more voluntary exits triggered by the failure to achieve comparatively (i.e. relative to isolated locations) higher expectations of performance.

The lack of support for adverse selection effects is particularly interesting, given that they have been previously posited as a viable explanation for the higher observed startup failure rates within clusters (Kalnins and Chung, 2004; McCann and Folta, 2008). To the contrary, the evidence suggests that there might not be any selective sorting when comparing clusters to isolated locations , and positive selection effects when considering entry into very large clusters (defined as clusters with a LQ>1.46, that is in the 90^{th} percentile or higher of the study sample). These findings also lend support to theories of entrepreneurial location choice that argue against adverse selection effects. For instance, some scholars have suggested that adverse selection explanations overly simplify the balance between positive and negative spillover effects in clusters. Instead they have argued that the stronger capabilities and higher absorptive capacity of high quality firms might actually allow them to benefit asymmetrically from clusters relative to any costs that they might incur through co-location (Almeida and Kogut, 1997; Baldwin and Okubo, 2006; McCann and Folta, 2011). Furthermore, other studies that have examined the process by which entrepreneurs make location decisions also suggest that they do not systematically pay attention to agglomeration externalities when deciding where to locate; instead entrepreneurs often start firms in places where they already live or work thus endogenously forming geographic clusters (Agarwal and Braguinsky, 2014; Dahl and Sorenson, 2012; Klepper, 2010; Sorenson and Audia, 2000). Thus, while adverse selection might explain the location choices of more established entrants from distant locales (e.g. Shaver & Flyer looked at the location choices of foreign firms choosing strategically among various U.S. locations), these theories indicate they are also potentially less applicable in the startup context (Folta *et al.*, 2006a). These contrasting theoretical predictions and my empirical findings also lead to some interesting possibilities for future research. For instance, it might be that the particular kind of selection effect observed (i.e. adverse vs. self-selection) varies as a function of different entrant types (e.g. de novo, spinoff, de alio). That is, while spinoff firms and de novo firms might contribute to positive selection effects in clusters, established de alio firms might avoid them and contribute to adverse selection. Future studies with more fine-grained data on the origins of startups might be able to tackle this interesting research question.

The most interesting set of findings from this study is that startups within clusters require a performance premium to persist with operations. For instance, survival analyses indicated that they are more likely to close operations for comparable levels of economic performance. They are also likely to have higher exit thresholds, and less likely to be subjectively happy with comparable levels of financial performance relative to more isolated startups. This finding means that while cluster-based diseconomies might undoubtedly contribute to higher rates of startup failure, they only offer a partial explanation. In particular, from a behavioral perspective, high levels of geographic clustering might also contribute to hyper-competitive dynamics where entrepreneurs within clusters form unrealistic explanations, and implement unwise strategies such as pursuing excessive growth and engaging in head-on competition with dominant incumbents (Churchill and Mullins, 2001; Fan, 2010; Markman and Gartner, 2002; Pierce and Aguinis, 2013). Yet, on the positive side, one could also potentially argue that the quicker failure induced by such behavior within clusters can also lead to faster and more effective learning (Blank, 2013; Cope, 2011; Ries, 2011; Ucbasaran *et al.*, 2013), and is less psychologically damaging (Shepherd, 2003) to entrepreneurs in the long run. It is therefore also theoretically possible that the marginal impacts of clustering on performance premium effects vary as a function of both prior entrepreneurial experiences. While examining these more complicated dynamics was outside the scope of this study, these are interesting questions that could be answered by future research with more granular detail on both the volume and nature of prior entrepreneurial experiences (Toft-Kehler, Wennberg, and Kim, 2014). Beyond these opportunities for future research, these findings also raise some interesting policy implications.

For instance, from a societal standpoint, some scholars have suggested that firm failure can be potentially beneficial as it frees up resources that might otherwise be tied up in inefficient or incompetent enterprises (Knott and Posen, 2005; Pe'er and Vertinsky, 2008; Shane, 2009). To the extent that startups within clusters are more likely to voluntarily exit, particularly if they are performing poorly, these findings suggest that cluster-based policies to encourage entrepreneurship (Gilbert *et al.*, 2004; Rocha, 2004; Rocha and Sternberg, 2005) might make the entrepreneurial process less forgiving but also potentially more efficient in the long-term.

This study also some potential limitations. First, I am not able to ascertain the true underlying motives behind exit decisions. One could legitimately argue that a sale or merger is not necessarily reflective of a positive exit and closure a negative exit event. For instance, firms might undergo distress sales (Wennberg *et al.*, 2010), be acquired by other firms despite poor performance (Bates, 2005) , or shut down due to personal, non-financial reasons (DeTienne and Chirico, 2013). While I cannot account for this possibility in my data, the sensitivity analyses in table 11 (models 6 and 7) suggest that using closures as measure of failure is a good approach. Second, the survey-based nature of the KFS also means that the data is self-reported. Entrepreneurs might have an incentive to conceal or inflate reported metrics, particularly with respect to financial performance. While the KFS data is carefully checked for such inconsistencies at the NORC prior to being released for analyses, I also accounted for such factors through relevant statistical procedures (e.g. logically imputing missing data, excluding variables with large amounts of missing data). And lastly, all firms in the sample are U.S. based, thus limiting the ability to generalize findings to other national contexts. Yet given the wide range of industries and large number of geographies examined, it is unlikely that these findings are idiosyncratic to the U.S.

Concluding remarks

In this study, I investigate the underlying mechanisms behind startup failure in clusters. In particular, I examine whether the higher observed failure rates within clusters can be explained by a combination of adverse selection and performance premium effects. While I do not find evidence of adverse selection effects, I do find that geographic clustering leads to performance premium effects. These findings are important as they demonstrate that *post-entry* behavioral dynamics can play an important role, both independent from and in combination with the agglomeration diseconomies that scholars have focused on in studying startup failure within clusters.

SUMMARY OF FINDINGS AND CONTRIBUTION TO THE LITERATURE

The purpose of this dissertation was to study the strategic implications of geographic factors on startup firms. In particular, I suggested that the existing literature on geography and entrepreneurship suffers from two major gaps. First, the literature is largely focused on explaining *regional level* variability in entrepreneurial activity (Buhr and Owen-Smith, 2010). This topic is undoubtedly deserving of the attention it has received, given the very real impacts of entrepreneurial activity on regional economic growth and prosperity (Audretsch *et al.*, 2006; Wennekers and Thurik, 1999). And yet, within this stream, scholars have largely operated in siloes within disciplinary boundaries. For instance, economic geographers have traditionally focused on factors such as clusters (Delgado *et al.*, 2010), sociologists on the role of institutions (Tolbert *et al.*, 2011), and strategy scholars on endogenous industry evolution dynamics (Agarwal *et al.*, 2007; Agarwal and Braguinsky, 2014; Klepper, 2010). As a result, there has been very little integration and research across disciplinary boundaries. Second, the emphasis on understanding the causes and consequences of regional variability in entrepreneurship has also meant that we know far less about the *implications* of location choice for startup firms. In fact, some studies suggest that entrepreneurial location choices are devoid of any complex strategy at all, with most individuals just starting firms in the places where they already live or work (Dahl and Sorenson, 2012; Nanda and Sørensen, 2010; Stam, 2007). But what are the consequences of these decisions? That is, how does location choice impact startup performance and startup behavior post-entry?

The papers in this dissertation therefore attempt to address both these gaps. Paper 1 focuses on developing a more integrated theory of entrepreneurial entry at the regional level, combining perspectives from industrial economics and sociological institutional theory. Paper 2 and 3 instead focus on firm-level issues. In particular, paper 2 uses a variance decomposition approach to examine the *relative extent* (i.e. in comparison to other factors such as firm capabilities and industry affiliation) to which geographic factors have a marked influence on startup performance. And paper 3 looks at the impacts of a specific regional factor (geographic clustering) on startup failure, with an emphasis on how geographic clustering levels can impact startup decision making processes related to voluntary exit decisions (DeTienne *et al.*, 2008; Gimeno *et al.*, 1997).

Theoretical insights

The key theoretical implication from the first paper is that regional entrepreneurial entry is not solely a function of endogenous industry dynamics such as localized knowledge spillovers emanating from industry incumbents (Agarwal *et al.*, 2007; Agarwal and Braguinsky, 2014; Klepper, 2010). In particular, I argue that the sociopolitical ideologies held by communities in a region can help potential entrants both recognize and be more motivated to orchestrate the resources that they need to pursue entrepreneurial opportunities. Furthermore, while prior sociological studies have largely focused on the *direct* impacts that sociocultural institutions can have on entrepreneurial action (Tolbert *et al.*, 2011), I show that such institutions can also have significant *indirect* impacts. Namely, they are able to shape the evolutionary dynamics within emerging industries, by conditioning the symbiotic relationship (Baumol, 2002; Hockerts and Wüstenhagen, 2010) between incumbents and startup firms. Furthermore, my finding that the influence of ideologies wanes as regions become more specialized in cleantech is also theoretically important as it highlights the dynamic role of institutional factors (Nelson, 1994) . In particular my findings suggest that while sociocultural factors such as ideologies are important initially, as regions become more focused and specialized within particular technology trajectories (Martin and Sunley, 2006) the endogenous industry dynamics (Agarwal *et al.*, 2007) emphasized by strategy scholars are more likely to be more salient.

Since paper 2 uses a variance decomposition approach, it does not have the ability to isolate the myriad of underlying theoretical factors (e.g. knowledge spillovers, labor market pools, supplier linkages, regional inventor networks, sociocultural institutions) that could potentially contribute to performance advantages for startups. However the findings do suggest that, at least for venture-backed startups, the locus of competitive advantage lies within firm boundaries consistent with the resource-based view of strategy (Barney, 1991). The larger impacts of regions seem to be instead limited to less developed startups and those operating in emerging industry sectors. Thus, the key theoretical insight from the second paper is to suggest that, in general, the way that regions can lead to strategic advantages for startups is by providing them with a platform that allows important resources at founding, such as human capital, to be more easily accessed and internalized (Packalen, 2007; Wright *et al.*, 2007). To the degree that certain urban environments are more conducive to this process of resource orchestration (Mendes *et al.*, 2010; Shimer and Smith, 2000; Sirmon *et al.*, 2011), regions therefore still play an important but largely supporting role in terms of enabling firms to compete effectively. This interpretation of regions thus complementing but not usurping firm-specific capabilities is also consistent with some recent empirical evidence in both entrepreneurship (Gilbert *et al.*, 2008) and economic geography (Huber, 2012) that suggests that the importance of common-pool resources such as knowledge spillovers might be somewhat overstated in terms of explaining firm-level performance advantages; instead the ability of firms to take advantage of available regional endowments (McCann and Folta, 2011; Pe'er and Keil, 2013) is likely to be far more important.

In paper 3, I focus more specifically on the role of one specific regional factor, geographic clustering, studying how it impacts startup failure. The most interesting findings from this study are that geographic clustering can lead to performance premium effects; startups within clusters are more likely to close operations for a given level of performance, have higher exit thresholds, and are subjectively less satisfied with a given objective level of performance. Thus consistent with other work emphasizing the importance of behavioral dynamics on startup exit decisions (DeTienne *et al.*, 2008; Gimeno *et al.*, 1997), the findings from this study show that existing theories of startup failure within clusters solely based on theories of resource-competition within markets (Baum and Mezias, 1992; Sorenson and Audia, 2000) might be incomplete in that they ignore the voluntary dimension of exit. The key theoretical contribution of this study is therefore to highlight and demonstrate that regional factors, such as clusters, can have important *behavioral impacts* on startups specifically, and on firms more broadly (Porter, 2000). It also answers calls for research on how locational factors shape managerial behavior and decision making processes within firms (Marquis and Battilana, 2009). This is an area that has received surprisingly little scholarly interest, despite the fact that firms do have the agency to differentially respond to common contextual pressures (Oliver, 1991).

Insights for policy makers

Through this research I am also able to offer a number of policy insights. For instance, the results from the first paper indicate that policy makers interested in regional economic development through entrepreneurship can be more strategic with respect to how they choose to incentivize and support entrepreneurs. The existing evidence suggests that public policy efforts to boost regional rates of entrepreneurship have largely failed (Lerner, 2009, 2010). Yet, most regional programs have tried to stimulate entrepreneurship through the use of financial incentives (Shane, 2009). However such a policy approach that ignores the local sociocultural environment might inadvertently benefit more established companies to the exclusion of nascent startups (York and Lenox, 2014). Instead an emphasis on shaping community attitudes towards specific technology sectors, both directly (Walker *et al.*, 2010) as well as through third party actors such as social movement organizations (Pacheco *et al.*, 2014) might be more fruitful, particularly at the nascent stages of technological development within a region.

My findings from the second paper also have interesting policy implications. The fact that, on average, startup performance differences are largely driven by firm-specific effects means that policy makers should focus on actions that can directly *improve the quality* of startup firms. This could for instance be achieved through an increased emphasis on entrepreneurial training (Drucker, 2014; Katz, 2003, 2008; Kuratko, 2005; Neck, Greene, and Brush, 2014), which could translate to founding teams with stronger entrepreneurship-specific human capital (Erikson, 2002). In contrast, using policy instruments that simply encourage more people within regions to become entrepreneurs, irrespective of their capabilities, are less likely to be successful in terms of effectively improving startup performance levels (Shane, 2009). Along similar lines, my finding of relatively small industry effects also indicates that policy makers should avoid focusing their support too narrowly on a particular set of industry sectors and technology domains (Mason and Brown, 2013b). And lastly, my split-sample analyses (i.e. by startup development stage and industry maturity level) also give guidance on the kinds of firms that are most likely to be impacted by changes in the regional policy regime. For instance, a shift in regional labor market policy that modifies the accessibility of regional common-pool resources to firms, such as a change in the mobility of skilled workers through a differential enforcement of non-compete agreements (Garmaise, 2011; Marx, 2011), is most likely to impact less developed startups and those operating within nascent industry sectors.

The findings from paper 3 are also interesting in terms of their policy implications. For instance, they suggest that cluster-based policies to encourage entrepreneurship (Rocha, 2004; Rocha and Sternberg, 2005) might indeed motivate entrepreneurs to aspire to higher levels of performance as proponents of clusters have argued for (Porter, 2000). However the results suggest that high levels of clustering can also lead entrepreneurs to be less satisfied and aspire to possibly unrealistic levels of performance (Ordóñez *et al.*, 2009). If these dynamics lead startups to pursue potentially unwise strategies such as excessive growth (Churchill and Mullins, 2001; Markman and Gartner, 2002) or head-on competition with incumbents (Fan, 2010), an excessive emphasis on cluster-based strategies might not be ideal and lead to excessive failure among startup firms. However, to the extent that the failure induced by such behavior also leads to subsequent learning (Toft-Kehler *et al.*, 2014) and is less psychologically damaging to entrepreneurs (Shepherd, 2003), clusters might also make the entrepreneurial process more efficient at the societal level in the long run (Knott and Posen, 2005).

Potential avenues for future research

Some of the findings from this dissertation also lay the foundation for future work. For instance, while paper 1 focuses on how exogenous sociological institutions such as ideologies can shape the cooperative relationship between incumbents and startup firms, it stands to reason that such institutions should also impact their competitive dynamics post-entry. Thus, it would be interesting to examine how sociocultural institutions impact the ability of startups entering these emerging industries to compete effectively with incumbents. Similarly, since the focus of this study is on how innovation rates impact entrepreneurship rates at the aggregate regional level, I did not focus on how specific types of innovations (e.g. radical vs. incremental) are brought to market. For instance, an implicit underlying assumption of the knowledge spillover based theory of entrepreneurship (Agarwal *et al.*, 2007) is that startup firms are most likely to commercialize the most radical innovations (Acs *et al.*, 2009; Christensen and Rosenbloom, 1995; Tripsas and Gavetti, 2000). Yet others have challenged this view arguing that incumbent firms can and do overcome disadvantages that they face in commercializing radical innovations (Hill and Rothaermel, 2003), for instance by cooperating with new entrants (Rothaermel, 2001). To the extent that exogenous social forces shape competitive and cooperative dynamics between entrants and incumbents, this context offers an opportunity to extend our understanding of innovation dynamics within emerging industries. Similarly with respect to paper 3, the dearth of research on the behavioral impacts of locational factors on startup firms means that this is a particularly fruitful area for future research. For instance, an interesting implication of the peer-pressure effects of clusters is that social comparison mechanisms should potentially help counteract agency costs (Porter, 2000), such as shirking behavior. While this is less likely to be problematic in startups where owner-operators own much of the firm and are thus both financially and emotionally (Cardon *et al.*, 2005; DeTienne and Chirico, 2013) invested in the enterprise, it is likely to be quite important for firms with a larger share of outsider equity or for startups that rely on outside contracts as a source of financing (e.g. government grants with little scope for direct enforcement). Thus future work could examine both the occurrence of, as well as the degree to which geographic clustering influences, these other kinds of managerial behaviors within startup firms.

Concluding remarks

Geography is important to study given that entrepreneurship is a spatially heterogeneous phenomenon. Furthermore, choosing where to locate is a fundamental decision that all startups need to make. And yet, the intricacies and strategic implications of these choices are poorly understood, despite their significant impacts on both startups as well as the broader economy. It is therefore my hope that the papers in this dissertation can form an initial step in improving our understanding of this important area of research.

BIBLIOGRAPHY

Abrahamson E. 1996. Management fashion. *Academy of Management Review* **21**: 254–285.

Acs ZJ, Braunerhjelm P, Audretsch DB, Carlsson B. 2009. The knowledge spillover theory of entrepreneurship. *Small Business Economics* **32**(1): 15–30.

Acs ZJ, Plummer LA. 2005. Penetrating the'knowledge filter'in regional economies. *The Annals of Regional Science* **39**(3): 439–456.

Acs ZJ, Sanders M. 2012. Patents, knowledge spillovers, and entrepreneurship. *Small Business Economics* **39**(4): 801–817.

Acs ZJ, Szerb L. 2007. Entrepreneurship, economic growth and public policy. *Small Business Economics* **28**(2-3): 109–122.

Agarwal R, Audretsch D, Sarkar MB. 2007. The process of creative construction: knowledge spillovers, entrepreneurship, and economic growth. *Strategic Entrepreneurship Journal* **1**(3-4): 263–286.

Agarwal R, Braguinsky S. 2014. Industry evolution and entrepreneurship: Steven Klepper's contributions to industrial organization, strategy, technological change and entrepreneurship. *Strategic Entrepreneurship Journal (forthcoming).*

Agarwal R, Echambadi R, Franco AM, Sarkar MB. 2004. Knowledge transfer through inheritance: Spin-out generation, development, and survival. *The Academy of Management Journal* **47**(4): 501–522.

Aharonson BS, Baum JA, Feldman MP. 2007. Desperately seeking spillovers? Increasing returns, industrial organization and the location of new entrants in geographic and technological space. *Industrial and Corporate Change* **16**(1): 89–130.

Alcácer J, Chung W. 2007. Location strategies and knowledge spillovers. *Management Science* **53**(5): 760–776.

Alcácer J, Chung W. 2014. Location strategies for agglomeration economies. *Strategic Management Journal* **35**(12): 1749–1761.

Aldrich HE, Fiol CM. 1994. Fools rush in? The institutional context of industry creation. *Academy of Management Review* **19**(4): 645–670.

Aldrich H, Stern RN. 1983. Resource Mobilization and the Creation of US Producer's Cooperatives, 1835-1935. *Economic and Industrial Democracy* **4**(3): 371–406.

AlLaham A, Souitaris V. 2008. Network embeddedness and new-venture internationalization: Analyzing international linkages in the German biotech industry. *Journal of Business Venturing* **23**(5): 567–586.

Allison PD. 2009. *Fixed effects regression models*. SAGE publications: Thousand Oaks, CA.

Allison PD, Waterman RP. 2002. Fixed–effects negative binomial regression models. *Sociological Methodology* **32**(1): 247–265.

Almeida P, Dokko G, Rosenkopf L. 2003. Startup size and the mechanisms of external learning: increasing opportunity and decreasing ability? *Research Policy* **32**(2): 301–315.

Almeida P, Kogut B. 1997. The exploration of technological diversity and geographic localization in innovation: start-up firms in the semiconductor industry. *Small Business Economics* **9**(1): 21–31.

Almeida P, Kogut B. 1999. Localization of knowledge and the mobility of engineers in regional networks. *Management Science* **45**(7): 905–917.

Almeida P, Song J, Grant RM. 2002. Are firms superior to alliances and markets? An empirical test of cross-border knowledge building. *Organization Science* **13**(2): 147–161.

Alvarez SA, Barney JB. 2007. Discovery and creation: Alternative theories of entrepreneurial action. *Strategic Entrepreneurship Journal* **1**(1-2): 11–26.

Amin A, Cohendet P. 2005. Geographies of knowledge formation in firms. *Industry and Innovation* **12**(4): 465–486.

Appold S. 2005. Location patterns of US industrial research: Mimetic isomorphism and the emergence of geographic charisma. *Regional Studies* **39**(1): 17–39.

Arnott R. 2007. Congestion tolling with agglomeration externalities. *Journal of Urban Economics* **62**(2): 187–203.

Arora A, Gambardella A. 1994. The changing technology of technological change: general and abstract knowledge and the division of innovative labour. *Research Policy* **23**(5): 523–532.

Arrow K. 1962. Economic Welfare and the Allocation of Resources for Invention. In *The Rate and Direction of Inventive Activity: Economic and Social Factors*. Princeton University Press: Princeton, NJ: 609–626.

Asheim BT, Boschma R, Cooke P. 2011. Constructing regional advantage: platform policies based on related variety and differentiated knowledge bases. *Regional Studies* **45**(7): 893–904.

Audia PG, Rider CI. 2010. Close, but not the same: locally headquartered organizations and agglomeration economies in a declining industry. *Research Policy* **39**(3): 360–374.

Audretsch DB. 2007. Entrepreneurship capital and economic growth. *Oxford Review of Economic Policy* **23**(1): 63–78.

Audretsch DB, Bönte W, Keilbach M. 2008. Entrepreneurship capital and its impact on knowledge diffusion and economic performance. *Journal of Business Venturing* **23**(6): 687–698.

Audretsch DB, Feldman MP. 1996. R&D spillovers and the geography of innovation and production. *The American Economic Review* **86**(3): 630–640.

Audretsch DB, Keilbach M. 2004a. Does entrepreneurship capital matter? *Entrepreneurship Theory and Practice* **28**(5): 419–429.

Audretsch DB, Keilbach M. 2004b. Entrepreneurship capital and economic performance. *Regional Studies* **38**(8): 949–959.

Audretsch DB, Keilbach M. 2005. Entrepreneurship capital and regional growth. *The Annals of Regional Science* **39**(3): 457–469.

Audretsch DB, Keilbach M. 2007. The Theory of Knowledge Spillover Entrepreneurship. *Journal of Management Studies* **44**(7): 1242–1254.

Audretsch DB, Keilbach M. 2008. Resolving the knowledge paradox: Knowledge-spillover entrepreneurship and economic growth. *Research Policy* **37**(10): 1697–1705.

Audretsch DB, Keilbach MC, Lehmann EE. 2006. *Entrepreneurship and economic growth*. Oxford University Press: Oxford, United Kingdom.

Audretsch DB, Lehmann EE, Warning S. 2005. University spillovers and new firm location. *Research Policy* **34**(7): 1113–1122.

Audretsch DB, Thurik AR. 2001. What's new about the new economy? Sources of growth in the managed and entrepreneurial economies. *Industrial and Corporate Change* **10**(1): 267–315.

Baldwin RE, Okubo T. 2006. Heterogeneous firms, agglomeration and economic geography: spatial selection and sorting. *Journal of Economic Geography* **6**(3): 323–346.

Barley SR, Kunda G. 1992. Design and devotion: Surges of rational and normative ideologies of control in managerial discourse. *Administrative Science Quarterly* **37**(3): 363–399.

Barnett WP, Miner AS. 1992. Standing on the shoulders of others: Career interdependence in job mobility. *Administrative Science Quarterly* **37**(2): 262–281.

Barnett WP, Pontikes EG. 2004. The Red Queen: History-dependent competition among organizations. *Research in Organizational Behavior* **26**: 351–371.

Barney J. 1991. Firm resources and sustained competitive advantage. *Journal of Management* **17**(1): 99–120.

Baron DP. 1995. Integrated strategy: market and nonmarket components. *California Management Review* **37**(2).

Baron RA. 2004. The cognitive perspective: a valuable tool for answering entrepreneurship's basic 'why' questions. *Journal of Business Venturing* **19**(2): 221–239.

Baron RA. 2006. Opportunity recognition as pattern recognition: How entrepreneurs 'connect the dots' to identify new business opportunities. *The Academy of Management Perspectives* **20**(1): 104–119.

Baron RA, Ensley MD. 2006. Opportunity recognition as the detection of meaningful patterns: Evidence from comparisons of novice and experienced entrepreneurs. *Management Science* **52**(9): 1331–1344.

Bates T. 2005. Analysis of young, small firms that have closed: delineating successful from unsuccessful closures. *Journal of Business Venturing* **20**(3): 343–358.

Baum JA, Haveman HA. 1997. Love thy neighbor? Differentiation and agglomeration in the Manhattan hotel industry, 1898-1990. *Administrative Science Quarterly* **42**(2): 304–338.

Baum JA, Mezias SJ. 1992. Localized competition and organizational failure in the Manhattan hotel industry, 1898-1990. *Administrative Science Quarterly* **37**(4): 580–604.

Baum JA, Silverman BS. 2004. Picking winners or building them? Alliance, intellectual, and human capital as selection criteria in venture financing and performance of biotechnology startups. *Journal of Business Venturing* **19**(3): 411–436.

Baumol WJ. 1996. Entrepreneurship: Productive, unproductive, and destructive. *Journal of Business Venturing* **11**(1): 3–22.

Baumol WJ. 2002. Entrepreneurship, innovation and growth: The David-Goliath symbiosis. *Journal of Entrepreneurial Finance* **7**(2): 1–10.

Baumol WJ, Litan RE, Schramm CJ. 2007. *Good Capitalism, Bad Capitalism, and the Economics of Growth and Prosperity*. Yale University Press: New Haven, CT.

Beckman CM. 2006. The influence of founding team company affiliations on firm behavior. *Academy of Management Journal* **49**(4): 741–758.

Bell GG. 2005. Clusters, networks, and firm innovativeness. *Strategic Management Journal* **26**(3): 287–295.

Benner MJ, Tripsas M. 2012. The influence of prior industry affiliation on framing in nascent industries: the evolution of digital cameras. *Strategic Management Journal* **33**(3): 277–302.

Berchicci L, King A, Tucci CL. 2011. Does the apple always fall close to the tree? The geographical proximity choice of spin-outs. *Strategic Entrepreneurship Journal* **5**(2): 120–136.

Bertoni F, Colombo MG, Grilli L. 2011. Venture capital financing and the growth of high-tech start-ups: Disentangling treatment from selection effects. *Research Policy* **40**(7): 1028–1043.

Bitektine A. 2011. Toward a theory of social judgments of organizations: The case of legitimacy, reputation, and status. *Academy of Management Review* **36**(1): 151–179.

Blank S. 2013. Why the lean start-up changes everything. *Harvard Business Review* **91**(5): 63–72.

Boschma R, Iammarino S. 2009. Related variety, trade linkages, and regional growth in Italy. *Economic Geography* **85**(3): 289–311.

Bradley SW, Shepherd DA, Wiklund J. 2011. The importance of slack for new organizations facing ‘tough’environments. *Journal of Management Studies* **48**(5): 1071–1097.

Brander JA, Amit R, Antweiler W. 2002. Venture-Capital Syndication: Improved Venture Selection vs. The Value-Added Hypothesis. *Journal of Economics & Management Strategy* **11**(3): 423–452.

Breschi S, Lissoni F. 2001a. Localised knowledge spillovers vs. innovative milieux: Knowledge ‘tacitness’ reconsidered. *Papers in Regional Science* **80**(3): 255–273.

Breschi S, Lissoni F. 2001b. Knowledge spillovers and local innovation systems: a critical survey. *Industrial and Corporate Change* **10**(4): 975–1005.

Bresnahan T, Gambardella A, Saxenian A. 2001. ‘Old economy’ inputs for ‘new economy’ outcomes: cluster formation in the new Silicon Valleys. *Industrial and Corporate Change* **10**(4): 835–860.

Brinckmann J, Grichnik D, Kapsa D. 2010. Should entrepreneurs plan or just storm the castle? A meta-analysis on contextual factors impacting the business planning–performance relationship in small firms. *Journal of Business Venturing* **25**(1): 24–40.

Brockner J, Higgins ET, Low MB. 2004. Regulatory focus theory and the entrepreneurial process. *Journal of Business Venturing* **19**(2): 203–220.

Buenstorf G, Klepper S. 2009. Heritage and Agglomeration: The Akron Tyre Cluster Revisited. *The Economic Journal* **119**(537): 705–733.

Buhr H, Owen-Smith J. 2010. Networks as institutional support: Law firm and venture capitalist relations and regional diversity in high-technology IPOs. *Research in the Sociology of Work* **21**: 95–126.

Busenitz LW, Barney JB. 1997. Differences between entrepreneurs and managers in large organizations: Biases and heuristics in strategic decision-making. *Journal of Business Venturing* **12**(1): 9–30.

Buttel FM, Flinn WL. 1978. The Politics of Environmental Concern The Impacts of Party Identification and Political Ideology on Environmental Attitudes. *Environment and Behavior* **10**(1): 17–36.

Camerer C, Lovallo D. 1999. Overconfidence and excess entry: An experimental approach. *The American Economic Review* **89**(1): 306–318.

Cameron AC, Trivedi PK. 2013. *Regression analysis of count data*. Cambdrige University Press: Cambridge, United Kingdom, 53.

Cantwell J, Santangelo GD. 2003. The new geography of corporate research in information and communications technology (ICT). In *Change, Transformation and Development*. Springer: Berlin, Germany: 343–377.

Capello R. 1999. Spatial transfer of knowledge in high technology milieux: learning versus collective learning processes. *Regional Studies* **33**(4): 353–365.

Cardon MS, Zietsma C, Saparito P, Matherne BP, Davis C. 2005. A tale of passion: New insights into entrepreneurship from a parenthood metaphor. *Journal of Business Venturing* **20**(1): 23–45.

Cassar G. 2006. Entrepreneur opportunity costs and intended venture growth. *Journal of Business Venturing* **21**(5): 610–632.

Cassar G. 2014. Industry and startup experience on entrepreneur forecast performance in new firms. *Journal of Business Venturing* **29**(1): 137–151.

Cassiman B, Ueda M. 2006. Optimal project rejection and new firm start-ups. *Management Science* **52**(2): 262–275.

Castellaneta F, Gottschalg O. 2014. Does ownership matter in private equity? The sources of variance in buyouts' performance. *Strategic Management Journal (forthcoming).*

Castrogiovanni GJ. 1996. Pre-startup planning and the survival of new small businesses: Theoretical linkages. *Journal of Management* **22**(6): 801–822.

Chan CM, Makino S, Isobe T. 2010. Does subnational region matter? Foreign affiliate performance in the United States and China. *Strategic Management Journal* **31**(11): 1226–1243.

Chatterji A, Glaeser E, Kerr W. 2014. Clusters of entrepreneurship and innovation. *Innovation Policy and the Economy* **14**(1): 129–166.

Chen H, Gompers P, Kovner A, Lerner J. 2010. Buy local? The geography of venture capital. *Journal of Urban Economics* **67**: 90–102.

Christensen CM, Rosenbloom RS. 1995. Explaining the attacker's advantage: Technological paradigms, organizational dynamics, and the value network. *Research Policy* **24**(2): 233–257.

Chung W, Alcácer J. 2002. Knowledge seeking and location choice of foreign direct investment in the United States. *Management Science* **48**(12): 1534–1554.

Churchill NC, Mullins JW. 2001. How fast can your company afford to grow. *Harvard Business Review* **79**(5): 135–142.

Chwolka A, Raith MG. 2012. The value of business planning before start-up—A decision-theoretical perspective. *Journal of Business Venturing* **27**(3): 385–399.

De Clercq D, Fried VH, Lehtonen O, Sapienza HJ. 2006. An Entrepreneur's Guide to the Venture Capital Galaxy. *The Academy of Management Perspectives* **20**: 90–112.

Cohen B, Winn MI. 2007. Market imperfections, opportunity and sustainable entrepreneurship. *Journal of Business Venturing* **22**(1): 29–49.

Coleman S, Cotei C, Farhat J. 2013. A Resource-Based View Of New Firm Survival: New Perspectives On The Role Of Industry And Exit Route. *Journal of Developmental Entrepreneurship* **18**(1).

Coleman S, Robb A. 2009. A comparison of new firm financing by gender: evidence from the Kauffman Firm Survey data. *Small Business Economics* **33**(4): 397–411.

Colombo MG, Grilli L. 2010. On growth drivers of high-tech start-ups: Exploring the role of founders' human capital and venture capital. *Journal of Business Venturing* **25**(6): 610–626.

Combes P-P, Duranton G. 2006. Labour pooling, labour poaching, and spatial clustering. *Regional Science and Urban Economics* **36**(1): 1–28.

Conti R. 2014. Do non-competition agreements lead firms to pursue risky R&D projects? *Strategic Management Journal* **35**(8): 1230–1248.

Cooke P. 2008. Regional innovation systems, clean technology & jacobian cluster-platform policies. *Regional Science Policy & Practice* **1**(1): 23–45.

Coombs JE, Deeds DL, Ireland RD. 2009. Placing the choice between exploration and exploitation in context: a study of geography and new product development. *Strategic Entrepreneurship Journal* **3**(3): 261–279.

Coombs JE, Mudambi R, Deeds DL. 2006. An examination of the investments in US biotechnology firms by foreign and domestic corporate partners. *Journal of Business Venturing* **21**(4): 405–428.

Cooper A, Folta T. 2000. Entrepreneurship and high-technology clusters. In *The Blackwell Handbook of Entrepreneurship*. John Wiley & Sons: Hoboken, NJ: 348–367.

Cope J. 2011. Entrepreneurial learning from failure: An interpretative phenomenological analysis. *Journal of Business Venturing* **26**(6): 604–623.

Cowan R, David PA, Foray D. 2000. The explicit economics of knowledge codification and tacitness. *Industrial and Corporate Change* **9**(2): 211–253.

Cumming D, Dai N. 2010. Local bias in venture capital investments. *Journal of Empirical Finance* **17**: 362–380.

Czarnitzki D, Rammer C, Toole AA. 2014. University spin-offs and the ‘performance premium’. *Small Business Economics* **43**(2): 309–326.

Dahl MS, Sorenson O. 2012. Home sweet home: Entrepreneurs’ location choices and the performance of their ventures. *Management Science* **58**(6): 1059–1071.

Davidsson P, Honig B. 2003. The role of social and human capital among nascent entrepreneurs. *Journal of Business Venturing* **18**(3): 301–331.

Davies AR. 2013. Cleantech clusters: transformational assemblages for a just, green economy or just business as usual? *Global Environmental Change* **23**(5): 1285–1295.

Davila A, Foster G. 2007. Management control systems in early-stage startup companies. *The Accounting Review* **82**(4): 907–937.

Davila A, Foster G, Gupta M. 2003. Venture capital financing and the growth of startup firms. *Journal of Business Venturing* **18**(6): 689–708.

Dean TJ, McMullen JS. 2007. Toward a theory of sustainable entrepreneurship: Reducing environmental degradation through entrepreneurial action. *Journal of Business Venturing* **22**(1): 50–76.

Decarolis DM, Deeds DL. 1999. The impact of stocks and flows of organizational knowledge on firm performance: an empirical investigation of the biotechnology industry. *Strategic Management Journal* **20**(10): 953–968.

Deeds DL, Mang PY, Frandsen ML. 2004. The influence of firms' and industries' legitimacy on the flow of capital into high-technology ventures. *Strategic Organization* **2**(1): 9–34.

Delgado M, Ketels C, Porter ME, Stern S. 2012. *The determinants of national competitiveness*. National Bureau of Economic Research. Available at: http://www.nber.org/papers/w18249.

Delgado M, Porter ME, Stern S. 2010. Clusters and entrepreneurship. *Journal of Economic Geography* **10**(4): 495–518.

Delgado M, Porter ME, Stern S. 2014a. Clusters, convergence, and economic performance. *Research Policy* **43**(10): 1785–1799.

Delgado M, Porter ME, Stern S. 2014b. *Defining clusters of related industries*. National Bureau of Economic Research. Available at: http://www.nber.org/papers/w20375.

Delmar F, Shane S. 2003. Does business planning facilitate the development of new ventures? *Strategic Management Journal* **24**(12): 1165–1185.

Delmas M, Russo MV, Montes-Sancho MJ. 2007. Deregulation and environmental differentiation in the electric utility industry. *Strategic Management Journal* **28**(2): 189–209.

Denzau AT, North DC. 1994. Shared mental models: ideologies and institutions. *Kyklos* **47**(1): 3–31.

DeTienne DR. 2010. Entrepreneurial exit as a critical component of the entrepreneurial process: Theoretical development. *Journal of Business Venturing* **25**(2): 203–215.

DeTienne DR, Chirico F. 2013. Exit strategies in family firms: How socioemotional wealth drives the threshold of performance. *Entrepreneurship Theory and Practice* **37**(6): 1297–1318.

DeTienne DR, Shepherd DA, De Castro JO. 2008. The fallacy of 'only the strong survive': The effects of extrinsic motivation on the persistence decisions for under-performing firms. *Journal of Business Venturing* **23**(5): 528–546.

Dew N, Read S, Sarasvathy SD, Wiltbank R. 2009. Effectual versus predictive logics in entrepreneurial decision-making: differences between experts and novices. *Journal of Business Venturing* **24**(4): 287–309.

Dew N, Velamuri SR, Venkataraman S. 2004. Dispersed Knowledge and an Entrepreneurial Theory of the Firm. *Journal of Business Venturing* **19**(5): 659–679.

DiAddario S. 2011. Job search in thick markets. *Journal of Urban Economics* **69**(3): 303–318.

Dimov D, de Holan M, Milanov H. 2012. Learning Patterns in Venture Capital Investing in New Industries. *Industrial and Corporate Change* : 1–38.

Dimov D, Murray G. 2008. Determinants of the incidence and scale of seed capital investments by venture capital firms. *Small Business Economics* **30**(2): 127–152.

Drucker P. 2014. *Innovation and entrepreneurship*. Routledge: London, United Kingdom.

Dushnitsky G, Lenox MJ. 2005. When do firms undertake R&D by investing in new ventures? *Strategic Management Journal* **26**(10): 947–965.

Dushnitsky G, Shaver JM. 2009. Limitations to interorganizational knowledge acquisition: the paradox of corporate venture capital. *Strategic Management Journal* **30**: 1045–1064.

Dyer JH. 1996. Specialized supplier networks as a source of competitive advantage: evidence from the auto industry. *Strategic Management Journal* **17**(4): 271–291.

Dyer JH, Singh H. 1998. The relational view: cooperative strategy and sources of interorganizational competitive advantage. *Academy of Management Review* **23**(4): 660–679.

Eckhardt JT, Shane SA. 2003. Opportunities and entrepreneurship. *Journal of Management* **29**: 333–349.

Elster J. 1989. Social norms and economic theory. *The Journal of Economic Perspectives* **3**(4): 99–117.

Erikson T. 2002. Entrepreneurial capital: the emerging venture’s most important asset and competitive advantage. *Journal of Business Venturing* **17**(3): 275–290.

Fallick B, Fleischman CA, Rebitzer JB. 2006. Job-hopping in Silicon Valley: some evidence concerning the microfoundations of a high-technology cluster. *The Review of Economics and Statistics* **88**(3): 472–481.

Fan TPC. 2010. De novo venture strategy: Arch incumbency at inaugural entry. *Strategic Management Journal* **31**(1): 19–38.

Feld B. 2012. *Startup communities: building an entrepreneurial ecosystem in your city*. John Wiley & Sons: Hoboken, NJ.

Feldman M, Francis J, Bercovitz J. 2005. Creating a cluster while building a firm: Entrepreneurs and the formation of industrial clusters. *Regional Studies* **39**(1): 129–141.

Feldman MP. 2001. The entrepreneurial event revisited: firm formation in a regional context. *Industrial and Corporate Change* **10**(4): 861–891.

Fernhaber SA, Gilbert BA, McDougall PP. 2008. International entrepreneurship and geographic location: an empirical examination of new venture internationalization. *Journal of International Business Studies* **39**(2): 267–290.

Fischer MM. 2001. Innovation, knowledge creation and systems of innovation. *The Annals of Regional Science* **35**(2): 199–216.

Fitza M, Matusik SF, Mosakowski E. 2009. Do VCs matter? The importance of owners on performance variance in start-up firms. *Strategic Management Journal* **30**(4): 387–404.

Fleming L, Frenken K. 2007. The evolution of inventor networks in the Silicon Valley and Boston regions. *Advances in Complex Systems* **10**(01): 53–71.

Fleming L, King C, Juda AI. 2007. Small worlds and regional innovation. *Organization Science* **18**(6): 938–954.

Fleming L, Sorenson O. 2004. Science as a map in technological search. *Strategic Management Journal* **25**(8-9): 909–928.

Florida R. 2002. Bohemia and economic geography. *Journal of Economic Geography* **2**(1): 55–71.

Flyer F, Shaver JM. 2003. Location choices under agglomeration externalities and strategic interaction. *Advances in Strategic Management* **20**: 193–214.

Folta TB, Cooper AC, Baik Y. 2006a. Geographic cluster size and firm performance. *Journal of Business Venturing* **21**(2): 217–242.

Folta TB, Johnson DR, O'Brien J. 2006b. Uncertainty, irreversibility, and the likelihood of entry: An empirical assessment of the option to defer. *Journal of Economic Behavior & Organization* **61**(3): 432–452.

Folta TB, O'Brien JP. 2008. Determinants of firm-specific thresholds in acquisition decisions. *Managerial and Decision Economics* **29**(2-3): 209–225.

Forlani D, Mullins JW. 2000. Perceived risks and choices in entrepreneurs' new venture decisions. *Journal of Business Venturing* **15**(4): 305–322.

Freedman ML. 2008. Job hopping, earnings dynamics, and industrial agglomeration in the software publishing industry. *Journal of Urban Economics* **64**(3): 590–600.

Friedman TL. 2005. It's a flat world, after all. *The New York Times* : 33–37.

Gaglio CM, Katz JA. 2001. The psychological basis of opportunity identification: Entrepreneurial alertness. *Small Business Economics* **16**(2): 95–111.

Gardner M, Steinberg L. 2005. Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: an experimental study. *Developmental Psychology* **41**(4): 625.

Garmaise MJ. 2011. Ties that truly bind: Noncompetition agreements, executive compensation, and firm investment. *Journal of Law, Economics, and Organization* **27**(2): 376–425.

Garud R. 1997. On the distinction between know-how, know-what, and know-why. In *Advances in Strategic Management*. Emerald Group Publishing: Bingley, United Kingdom, 14: 81–101.

Garud R, Karnøe P. 2003. Bricolage versus breakthrough: distributed and embedded agency in technology entrepreneurship. *Research Policy* **32**(2): 277–300.

George G. 2005. Slack resources and the performance of privately held firms. *Academy of Management Journal* **48**(4): 661–676.

Ghio N, Guerini M, Lehmann EE, Rossi-Lamastra C. 2014. The emergence of the knowledge spillover theory of entrepreneurship. *Small Business Economics* **44**(1): 1–18.

Gilbert BA, Audretsch DB, McDougall PP. 2004. The emergence of entrepreneurship policy. *Small Business Economics* **22**(3-4): 313–323.

Gilbert BA, McDougall PP, Audretsch DB. 2008. Clusters, knowledge spillovers and new venture performance: an empirical examination. *Journal of Business Venturing* **23**(4): 405–422.

Gilson RJ. 2003. Engineering a venture capital market: lessons from the American experience. *Stanford Law Review* **55**(4): 1067–1103.

Gimeno J, Folta TB, Cooper AC, Woo CY. 1997. Survival of the fittest? Entrepreneurial human capital and the persistence of underperforming firms. *Administrative Science Quarterly* **42**(4): 750–783.

Gittelman M. 2007. Does geography matter for science-based firms? Epistemic communities and the geography of research and patenting in biotechnology. *Organization Science* **18**(4): 724–741.

Glaeser EL, Kerr WR, Ponzetto GA. 2010. Clusters of entrepreneurship. *Journal of Urban Economics* **67**(1): 150–168.

Globerman S, Shapiro D, Vining A. 2005. Clusters and intercluster spillovers: their influence on the growth and survival of Canadian information technology firms. *Industrial and Corporate Change* : 1–34.

Gnyawali DR, Srivastava MK. 2013. Complementary effects of clusters and networks on firm innovation: A conceptual model. *Journal of Engineering and Technology Management* **30**(1): 1–20.

Goetz SJ, Freshwater D. 2001. State-level determinants of entrepreneurship and a preliminary measure of entrepreneurial climate. *Economic Development Quarterly* **15**(1): 58–70.

Gompers PA, Lerner J. 2001. *The money of invention: How venture capital creates new wealth*. Harvard Business Press: Cambridge, MA.

Gompers P, Lerner J. 2000. Money chasing deals? The impact of fund inflows on private equity valuation. *Journal of Financial Economics* **55**(2): 281–325.

Gompers P, Lerner J. 2003. Short-term America revisited? Boom and bust in the venture capital industry and the impact on innovation. In *Innovation Policy and the Economy*. MIT Press: Boston, MA, 3: 1–28.

Gorman M, Sahlman WA. 1989. What do venture capitalists do? *Journal of Business Venturing* **4**(4): 231–248.

Granovetter M. 2005. The impact of social structure on economic outcomes. *Journal of Economic Perspectives* **19**(1): 33–50.

Grant RM. 1996. Toward a knowledge-based theory of the firm. *Strategic Management Journal* **17**: 109–122.

Greve HR. 2002. An ecological theory of spatial evolution: Local density dependence in Tokyo banking, 1894–1936. *Social Forces* **80**(3): 847–879.

Gruber M. 2007. Uncovering the value of planning in new venture creation: A process and contingency perspective. *Journal of Business Venturing* **22**(6): 782–807.

Guillén MF. 1998. International management and the circulation of ideas. *Trends in Organizational Behavior* **5**: 47–64.

Guzman J, Stern S. 2015. Where is Silicon Valley? *Science* **347**(6222): 606–609.

Hambrick DC. 2007. The field of management's devotion to theory: too much of a good thing? *Academy of Management Journal* **50**(6): 1346–1352.

Hanson GH. 2001. Scale economies and the geographic concentration of industry. *Journal of Economic Geography* **1**(3): 255–276.

Hayek FA. 1945. The use of knowledge in society. *The American Economic Review* **35**(4): 519–530.

Hayter R. 2008. Environmental economic geography. *Geography Compass* **2**(3): 831–850.

Hayward ML, Shepherd DA, Griffin D. 2006. A hubris theory of entrepreneurship. *Management Science* **52**(2): 160–172.

Headd B. 2003. Redefining business success: Distinguishing between closure and failure. *Small Business Economics* **21**(1): 51–61.

Heeley MB, Jacobson R. 2008. The recency of technological inputs and financial performance. *Strategic Management Journal* **29**(7): 723–744.

Helfat CE. 2007. Stylized facts, empirical research and theory development in management. *Strategic Organization* **5**(2): 185–192.

Hellmann T, Puri M. 2002. Venture capital and the professionalization of start-up firms: empirical evidence. *The Journal of Finance* **57**(1): 169–197.

Hess DJ. 2005. Technology-and product-oriented movements: Approximating social movement studies and science and technology studies. *Science, Technology & Human Values* **30**(4): 515–535.

Hiatt S. 2010. *Institutional actors and entrepreneurial choices: New ventures in the biodiesel fuel industry*. Cornell University, Ithaca, NY. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1658425.

Hiatt SR, Sine WD, Tolbert PS. 2009. From Pabst to Pepsi: The deinstitutionalization of social practices and the creation of entrepreneurial opportunities. *Administrative Science Quarterly* **54**(4): 635–667.

Hill CW, Rothaermel FT. 2003. The performance of incumbent firms in the face of radical technological innovation. *Academy of Management Review* **28**(2): 257–274.

Hockerts K, Wüstenhagen R. 2010. Greening Goliaths versus emerging Davids—Theorizing about the role of incumbents and new entrants in sustainable entrepreneurship. *Journal of Business Venturing* **25**(5): 481–492.

Hoffman AJ. 1999. Institutional evolution and change: Environmentalism and the US chemical industry. *Academy of Management Journal* **42**(4): 351–371.

Hoopes DG, Madsen TL, Walker G. 2003. Guest editors' introduction to the special issue: why is there a resource-based view? Toward a theory of competitive heterogeneity. *Strategic Management Journal* **24**(10): 889–902.

Hsieh C, Nickerson JA, Zenger TR. 2007. Opportunity discovery, problem solving and a theory of the entrepreneurial firm. *Journal of Management Studies* **44**(7): 1255–1277.

Hsu DH. 2007. Experienced entrepreneurial founders, organizational capital, and venture capital funding. *Research Policy* **36**(5): 722–741.

Huber F. 2012. Do clusters really matter for innovation practices in Information Technology? Questioning the significance of technological knowledge spillovers. *Journal of Economic Geography* **12**(1): 107–126.

Ingram P, Simons T. 2000. State formation, ideological competition, and the ecology of Israeli workers' cooperatives, 1920–1992. *Administrative Science Quarterly* **45**(1): 25–53.

Isenberg DJ. 1986. Thinking and managing: A verbal protocol analysis of managerial problem solving. *Academy of Management Journal* **29**(4): 775–788.

Isenberg DJ. 2010. How to start an entrepreneurial revolution. *Harvard Business Review* **88**(6): 40–50.

Jacobsson S, Johnson A. 2000. The diffusion of renewable energy technology: an analytical framework and key issues for research. *Energy Policy* **28**(9): 625–640.

Jaffe AB, Trajtenberg M, Henderson R. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics* **108**(3): 577–598.

Jenkins M, Tallman S. 2010. The shifting geography of competitive advantage: clusters, networks and firms. *Journal of Economic Geography* **10**(4): 599–618.

Jennings PD, Greenwood R, Lounsbury MD, Suddaby R. 2013. Institutions, entrepreneurs, and communities: A special issue on entrepreneurship. *Journal of Business Venturing* **28**(1): 1–9.

Johnson B, Lorenz E, Lundvall B-A. 2002. Why all this fuss about codified and tacit knowledge? *Industrial and Corporate Change* **11**(2): 245–262.

Johnson-Laird PN. 1983. *Mental models: Towards a cognitive science of language, inference, and consciousness*. Harvard University Press: Cambridge, MA.

Jonsson S, Regnér P. 2009. Normative barriers to imitation: social complexity of core competences in a mutual fund industry. *Strategic Management Journal* **30**(5): 517–536.

Kahn ME. 2007. Do greens drive Hummers or hybrids? Environmental ideology as a determinant of consumer choice. *Journal of Environmental Economics and Management* **54**(2): 129–145.

Kahn ME, Vaughn RK. 2009. Green market geography: The spatial clustering of hybrid vehicles and LEED registered buildings. *The BE Journal of Economic Analysis & Policy* **9**(2).

Kaish S, Gilad B. 1991. Characteristics of opportunities search of entrepreneurs versus executives: Sources, interests, general alertness. *Journal of Business Venturing* **6**(1): 45–61.

Kalnins A, Chung W. 2004. Resource-seeking agglomeration: a study of market entry in the lodging industry. *Strategic Management Journal* **25**(7): 689–699.

Kapoor R, Furr NR. 2014. Complementarities and competition: Unpacking the drivers of entrants' technology choices in the solar photovoltaic industry. *Strategic Management Journal* **36**(3): 416–436.

Katila R. 2002. New product search over time: past ideas in their prime? *Academy of Management Journal* **45**(5): 995–1010.

Katz JA. 2003. The chronology and intellectual trajectory of American entrepreneurship education: 1876–1999. *Journal of Business Venturing* **18**(2): 283–300.

Katz JA. 2008. Fully mature but not fully legitimate: a different perspective on the state of entrepreneurship education. *Journal of Small Business Management* **46**(4): 550–566.

Kenney M. 2014. The Connected City. *The Buzz in Cities* : 45.

Kirzner IM. 1979. *Perception, opportunity, and profit: Studies in the theory of entrepreneurship*. University of Chicago Press: Chicago, IL.

Klepper S. 2009. Spinoffs: A review and synthesis. *European Management Review* **6**(3): 159–171.

Klepper S. 2010. The origin and growth of industry clusters: The making of Silicon Valley and Detroit. *Journal of Urban Economics* **67**(1): 15–32.

Klepper S, Sleeper S. 2005. Entry by spinoffs. *Management science* **51**(8): 1291–1306.

Klepper S, Thompson P. 2005. Spinoff entry in high-tech industries: motives and consequences. *Economic Perspectives on Innovation, Cambridge University Press* **6**: 187–218.

Knack S. 1992. Civic norms, social sanctions, and voter turnout. *Rationality and Society* **4**(2): 133–156.

Knight FH. 1921. *Risk, uncertainty and profit*. Houghton Mifflin: Boston, MA.

Knott AM, Posen HE. 2005. Is failure good? *Strategic Management Journal* **26**(7): 617–641.

Kortum S, Lerner J. 2000. Assessing the contribution of venture capital to innovation. *RAND Journal of Economics* **31**(4): 674–692.

Kotha S. 2010. Spillovers, spill-ins, and strategic entrepreneurship: America's first commercial jet airplane and Boeing's ascendancy in commercial aviation. *Strategic Entrepreneurship Journal* **4**(4): 284–306.

Krugman P. 1991. Increasing returns and economic geography. *The Journal of Political Economy* **99**(3): 483–499.

Kuratko DF. 2005. The emergence of entrepreneurship education: Development, trends, and challenges. *Entrepreneurship Theory and Practice* **29**(5): 577–598.

Lafuente E, Vaillant Y, Rialp J. 2007. Regional differences in the influence of role models: comparing the entrepreneurial process of rural Catalonia. *Regional Studies* **41**(6): 779–796.

Lambooy JG. 2010. Knowledge transfers, spillovers and actors: The role of context and social capital. *European Planning Studies* **18**(6): 873–891.

Landier A, Thesmar D. 2009. Financial contracting with optimistic entrepreneurs. *Review of Financial Studies* **22**(1): 117–150.

Land KC, McCall PL, Nagin DS. 1996. A comparison of Poisson, negative binomial, and semiparametric mixed Poisson regression models with empirical applications to criminal careers data. *Sociological Methods & Research* **24**(4): 387–442.

Langlois RN, Cosgel MM. 1993. Frank Knight on risk, uncertainty, and the firm: a new interpretation. *Economic Inquiry* **31**(3): 456–465.

Laursen K, Masciarelli F, Prencipe A. 2012. Regions matter: how localized social capital affects innovation and external knowledge acquisition. *Organization Science* **23**(1): 177–193.

Lawson C. 1999. Towards a competence theory of the region. *Cambridge Journal of Economics* **23**(2): 151–166.

Lawson C, Lorenz E. 1999. Collective learning, tacit knowledge and regional innovative capacity. *Regional Studies* **33**(4): 305–317.

Lenox M, York JG. 2011. Environmental entrepreneurship. *Oxford University Press, Oxford, UK.* Available at: http://leeds.colorado.edu/publication/300.

Lerner J. 2009. *Boulevard of broken dreams: why public efforts to boost entrepreneurship and venture capital have failed–and what to do about it*. Princeton University Press: Princeton, NJ.

Lerner J. 2010. The future of public efforts to boost entrepreneurship and venture capital. *Small Business Economics* **35**(3): 255–264.

Li G-C *et al.* 2014. Disambiguation and co-authorship networks of the US patent inventor database (1975–2010). *Research Policy* **43**(6): 941–955.

Lockett A, Siegel D, Wright M, Ensley MD. 2005. The creation of spin-off firms at public research institutions: Managerial and policy implications. *Research Policy* **34**(7): 981–993.

Lomi A. 1995. The population ecology of organizational founding: Location dependence and unobserved heterogeneity. *Administrative Science Quarterly* **40**(1): 111–144.

Louis MR, Sutton RI. 1991. Switching cognitive gears: From habits of mind to active thinking. *Human Relations* **44**(1): 55–76.

Lounsbury M, Glynn MA. 2001. Cultural entrepreneurship: Stories, legitimacy, and the acquisition of resources. *Strategic Management Journal* **22**(6-7): 545–564.

Lunn M, McNeil D. 1995. Applying Cox regression to competing risks. *Biometrics* **51**(2): 524–532.

Mann RJ, Sager TW. 2007. Patents, venture capital, and software start-ups. *Research Policy* **36**(2): 193–208.

March JG, Shapira Z. 1987. Managerial perspectives on risk and risk taking. *Management Science* **33**(11): 1404–1418.

March JG, Shapira Z. 1992. Variable risk preferences and the focus of attention. *Psychological Review* **99**(1): 172.

Markman GD, Gartner WB. 2002. Is Extraordinary Growth Profitable? A Study of Inc. 500 High-Growth Companies*. *Entrepreneurship Theory and Practice* **27**(1): 65–75.

Marquis C, Battilana J. 2009. Acting globally but thinking locally? The enduring influence of local communities on organizations. *Research in Organizational Behavior* **29**: 283–302.

Marshall A. 1890. *Principles of Economics*. Macmillan: London, United Kingdom.

Martin R, Sunley P. 2006. Path dependence and regional economic evolution. *Journal of Economic Geography* **6**(4): 395–437.

Marx M. 2011. The Firm Strikes Back Non-compete Agreements and the Mobility of Technical Professionals. *American Sociological Review* **76**(5): 695–712.

Marx M, Strumsky D, Fleming L. 2009. Mobility, skills, and the Michigan non-compete experiment. *Management Science* **55**(6): 875–889.

Mason C, Brown R. 2013a. Entrepreneurial ecosystems and growth oriented entrepreneurship. In *background paper for the International Workshop on Entrepreneurial Ecosystems and Growth Oriented Entrepreneurship*, 7.

Mason C, Brown R. 2013b. Creating good public policy to support high-growth firms. *Small Business Economics* **40**(2): 211–225.

Matusik SF, Fitza MA. 2012. Diversification in the venture capital industry: leveraging knowledge under uncertainty. *Strategic Management Journal* **33**(4): 407–426.

Ma X, Tong TW, Fitza M. 2013. How much does subnational region matter to foreign subsidiary performance? Evidence from Fortune Global 500 corporations' investment in China. *Journal of International Business Studies* **44**(1): 66–87.

McCann BT, Folta TB. 2008. Location matters: where we have been and where we might go in agglomeration research. *Journal of Management* **34**(3): 532–565.

McCann BT, Folta TB. 2011. Performance differentials within geographic clusters. *Journal of Business Venturing* **26**(1): 104–123.

McCann BT, Folta TB. 2012. Entrepreneurial entry thresholds. *Journal of Economic Behavior & Organization* **84**(3): 782–800.

McGahan AM, Porter ME. 2002. What do we know about variance in accounting profitability? *Management Science* **48**(7): 834–851.

McMullen JS, Shepherd DA. 2006. Entrepreneurial action and the role of uncertainty in the theory of the entrepreneur. *Academy of Management Review* **31**(1): 132–152.

Meek WR, Pacheco DF, York JG. 2010. The impact of social norms on entrepreneurial action: Evidence from the environmental entrepreneurship context. *Journal of Business Venturing* **25**(5): 493–509.

Mendes R, Van Den Berg GJ, Lindeboom M. 2010. An empirical assessment of assortative matching in the labor market. *Labour Economics* **17**(6): 919–929.

Mishina Y, Pollock TG, Porac JF. 2004. Are more resources always better for growth? Resource stickiness in market and product expansion. *Strategic Management Journal* **25**(12): 1179–1197.

Moran P, Ghoshal S. 1999. Markets, firms, and the process of economic development. *Academy of Management Review* **24**(3): 390–412.

Morgan K. 2004. The exaggerated death of geography: learning, proximity and territorial innovation systems. *Journal of Economic Geography* **4**(1): 3–21.

Naffziger DW, Hornsby JS, Kuratko DF. 1994. A proposed research model of entrepreneurial motivation. *Entrepreneurship Theory and Practice* **18**: 29–29.

Nanda R, Sørensen JB. 2010. Workplace peers and entrepreneurship. *Management Science* **56**(7): 1116–1126.

Nanda R, Younge K, Fleming L. 2013. Innovation and Entrepreneurship in Renewable Energy. In *The Changing Frontier: Rethinking Science and Innovation Policy*. University of Chicago Press. Available at: http://www.nber.org/chapters/c13048.pdf.

Neck HM, Greene PG, Brush CG. 2014. *Teaching entrepreneurship: A practice-based approach*. Edward Elgar Publishing: Cheltenham, United Kingdom.

Neffke F, Henning M, Boschma R, Lundquist K-J, Olander L-O. 2011. The dynamics of agglomeration externalities along the life cycle of industries. *Regional Studies* **45**(1): 49–65.

Nelson FD. 1977. Censored regression models with unobserved, stochastic censoring thresholds. *Journal of Econometrics* **6**(3): 309–327.

Nelson RR. 1994. The co-evolution of technology, industrial structure, and supporting institutions. *Industrial and Corporate Change* **3**(1): 47–63.

Nelson RR, Winter SG. 1982. *An evolutionary theory of economic change*. Belknap Press: Cambridge, MA.

Nesta L, Saviotti P-P. 2006. Firm knowledge and market value in biotechnology. *Industrial and Corporate Change* **15**(4): 625–652.

Nisar TM. 2005. Investor influence on portfolio company growth and development strategy. *The Journal of Private Equity* **9**(1): 22–35.

North DC. 1990. *Institutions, institutional change and economic performance*. Cambdrige University Press: Cambridge, MA.

Oakey RP, Cooper SY. 1989. High technology industry, agglomeration and the potential for peripherally sited small firms. *Regional Studies* **23**(4): 347–360.

Obama B. 2010. Presidential Summit on Entrepreneurship. *The White House* **26**.

O'Brien JP, Folta TB, Johnson DR. 2003. A real options perspective on entrepreneurial entry in the face of uncertainty. *Managerial and Decision Economics* **24**(8): 515–533.

Oliver C. 1991. Strategic responses to institutional processes. *Academy of Management Review* **16**(1): 145–179.

O'Mahoney J, Heusinkveld S, Wright C. 2013. Commodifying the commodifiers: the impact of procurement on management knowledge. *Journal of Management Studies* **50**(2): 204–235.

Operti E, Carnabuci G. 2014. Public Knowledge, Private Gain: The Effect of Spillover Networks on Firms' Innovative Performance. *Journal of Management* **40**(4): 1042–1074.

Ordóñez LD, Schweitzer ME, Galinsky AD, Bazerman MH. 2009. Goals gone wild: The systematic side effects of overprescribing goal setting. *The Academy of Management Perspectives* **23**(1): 6–16.

Pacheco DF, York JG, Hargrave TJ. 2014. The coevolution of industries, social movements, and institutions: wind power in the United States. *Organization Science* **25**(6): 1609–1632.

Packalen KA. 2007. Complementing capital: The role of status, demographic features, and social capital in founding teams' abilities to obtain resources. *Entrepreneurship Theory and Practice* **31**(6): 873–891.

Pe'er A, Keil T. 2013. Are all startups affected similarly by clusters? Agglomeration, competition, firm heterogeneity, and survival. *Journal of Business Venturing* **28**(3): 354–372.

Pe'er A, Vertinsky I. 2008. Firm exits as a determinant of new entry: Is there evidence of local creative destruction? *Journal of Business Venturing* **23**(3): 280–306.

Pe'er A, Vertinsky I, King A. 2008. Who enters, where and why? The influence of capabilities and initial resource endowments on the location choices of de novo enterprises. *Strategic Organization* **6**(2): 119–149.

Penrose ET. 1959. *The theory of the growth of the firm*. Oxford University Press: Oxford, United Kingdom.

Pernick R, Wilder C. 2007. *The clean tech revolution: The next big growth and investment opportunity*. HarperCollins: New York, NY.

Peteraf MA. 1993. The cornerstones of competitive advantage: A resource-based view. *Strategic Management Journal* **14**(3): 179–191.

Petkova A, Wadhwa A, Yao X, Jain S. 2014. Reputation and decision making under ambiguity: a study of us venture capital firms' investments in the emerging clean energy sector. *Academy of Management Journal* **57**(2): 422–448.

Pfeffer J. 2007. A modest proposal: How we might change the process and product of managerial research. *Academy of Management Journal* **50**(6): 1334–1345.

Phene A, Fladmoe-Lindquist K, Marsh L. 2006. Breakthrough innovations in the US biotechnology industry: the effects of technological space and geographic origin. *Strategic Management Journal* **27**(4): 369–388.

Pierce JR, Aguinis H. 2013. The too-much-of-a-good-thing effect in management. *Journal of Management* **39**(2): 313–338.

Pisano GP. 1994. Knowledge, integration, and the locus of learning: an empirical analysis of process development. *Strategic Management Journal* **15**(S1): 85–100.

Plummer LA, Acs ZJ. 2014. Localized competition in the knowledge spillover theory of entrepreneurship. *Journal of Business Venturing* **29**(1): 121–136.

Plummer LA, Pe'er A. 2010. The Geography of Entrepreneurship. In *Handbook of Entrepreneurship Research*. Springer: Berlin, Germany: 519–556.

Polanyi M. 1966. The logic of tacit inference. *Philosophy* **41**(155): 1–18.

Porter M. 2003. The economic performance of regions. *Regional Studies* **37**(6-7): 545–546.

Porter ME. 1990. *The competitive advantage of nations: creating and sustaining competitive advantage*. Simon and Schuster: New York, NY.

Porter ME. 1998. Clusters and the new economics of competition. *Harvard Business Review* **76**(6): 77–90.

Porter ME. 2000. Location, competition, and economic development: Local clusters in a global economy. *Economic Development Quarterly* **14**(1): 15–34.

Potter A, Watts HD. 2011. Evolutionary agglomeration theory: increasing returns, diminishing returns, and the industry life cycle. *Journal of Economic Geography* **11**(3): 417–455.

Pouder R, John CHS. 1996. Hot spots and blind spots: geographical clusters of firms and innovation. *Academy of Management Review* **21**(4): 1192–1225.

Powell WW, Koput KW, Bowie JI, Smith-Doerr L. 2002. The spatial clustering of science and capital: accounting for biotech firm-venture capital relationships. *Regional Studies* **36**(3): 291–305.

Prevezer M. 1997. The dynamics of industrial clustering in biotechnology. *Small Business Economics* **9**(3): 255–271.

Puga D. 2010. The magnitude and causes of agglomeration economies. *Journal of Regional Science* **50**(1): 203–219.

Rao H. 2004. Institutional activism in the early American automobile industry. *Journal of Business Venturing* **19**(3): 359–384.

Richards D. 2001. Coordination and shared mental models. *American Journal of Political Science* **45**(2): 259–276.

Richards W, McKay BD, Richards D. 2002. The probability of collective choice with shared knowledge structures. *Journal of Mathematical Psychology* **46**(3): 338–351.

Ries E. 2011. *The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses*. Random House LLC: New York, NY.

Rindova V, Ferrier WJ, Wiltbank R. 2010. Value from gestalt: how sequences of competitive actions create advantage for firms in nascent markets. *Strategic Management Journal* **31**(13): 1474–1497.

Rindova VP, Petkova AP, Kotha S. 2007. Standing out: how new firms in emerging markets build reputation. *Strategic Organization* **5**(1): 31–70.

Ritzer G, Ryan JM. 2010. *The concise encyclopedia of sociology*. John Wiley & Sons: Hoboken, NJ.

Robb AM, Watson J. 2012. Gender differences in firm performance: evidence from new ventures in the United States. *Journal of Business Venturing* **27**(5): 544–558.

Rocha HO. 2004. Entrepreneurship and development: The role of clusters. *Small Business Economics* **23**(5): 363–400.

Rocha HO, Sternberg R. 2005. Entrepreneurship: The role of clusters theoretical perspectives and empirical evidence from Germany. *Small Business Economics* **24**(3): 267–292.

Romanelli E, Khessina OM. 2005. Regional industrial identity: Cluster configurations and economic development. *Organization Science* **16**(4): 344–358.

Rosa P. 1998. Entrepreneurial processes of business cluster formation and growth by'habitual'entrepreneurs. *Entrepreneurship Theory and Practice* **22**: 43–62.

Rosenkopf L, Almeida P. 2003. Overcoming local search through alliances and mobility. *Management Science* **49**(6): 751–766.

Rosenthal SS, Strange WC. 2003. Geography, industrial organization, and agglomeration. *Review of Economics and Statistics* **85**(2): 377–393.

Rothaermel FT. 2001. Incumbent's advantage through exploiting complementary assets via interfirm cooperation. *Strategic Management Journal* **22**(6-7): 687–699.

Roth W-M. 2009. Radical uncertainty in scientific discovery work. *Science, Technology & Human Values* **34**(3): 313–336.

Ruhnka JC, Young JE. 1987. A venture capital model of the development process for new ventures. *Journal of Business Venturing* **2**(2): 167–184.

Rumelt RP. 1991. How much does industry matter? *Strategic Management Journal* **12**(3): 167–185.

Russo MV. 2003. The emergence of sustainable industries: building on natural capital. *Strategic Management Journal* **24**(4): 317–331.

Samila S, Sorenson O. 2010. Venture capital as a catalyst to commercialization. *Research Policy* **39**(10): 1348–1360.

Samila S, Sorenson O. 2011. Venture capital, entrepreneurship, and economic growth. *The Review of Economics and Statistics* **93**(1): 338–349.

Samuelsson M, Davidsson P. 2009. Does venture opportunity variation matter? Investigating systematic process differences between innovative and imitative new ventures. *Small Business Economics* **33**(2): 229–255.

Sanders WM, Boivie S. 2004. Sorting things out: Valuation of new firms in uncertain markets. *Strategic Management Journal* **25**(2): 167–186.

Sandri S, Schade C, Musshoff O, Odening M. 2010. Holding on for too long? An experimental study on inertia in entrepreneurs' and non-entrepreneurs' disinvestment choices. *Journal of Economic Behavior & Organization* **76**(1): 30–44.

Santos FM, Eisenhardt KM. 2009. Constructing markets and shaping boundaries: entrepreneurial power in nascent fields. *Academy of Management Journal* **52**(4): 643–671.

Sapienza HJ. 1992. When do venture capitalists add value? *Journal of Business Venturing* **7**(1): 9–27.

Saxenian A. 1996. *Regional advantage: culture and competition in Silicon Valley and Route 128*. Harvard University Press: Cambridge, MA.

Schilling MA, Esmundo M. 2009. Technology S-curves in renewable energy alternatives: Analysis and implications for industry and government. *Energy Policy* **37**(5): 1767–1781.

Shah KS, Smith SW. 2010. Intellectual property, prior knowledge and the survival of new firms. Available at: http://mackinstitute.wharton.upenn.edu/wp-content/uploads/2012/12/ShahWinston-Smith_Intellectual-Property-Prior-Knowledge-the-Survival-of-New-Firms.pdf.

Shane S. 2000. Prior knowledge and the discovery of entrepreneurial opportunities. *Organization Science* **11**(4): 448–469.

Shane S. 2001. Technology regimes and new firm formation. *Management science* **47**(9): 1173–1190.

Shane S. 2008. *The handbook of technology and innovation management*. John Wiley & Sons: Hoboken, NJ.

Shane S. 2009. Why encouraging more people to become entrepreneurs is bad public policy. *Small Business Economics* **33**(2): 141–149.

Shane S, Venkataraman S. 2000. The promise of entrepreneurship as a field of research. *Academy of Management Review* **25**(1): 217–226.

Sharpe WF. 1975. Adjusting for risk in portfolio performance measurement. *The Journal of Portfolio Management* **1**(2): 29–34.

Shaver JM, Flyer F. 2000. Agglomeration economies, firm heterogeneity, and foreign direct investment in the United States. *Strategic Management Journal* **21**(12): 1175–1194.

Shepherd DA. 2003. Learning from business failure: Propositions of grief recovery for the self-employed. *Academy of Management Review* **28**(2): 318–328.

Shepherd DA, Williams TA, Patzelt H. 2015. Thinking About Entrepreneurial Decision Making Review and Research Agenda. *Journal of Management* **41**(1): 11–46.

Shimer R, Smith L. 2000. Assortative matching and search. *Econometrica* **68**(2): 343–369.

Short JC, McKelvie A, Ketchen DJ, Chandler GN. 2009. Firm and industry effects on firm performance: a generalization and extension for new ventures. *Strategic Entrepreneurship Journal* **3**(1): 47–65.

Short JC, Palmer TB. 2003. Organizational performance referents: An empirical examination of their content and influences. *Organizational Behavior and Human Decision Processes* **90**(2): 209–224.

Sidanius J. 1985. Cognitive functioning and sociopolitical ideology revisited. *Political Psychology* : 637–661.

Simon M, Houghton SM, Aquino K. 2000. Cognitive biases, risk perception, and venture formation: How individuals decide to start companies. *Journal of Business Venturing* **15**(2): 113–134.

Simons T, Ingram P. 2004. An ecology of ideology: Theory and evidence from four populations. *Industrial and Corporate Change* **13**(1): 33–59.

Sine WD, David RJ. 2010. Institutions and entrepreneurship. *Research in the Sociology of Work* **21**: 1–26.

Sine WD, Lee BH. 2009. Tilting at windmills? The environmental movement and the emergence of the US wind energy sector. *Administrative Science Quarterly* **54**(1): 123–155.

Sirmon DG, Hitt MA, Ireland RD. 2007. Managing firm resources in dynamic environments to create value: Looking inside the black box. *Academy of Management Review* **32**(1): 273–292.

Sirmon DG, Hitt MA, Ireland RD, Gilbert BA. 2011. Resource orchestration to create competitive advantage breadth, depth, and life cycle effects. *Journal of Management* **37**(5): 1390–1412.

Sitkin SB, Pablo AL. 1992. Reconceptualizing the determinants of risk behavior. *Academy of management review* **17**(1): 9–38.

Smith BR, Matthews CH, Schenkel MT. 2009. Differences in entrepreneurial opportunities: the role of tacitness and codification in opportunity identification. *Journal of Small Business Management* **47**(1): 38–57.

Somaya D, Williamson IO, Lorinkova N. 2008. Gone but not lost: The different performance impacts of employee mobility between cooperators versus competitors. *Academy of Management Journal* **51**(5): 936–953.

Song J, Almeida P, Wu G. 2003. Learning-by-hiring: When is mobility more likely to facilitate interfirm knowledge transfer? *Management Science* **49**(4): 351–365.

Sorensen JB, Sorenson O. 2003. From conception to birth: Opportunity perception and resource mobilization in entrepreneurship. *Advances in Strategic Management* **20**: 89–118.

Sorenson O, Audia PG. 2000. The Social Structure of Entrepreneurial Activity: Geographic Concentration of Footwear Production in the United States, 1940–19891. *American Journal of Sociology* **106**(2): 424–462.

Sorenson O, Baum JA. 2003. Editors' introduction: geography and strategy: the strategic management of space and place. *Advances in Strategic Management* **20**: 1–22.

Spencer GM, Vinodrai T, Gertler MS, Wolfe DA. 2010. Do clusters make a difference? Defining and assessing their economic performance. *Regional Studies* **44**(6): 697–715.

Srivastava MK, Gnyawali DR. 2011. When do relational resources matter? Leveraging portfolio technological resources for breakthrough innovation. *Academy of Management Journal* **54**(4): 797–810.

Staber U. 2001. Spatial Proximity and Firm Survival in a Declining Industrial District: The Case of Knitwear Firms in Baden-Wurttemberg. *Regional Studies* **35**(4): 329–341.

Stam E. 2007. Why butterflies don‘t leave: Locational behavior of entrepreneurial firms. *Economic Geography* **83**(1): 27–50.

Stam E. 2010. Entrepreneurship, evolution and geography. In *The Handbook of Evolutionary Economic Geography*: 307–348.

Stern PC, Dietz T, Abel T, Guagnano GA, Kalof L. 1999. A value-belief-norm theory of support for social movements: The case of environmentalism. *Human Ecology Review* **6**(2): 81–98.

Storper M. 1995. The resurgence of regional economies, ten years later the region as a nexus of untraded interdependencies. *European Urban and Regional Studies* **2**(3): 191–221.

Stuart TE, Podolny JM. 1996. Local search and the evolution of technological capabilities. *Strategic Management Journal* **17**(S1): 21–38.

Stuart TE, Sorenson O. 2003. Liquidity events and the geographic distribution of entrepreneurial activity. *Administrative Science Quarterly* **48**(2): 175–201.

Suddaby R, Greenwood R. 2001. Colonizing knowledge: commodification as a dynamic of jurisdictional expansion in professional service firms. *Human Relations* **54**(7): 933–953.

Sullivan SE, Arthur MB. 2006. The evolution of the boundaryless career concept: Examining physical and psychological mobility. *Journal of Vocational Behavior* **69**(1): 19–29.

Tallman S, Jenkins M, Henry N, Pinch S. 2004. Knowledge, clusters, and competitive advantage. *Academy of Management Review* **29**(2): 258–271.

Teece DJ. 1986. Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy* **15**(6): 285–305.

Tempel A, Walgenbach P. 2007. Global standardization of organizational forms and management practices? What new institutionalism and the business-systems approach can learn from each other. *Journal of Management Studies* **44**(1): 1–24.

Ter Wal AL. 2013. The dynamics of the inventor network in German biotechnology: geographic proximity versus triadic closure. *Journal of Economic Geography* : 1–32.

Thornton PH, Ribeiro-Soriano D, Urbano D. 2011. Socio-cultural factors and entrepreneurial activity: An overview. *International Small Business Journal* **29**(2): 1–14.

Tian X. 2011. The causes and consequences of venture capital stage financing. *Journal of Financial Economics* **101**(1): 132–159.

Toft-Kehler R, Wennberg K, Kim PH. 2014. Practice makes perfect: Entrepreneurial-experience curves and venture performance. *Journal of Business Venturing* **29**(4): 453–470.

Tolbert PS, David RJ, Sine WD. 2011. Studying choice and change: The intersection of institutional theory and entrepreneurship research. *Organization Science* **22**(5): 1332–1344.

Townsend DM, Hart TA. 2008. Perceived institutional ambiguity and the choice of organizational form in social entrepreneurial ventures. *Entrepreneurship Theory and Practice* **32**(4): 685–700.

Tripsas M. 1997. Unraveling the process of creative destruction: Complementary assets and incumbent survival in the typesetter industry. *Strategic Management Journal* **18**(s1): 119–142.

Tripsas M, Gavetti G. 2000. Capabilities, cognition, and inertia: Evidence from digital imaging. *Strategic Management Journal* **21**(10-11): 1147–1161.

Ucbasaran D, Shepherd DA, Lockett A, Lyon SJ. 2013. Life after business failure the process and consequences of business failure for entrepreneurs. *Journal of Management* **39**(1): 163–202.

Uzzi B. 1996. The sources and consequences of embeddedness for the economic performance of organizations: The network effect. *American Sociological Review* **61**(4): 674–698.

Vaghely IP, Julien P-A. 2010. Are opportunities recognized or constructed?: An information perspective on entrepreneurial opportunity identification. *Journal of Business Venturing* **25**(1): 73–86.

Venkataraman S. 1997. The distinctive domain of entrepreneurship research: An editor’s perspective. *Advances in Entrepreneurship, Firm Emergence, and Growth* **3**: 119–138.

Venkataraman S. 2004. Regional transformation through technological entrepreneurship. *Journal of Business Venturing* **19**(1): 153–167.

Vissa B, Chacar AS. 2009. Leveraging ties: the contingent value of entrepreneurial teams' external advice networks on Indian software venture performance. *Strategic Management Journal* **30**(11): 1179–1191.

Vogel P. 2013. The employment outlook for youth: Building entrepreneurial ecosystems as a way forward. In *Conference Paper for the G20 Youth Forum*. Available at: http://www.entrepreneurssship.org/uploads/1/0/6/4/10642206/ecosystems_paper_petervogel.pdf.

Wadhwa A, Basu S. 2013. Exploration and resource commitments in unequal partnerships: An examination of corporate venture capital investments. *Journal of Product Innovation Management* **30**(5): 916–936.

Wadhwa A, Kotha S. 2006. Knowledge creation through external venturing: Evidence from the telecommunications equipment manufacturing industry. *Academy of Management Journal* **49**(4): 819–835.

Walker G, Devine-Wright P, Hunter S, High H, Evans B. 2010. Trust and community: Exploring the meanings, contexts and dynamics of community renewable energy. *Energy Policy* **38**(6): 2655–2663.

Weigel RH. 1977. Ideological and demographic correlates of proecology behavior. *The Journal of Social Psychology* **103**(1): 39–47.

Wei LJ. 1992. The accelerated failure time model: a useful alternative to the Cox regression model in survival analysis. *Statistics in Medicine* **11**(14-15): 1871–1879.

Wenger E. 2000. Communities of practice and social learning systems. *Organization* **7**(2): 225–246.

Wennberg K, Lindqvist G. 2010. The effect of clusters on the survival and performance of new firms. *Small Business Economics* **34**(3): 221–241.

Wennberg K, Wiklund J, DeTienne DR, Cardon MS. 2010. Reconceptualizing entrepreneurial exit: Divergent exit routes and their drivers. *Journal of Business Venturing* **25**(4): 361–375.

Wennekers S, Thurik R. 1999. Linking entrepreneurship and economic growth. *Small Business Economics* **13**(1): 27–56.

Wheeler CH. 2001. Search, sorting, and urban agglomeration. *Journal of Labor Economics* **19**(4): 879–899.

Wheeler CH. 2008. Local market scale and the pattern of job changes among young men. *Regional Science and Urban Economics* **38**(2): 101–118.

Whittington KB, Owen-Smith J, Powell WW. 2009. Networks, propinquity, and innovation in knowledge-intensive industries. *Administrative Science Quarterly* **54**(1): 90–122.

Wijbenga FH, Postma TJ, Stratling R. 2007. The influence of the venture capitalist's governance activities on the entrepreneurial firm's control systems and performance. *Entrepreneurship Theory and Practice* **31**(2): 257–277.

Wilson J. 1973. *Introduction to social movements*. Basic Books: New York, NY.

Winter SG. 1984. Schumpeterian competition in alternative technological regimes. *Journal of Economic Behavior & Organization* **5**(3): 287–320.

Wright M, Hmieleski KM, Siegel DS, Ensley MD. 2007. The role of human capital in technological entrepreneurship. *Entrepreneurship Theory and Practice* **31**(6): 791–806.

Yang H, Phelps C, Steensma HK. 2010. Learning from what others have learned from you: The effects of knowledge spillovers on originating firms. *Academy of Management Journal* **53**(2): 371–389.

York JG, Lenox MJ. 2014. Exploring the sociocultural determinants of de novo versus de alio entry in emerging industries. *Strategic Management Journal* **35**(13): 1930–1951.

York JG, Venkataraman S. 2010. The entrepreneur–environment nexus: Uncertainty, innovation, and allocation. *Journal of Business Venturing* **25**(5): 449–463.

Zacharakis AL, Shepherd DA, Coombs JE. 2003. The development of venture-capital-backed internet companies: An ecosystem perspective. *Journal of Business Venturing* **18**(2): 217–231.

Zimmerman MA, Zeitz GJ. 2002. Beyond survival: Achieving new venture growth by building legitimacy. *Academy of Management Review* **27**(3): 414–431.

Zook MA. 2002. Grounded capital: venture financing and the geography of the Internet industry, 1994–2000. *Journal of Economic Geography* **2**(2): 151–177.

Zucker LG, Darby MR, Brewer MB. 1998. Intellectual human capital and the birth of US biotechnology enterprises. *American Economic Review* **88**(1): 290–306.