ESSAYS IN HIGH-IMPACT COMPANIES
AND HIGH-IMPACT ENTREPRENEURSHIP

by

Anamaria Berea
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
In Partial fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Computational Social Science

Committee:

Chair of Committee

Date:  Summer Semester 2012
George Mason University
Fairfax, VA
Essays in High-Impact Companies and High-Impact Entrepreneurship

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

By

Anamaria Berea
Doctor
Academy of Economic Studies, Romania, 2010
Master of Arts
Academy of Economic Studies, Romania, 2003
Bachelor of Arts
Academy of Economic Studies, Romania, 2001

Director: Maksim Tsvetovat, Professor
Department of Department of Computational Social Science

Summer Semester 2012
George Mason University
Fairfax, VA
Copyright © 2012 by Anamaria Berea
All Rights Reserved
Dedication

To mom and dad. Especially to my mom, the most extraordinary woman I know.
Acknowledgments

I would like to thank the following people without whom this research endeavor would have been much harder. I would like to thank my adviser, Max Tsvetovat, who has been pushing me to be better. Special thanks to Spencer Tracy and Corporate Research Board, who has provided most of the data used in this research, without which this dissertation would have not been possible. Also special thanks to Rob Axtell, who has suggested this research topic to me and provided very helpful comments on my work in progress and to Richard Wagner, who believed in my potential to become a scholar. I would also like to thank my first adviser, Costea Munteanu, who was a mentor to me and inspired me to do academic research. I would also like to thank my friends and colleagues, Sarah Wise, Dan Maxwell, Charles Twardy, David Kirsch, Matthew Handrey, Tim Gulden and Bill Kennedy, for very helpful and useful comments and more than that, for moral support. And last but not least, to my dear friend Richard Kaback, who, although passed way in 2009, still inspires me to achieve my dreams and stay true to myself.
Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
<td>viii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>vii</td>
</tr>
<tr>
<td>Abstract</td>
<td>xi</td>
</tr>
<tr>
<td>1 Chapter 1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Research Background</td>
<td>2</td>
</tr>
<tr>
<td>1.2 Research Problem</td>
<td>4</td>
</tr>
<tr>
<td>1.3 Goal of Dissertation</td>
<td>6</td>
</tr>
<tr>
<td>1.4 Research Methodology</td>
<td>6</td>
</tr>
<tr>
<td>1.5 Organization of Dissertation</td>
<td>7</td>
</tr>
<tr>
<td>1.6 List of Hypotheses</td>
<td>7</td>
</tr>
<tr>
<td>2 Chapter 2. Theoretical and Qualitative Research in Entrepreneurship</td>
<td>9</td>
</tr>
<tr>
<td>2.1 Entrepreneurship as Social Phenomena Inside and Outside Firms</td>
<td>9</td>
</tr>
<tr>
<td>2.2 Economic Theories of the Firm and Entrepreneurship Theories</td>
<td>10</td>
</tr>
<tr>
<td>2.3 Current Methods in Entrepreneurship Research</td>
<td>11</td>
</tr>
<tr>
<td>2.4 Entrepreneurs as Procedural Actors</td>
<td>13</td>
</tr>
<tr>
<td>2.5 Innovation and Entrepreneurial Skills</td>
<td>18</td>
</tr>
<tr>
<td>2.5.1 Innovation</td>
<td>18</td>
</tr>
<tr>
<td>2.5.2 Entrepreneurial skills</td>
<td>21</td>
</tr>
<tr>
<td>2.6 &quot;Gazelle&quot; vs. High-Growth vs. High-Impact Companies</td>
<td>22</td>
</tr>
<tr>
<td>2.6.1 High-growth companies</td>
<td>23</td>
</tr>
<tr>
<td>2.6.2 &quot;Gazelle&quot; companies</td>
<td>26</td>
</tr>
<tr>
<td>2.6.3 High-Impact Companies</td>
<td>30</td>
</tr>
<tr>
<td>3 Chapter 3. The World of High-Impact Companies: Popular Conjectures</td>
<td>32</td>
</tr>
<tr>
<td>Empirical Facts</td>
<td>32</td>
</tr>
<tr>
<td>3.1 The Growth Quantifiers</td>
<td>32</td>
</tr>
<tr>
<td>3.2 Power law distribution tests in High-Impact Companies</td>
<td>37</td>
</tr>
<tr>
<td>3.3 Geographical Distribution of High-Impact Companies</td>
<td>41</td>
</tr>
<tr>
<td>3.3.1 Statistical tests for geographical distribution</td>
<td>46</td>
</tr>
</tbody>
</table>
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>The High-Growth Companies Ratio in Selected European Countries</td>
<td>25</td>
</tr>
<tr>
<td>3.1</td>
<td>The US industries by SIC codes</td>
<td>51</td>
</tr>
<tr>
<td>4.1</td>
<td>The Parameter Values in the Base Model</td>
<td>74</td>
</tr>
<tr>
<td>4.2</td>
<td>The Variables computed in the model</td>
<td>75</td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>The frequency of the words “Entrepreneurship”, ”Entrepreneur”</td>
</tr>
<tr>
<td>2.1</td>
<td>The Stages of An Entrepreneur</td>
</tr>
<tr>
<td>2.2</td>
<td>The 2 dimensions of growth for a gazelle or high-growth company in OECD methodology.</td>
</tr>
<tr>
<td>2.3</td>
<td>A Comparison Between American and European Gazelles.</td>
</tr>
<tr>
<td>2.4</td>
<td>Accumulation Of High Growth Companies</td>
</tr>
<tr>
<td>2.5</td>
<td>The subset of High-Impact Companies in High Growth Entrepreneurship.</td>
</tr>
<tr>
<td>3.1</td>
<td>The basic statistical measurements for HICs in US, 1995-2006 data</td>
</tr>
<tr>
<td>3.2</td>
<td>EGQ of High-Impact Companies in US, 2007-2010</td>
</tr>
<tr>
<td>3.3</td>
<td>The employment and revenue growth quantifiers in HICs split in 2 periods of 2 years each.</td>
</tr>
<tr>
<td>3.4</td>
<td>The Employment and Revenue Quantifiers for the 326,509 HICs in US, 2007-2010 data</td>
</tr>
<tr>
<td>3.5</td>
<td>Employment Size Distribution in HICs, 2007-2010 data</td>
</tr>
<tr>
<td>3.6</td>
<td>The log-log distribution of HICs on employment size, 1995-2006 data</td>
</tr>
<tr>
<td>3.7</td>
<td>The distribution of HICs on job density and markets in US, 2007-2010 data</td>
</tr>
<tr>
<td>3.8</td>
<td>Density maps of High-Impact Companies in US by location, 2007-2010 data</td>
</tr>
<tr>
<td>3.9</td>
<td>Distribution of HICs both by industry and submarkets, 2007-2010 data</td>
</tr>
<tr>
<td>3.10</td>
<td>The Heat Map of High-Impact Companies, 2007-2010 data</td>
</tr>
<tr>
<td>3.11</td>
<td>The log-log plot of High-Impact Companies distribution on zip codes in US, 1995-2006 data</td>
</tr>
<tr>
<td>3.12</td>
<td>The log-log plot of High-Impact Companies distribution on MSAs in US, 1995-2006 data</td>
</tr>
<tr>
<td>3.13</td>
<td>The log-log plot of High-Impact Companies distribution on county codes in US, 1995-2006 data</td>
</tr>
<tr>
<td>3.14</td>
<td>The distribution of HICs by industry (SIC codes), 2007-2010 data</td>
</tr>
<tr>
<td>3.15</td>
<td>The distribution of all US companies on industry, NAICS codes, 2009 data</td>
</tr>
</tbody>
</table>
3.16 The dependence of managerial activity on other occupations in HICs (2007-2010 dataset) ............................................................... 53
3.17 The distribution of different activities in HICs. (2007-2010 dataset) ........ 54
3.18 The distribution of different activities in HICs. (2007-2010 dataset) ........ 55
3.19 The correlation between founding year and high-impact start year in HICs, 1995-2006 data ................................................................. 56
3.20 The Geographical Distribution of Super High-Impact Companies in US, 1995-2006 data ................................................................. 57
3.21 The Pearson correlation between employment size and zip codes or FIPS county codes, 1995-2006 data ........................................... 58
3.22 The correlation between EGQ and RGQ in super High-Impact Companies, 1995-2006 data ................................................................. 59
3.23 Correlations between super high impact companies size and industry type, 1995-2006 data ................................................................. 61
3.24 The market entrepreneurship indicator for HICs, 2007-2010 data .......... 63
3.25 The distribution of HICs in frequency and percentages on different job dynamics, 2007-2010 data ................................................. 64
4.1 The setup of the model of the growth of a company based on team formation and employee skills ............................................................ 69
4.2 Team formation in the ABM ............................................................ 71
4.3 The pseudo-code for the initialization of the model ............................. 75
4.4 The formation of teams in the model ................................................ 77
4.5 The assignment of the managers of the teams .................................... 77
4.6 The hiring of the employees ............................................................ 78
4.7 The firing of the employees ............................................................ 79
4.8 The company works by completing the task and paying the wages .......... 79
4.9 Company internal accounting ........................................................ 82
4.10 A Typical Run of the Base Model .................................................. 83
4.11 Other typical runs of the base case of the model ......................... 84
4.12 The simulation of a company that grows or not ................................ 84
4.13 A simulation with 3 founders or high learning rate .......................... 85
4.14 A typical run with maximum values of lw and cc and high number of founders .......................................................... 86
4.15 Scaling (employment growth stages) in Internet Start-ups, 2011 .......... 92
4.16 The output in employment and revenue of one company the simulation that shows high-impact phase ................................................. 93
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.17</td>
<td>The employment change in 100 time ticks for 100 computer generated companies.</td>
<td>94</td>
</tr>
<tr>
<td>4.18</td>
<td>Sensitivity analysis</td>
<td>96</td>
</tr>
<tr>
<td>4.19</td>
<td>Regression coefficients of the maximum employment/tick on learning, founders size and failure</td>
<td>97</td>
</tr>
<tr>
<td>4.20</td>
<td>The number of founders in the simulation</td>
<td>99</td>
</tr>
<tr>
<td>4.21</td>
<td>The log-log plot of the frequency of generated companies and employment size in the agent-based model</td>
<td>100</td>
</tr>
<tr>
<td>4.22</td>
<td>The scatterplots of learning, failed projects and maximum changes in employment in 4000 generated companies</td>
<td>102</td>
</tr>
<tr>
<td>4.23</td>
<td>The box plot and heat map of the number of founders, learning parameter and failure rates</td>
<td>104</td>
</tr>
<tr>
<td>4.24</td>
<td>The correlation coefficients between learning, failure and maximum employment changes</td>
<td>105</td>
</tr>
<tr>
<td>4.25</td>
<td>Statistical analysis of employment spikes as dependent on learning and failure</td>
<td>106</td>
</tr>
<tr>
<td>5.1</td>
<td>The Leadership for Learning Framework</td>
<td>108</td>
</tr>
</tbody>
</table>
Abstract

ESSAYS IN HIGH-IMPACT COMPANIES AND HIGH-IMPACT ENTREPRENEURSHIP
Anamaria Berea, PhD
George Mason University, 2012
Dissertation Director: Maksim Tsvetovat

Evidence shows that only a small number of entrepreneurial endeavors - high-impact companies - create most of the new employment in the US. This research is looking for the causes of emergence and uses computational methods to analyze specific aspects of these companies. First, this research proves or disproves some "popular conjectures" regarding the age, location, industry and entrepreneurial character of high-impact companies. Secondly, the agent-based model shows how a company grows in employment and revenue based on two layers of organization: 1) one is the heterogeneous team formation and interaction at the mezzo level of a company organization and 2) another one is the heterogeneous employees skills and interaction at the individual level of a company organization. The model advances and tests the hypothesis that companies that learn the "fastest" from failed projects while retaining access to capital are more likely to become high-impact companies. The experiments replicate the high-impact rate given by the real life data and show that the high-impact phase of a company growth is achieved for specific learning parameters and failed projects. The purpose is to provide a coherent theory-model-evidence analysis on high-impact entrepreneurship, that adds new insights for researchers, policy makers and business practitioners, in addition to the qualitative information, informal knowledge or hands-on experience that they currently posses.
Private companies and business organizations are fundamental for economic growth. Entrepreneurship is a highly interdisciplinary area of research in social science, that studies the formation of these private business organizations from individual entrepreneurial actions. The entrepreneur is a microeconomic actor, yet there are differences between the economic theory of the firm and entrepreneurship theory.

High-Impact Companies are highly adaptable firms that grow very fast in employment and revenue and are responsible for most new job creation in an economy. They are usually their industry leaders, can be found in all industries and all over the territory of US, can have different ages and seem to be robust to macroeconomic cycles. They also exhibit a low concentration of managerial activity and are usually created by 2-3 founders. These business “outliers” represent only 2-4% of all the firms in US and are not necessarily start-ups. Technically, high-impact entrepreneurship is a phase of growth in a business organization that is defined by specific revenue and employment growth during a certain period of time.

Researchers from different fields have defined entrepreneurship as a phenomenon of emergence; however, there is still an unclear definition of what constitutes an entrepreneur or why some firms grow and others don’t. By shifting the focus of entrepreneurship towards the processes and interactions between founders, employees and teams in a complex environment, this research aims to discover bottom-up and top-down factors that allocate entrepreneurial resources towards firm growth. These investigations are based on an agent-based model where the agents are employees endowed with heterogeneous skills and incentives that ultimately lead to the firm’s growth or decline.

High-Impact Entrepreneurship (HIE) is a particular class of entrepreneurship [1] that is distinguishable from lifestyle, social or other forms of entrepreneurship. HIE involves the growth of a business, not merely its creation. Acs notices that the difference between

1
HIE and any other forms of entrepreneurship lays in “leverage” - being a growth business and being involved in disruptive innovation that shifts the wealth creation curve [1]. This dissertation analyzes the specific causes of those businesses that gain leverage and achieve the high-impact phase of growth. Therefore, the purpose of this dissertation is to ultimately provide qualitative and quantitative insights into entrepreneurship theory.

1.1 Research Background

There are still many inconsistencies in definitions and a lack of consensus in understanding entrepreneurship conceptually. Entrepreneurship is a fundamental social phenomenon that can be found anywhere around the world and any time in the history of human civilization, but there is a lack of theoretical and empirical work that would improve its understanding and formalization in a robust\(^1\) way. Researchers from different fields of study (management, business, organization studies, economics [3–7] have defined entrepreneurship as a phenomenon of emergence; however, most entrepreneurship research has focused on questions regarding new ventures characteristics and outcomes after a new venture is started [7].

In a foundational paper, Baumol analyzes the implications of different incentives for the allocation of entrepreneurial resources towards wealth creation or wealth depletion; he categorizes them as productive, unproductive and destructive [3]. Therefore entrepreneurship in general is not only about economic growth, only productive entrepreneurship leads to economic growth.

Entrepreneurship, in its most general sense, can take numerous forms and is an omnipersistent phenomenon. But only very few entrepreneurial endeavors create economic growth and, in the general understanding, entrepreneurship is associated with economic growth. But this is an unclear conceptualization of entrepreneurship. Even in the academic literature, the definitional inconsistencies of “entrepreneurship” and “entrepreneur” reflect the incomplete picture that we have today in understanding this phenomenon and its actors.

\(^1\)“Robustness” refers to the property of complex adaptive systems to persist under perturbations. [2]
The specific language and metaphors used to describe productive entrepreneurship shows the "tacit" or contextual knowledge people have about this phenomenon, overall. Initially, the descriptions were mostly borrowed from biology: "gazelle", "elephants", "mice", "gorillas", "sharks", and many more. Lately, the language spectrum moved from concepts associated with evolutionary competition to concepts associated with social aspects of cooperation: "lifestyle", "enviropreneur", "social" entrepreneur, "next door" entrepreneur, and so on. Why is this important? Although a linguistic analysis is not a part of the current research, I believe that the language used in order to describe entrepreneurial activity shows how strongly entrepreneurship is embedded in the current issues of social life and how the understanding of the general public regarding entrepreneurship has evolved in time. In other words, it is a phenomenon that reflects some trends in social norms. On another hand, these concepts or metaphors through association are a depiction of the lack of formalization or clear understanding of the entrepreneurship phenomenon.

For example, Google Ngram Viewer shows the presence of the words “entrepreneur”, “entrepreneurship” and “business” in printed press throughout the last 300 years [8] and we can get an idea about the evolution of these concepts (see Figure 1.1).

The word ”entrepreneur”, originally from French, has become more predominant in the English language in the 20th century, when the modern firm as we know it today has emerged as an organization. The 20th century is also the time when the meaning of the word ”business” has shifted from a personal interest to one that has an organizational flavor. This was also the time when management has become a separate field of study [9].

Generally speaking, the entrepreneur is seen as a catalyst for change in the world of business [10], and the phenomenon of entrepreneurship is thus a “function” of the entrepreneur in its simplest theoretical form. During the past decades though, entrepreneurship has been associated with innovation and technological change. But, in “The Nature of Technology”, Brian Arthur asks an important question with respect to technological innovations: what is the “essence” of technology and innovation? [11] Innovation and entrepreneurship are very closely related, but what is the “essence” that distinguishes entrepreneurial action from
other types of human activity that are agents of change? More specifically, what is the “essence” of entrepreneurship that fosters economic change towards wealth creation?

The question of conceptually defining productive entrepreneurship is not an easy one. In this respect, this dissertation narrows down the research problem to a particular form of entrepreneurship - HIE - and its formation.

1.2 Research Problem

In this dissertation, my aim is to provide a coherent theory-model-evidence analysis that will lead to insights into that essence of entrepreneurship that is more likely to bring the type
of structural change that Arthur had described. How do entrepreneurs react to patterns and how do they influence patterns? Previous research points towards “gazelle” companies or "high-growth entrepreneurship” as being that specific type of entrepreneurship that brings most structural change and mostly influences industry patterns.

The research question of simultaneous economic and employment growth at the macroeconomic level is unarguably crucial for policy and institutional makers. The understanding of the microeconomic foundations of this growth through computational methods therefore represents the premise of this research.

There are at least four reasons why academic research should be interested in the high-impact, high-growth entrepreneurship question: to determine their impact over time on job growth and industry change, to create an information rich environment that will stimulate business formation and growth, to assess entrepreneurial value and to better understand the micro foundations of the markets.

Vernon Smith argued in his Nobel Prize Lecture that cultures that have evolved markets have enormously expanded resource specialization: “Emergent arrangements, even if initially constructivist in form, must have survival properties that take account of opportunity costs and environmental challenges invisible to our modeling efforts. (...) It is in private information environments where the market is aggregating information far beyond the reach of what each individual knows and is able to comprehend.”

This dissertation shows the value of initial constructivist arrangements (teams) that take account of invisible opportunities (exogenous tasks) and survive or not their failed projects (through learning and access to capital) in order to achieve the success and growth of a company.

Entrepreneurial outcomes can therefore be viewed as (at least partially) dependent on the interactions of entrepreneurial individuals with each other, and with the external world (the environment) regarding sources of opportunity. This interaction framework is a departure from more mainstream theories of the firm that focus on the traits and behaviors of individuals, and from network theories that examine the incidence and location of inter-firm
networks. These theories do not take into account whether contacts are oriented toward the exploitation of opportunity that is bounded by the geographic or social space. Equally important for this approach is the development of skills among local entrepreneurs (or aspiring entrepreneurs) in order to identify and exploit those opportunities [16].

1.3 Goal of Dissertation

This dissertation proposes two things: a new methodology for studying entrepreneurship phenomena and an explanation for the emergence of high-impact/high-growth entrepreneurship in the US. The core of the research is focused on two main directions: 1. a quantitative analysis of empirical facts that dismiss or sustain a series of hypotheses (or "myths" [16] or "popular conjectures") that have been circulating as "common sense" or business advice in the high-impact/high-growth entrepreneurship and business literature; and 2. a computational method that advances and tests the constructive failure hypothesis with respect to the emergence of High-Impact Companies. This constructive failure hypothesis asserts that the emergence of High-Impact Companies depends on how companies are learning from failed projects without interrupting their business process.

1.4 Research Methodology

This research proposes alternative methods to the classic qualitative surveys or econometric methods used in entrepreneurship research. Specifically, it uses computational methods for: 1. analyzing large data and 2. finding and testing causal explanations. These methods are geo-information systems (GIS) analysis, statistical analysis and agent-based modeling (ABM). In this research, particularly ABM has the advantage of overcoming the gaps in the current literature with respect to explaining high-impact growth. It adds insights into theory and policy where the research cannot be done analytically (such as the field of entrepreneurship, which heavily rests on qualitative research), where there are distributional impacts (such as employment growth), where heterogeneity matters to the problem (such
as employees incentives and skills or the undertaken projects), where there is imperfect information, fractal structures and dynamics of networks (such as team formation or organizational learning), where evolutions show different results from comparative statics alone (entrepreneurship by definition is an evolutionary phenomenon), when there are two or more systems that need comparisons and are out of equilibrium (such as companies), where the research is looking for causation and not correlation and where people react to change (such as learning or working in a team). All these aspects of ABM methodology are employed in Chapter 4.

1.5 Organization of Dissertation

This dissertation is organized in 5 chapters. The first chapter is an introduction into the research topic, research background, research goal and methodology. The second chapter is a survey of the literature on entrepreneurship and the broader developments of the theories of the firm, with an emphasis on the qualitative nature of entrepreneurship research. The third chapter introduces a quantitative analysis of empirical data with respect to High-Impact Companies that proves or disproves a series of popular conjectures that the qualitative research in entrepreneurship has advanced. The fourth chapter proposes an agent based model of company growth as an alternative method for explaining the emergence of High-Impact Companies and tests the main hypothesis through a series of experiments. The fifth chapter draws some theoretical implications derived from the computational methods and describes the advances and limitations of the current research, while enumerating future possible research to explore. The dissertation ends with the main conclusions from this research.

1.6 List of Hypotheses

These are the hypotheses that I am testing in this research:
Based on previous academic research:

1. High-Impact Companies are power law distributed by employment size.

Based on popular conjectures:

2. High-Impact Companies are located only in a few specific locations on the territory of US, such as Silicon Valley, major urban areas or the East Coast.

3. High-Impact Companies are activating only in high-technology industries.

4. High-Impact Companies employ only highly educated people.

Main Hypothesis:

5. The high-impact phase of growth of a company depends on both the organizational learning and rate of failed projects simultaneously.
Chapter 2: Theoretical and Qualitative Research in Entrepreneurship

2.1 Entrepreneurship as Social Phenomena Inside and Outside Firms

Entrepreneurs are more than often equated with small business owners [10], but there can be a difference between entrepreneurs and small business owners. While typical small business owners are more focused on running and managing their firms, the entrepreneurial action implies a focus on innovation, profitability and growth (see in section 2.4 below). The difference between the typical small business owner and the entrepreneur comes from their view on the development of their firms: the former is more strategic and aims towards sustainable growth, while the second is focused more on rapid growth and profitability [10].

There is also a notable difference between the entrepreneurial mindset in established organizations and the development of entrepreneurship as a phenomenon. The latter is concerned with the birth of new companies, although the creation of new business does not complete the picture. If the entrepreneurial mindset can be exhibited in business and non-business ventures, in for profit and not-for-profit companies, in individuals outside and inside organizations, with the purpose of acting on whatever creative ideas they might have, the phenomenon of entrepreneurship is a combination of both: integrating the entrepreneurial mindset with the creation of a business.

Sobel makes two important points about entrepreneurship [6]: that entrepreneurial ideas are more important than funding (i.e. venture capital or angel funding) and that there is not an accurate definition of what constitutes an entrepreneur. Whether s/he is a single
person or an organization and whether it refers to a start-up or to revitalizing an already existing organization, entrepreneurship is always about perceiving and undertaking a business opportunity\textsuperscript{1}. Whether s/he is a single individual or a group, the entrepreneur is ultimately the acting agent for creating business ventures on the market. These will afterwards prove to be successful or unsuccessful, having a short-life or long-life.

2.2 Economic Theories of the Firm and Entrepreneurship Theories

The economic theories of the firm and entrepreneurship theories are different. Until the ‘50s, all the references and research of entrepreneurship came from economics. In mainstream economics, entrepreneurs have been placed either on the supply side of the economy, as producers, on the demand side of the economy, as distributors, or replaced by perfectly informed auctioneers [17]. The first economic theories of entrepreneurship are centered around the microeconomic actor that bears the risk [18] and faces uncertainty [19]. In neoclassical microeconomics, the firm is a production function or production possibilities set. The Marshallian supply-demand and Walrasian general equilibrium frameworks do not take into account the role of the entrepreneur from the market. In the contractual theory of the firm, the role of the entrepreneur is one that internalizes transaction costs and coordination problems both from the market (external) and internal processes (i.e. incentives and information flows) [20]. But there still remains a theoretical debate on entrepreneurs as causal [21] or consequential [22] actors with respect to market disequilibria and their role for the markets. Later, Williamson laid the foundations for an organizational theory of the firm, where the problem of the boundaries of the firm and assets ownership is the task of the entrepreneur [23].

With the development of multi-agent view of the firm and agent-based modeling, the theory of the firm has made some substantial contributions towards a robust understanding.

\textsuperscript{1}In this dissertation, I use the terms “opportunity” and “niche” synonymously.
Axtell showed that the distribution of US companies by size of employment follows a power
law distribution [24].

An entrepreneur seeks profits through the voluntary and non-coercive exchange of prop-
erty rights [22], but this definition poses problems under the circumstances where property
rights are not well defined [25]. Entrepreneurial alertness is the innate ability of the en-
trepreneurs to coordinate their own future expectations and present actions [22], in the
context of the market process and their network of competitive entrepreneurs. The en-
trepreneurs are able to do this micro-level, internal type of coordination by becoming cre-
ators of relevant knowledge for their actions. Schumpeter’s [21] entrepreneur is an innovator,
while Kirzner’s entrepreneur is an arbitrageur; if the first actor is a creator at the ”core, the
second one is a discoverer at the margin. Schumpeter’s entrepreneur pulls the economy out
of equilibrium, Kirzner’s one pushes the economy towards equilibrium, yet neither remain
in equilibrium. Kirzner describes alertness as the fundamental quality of the entrepreneur,
while for Schumpeter it is innovation.

2.3 Current Methods in Entrepreneurship Research

One reason there is still a gap in entrepreneurship research concerning origins and emer-
gence and their formalization stems not only from theoretical ambiguities but also from the
difficulty in obtaining and tracking relevant data [7] with respect to entrepreneurial skills
and motivations.

Many scholars see an apparent contradiction between Schumpeter and Kirzner on equi-
librium [6]. But treating all entrepreneurs as one type or the other of homogeneous mi-
croeconomic actors can lead to significant theoretical and modeling departures from reality.
Fortunato and Alter [26] provide an integrated framework for entrepreneurship research,
called the entrepreneurs - institutions - opportunities nexus. Yet their framework lacks
formalization at the individual heterogeneity and phenomenological evolution levels as well,
treating the nexus as separate, static, bi-modal networks.
Entrepreneurship is highly correlated with the Economic Freedom Index [6], yet it is still unclear if entrepreneurship causes economic growth or vice-versa. Testing for economic causality using econometric methods is very difficult. Although many of the empirical studies on entrepreneurship found correlations between entrepreneurship and economic growth in different industries, there is still an ongoing debate on causal factors. In an analysis on the "gazelles" and structural industrial change in Netherlands [27], the authors use the Granger test in a PVAR econometric model without being able to fully support the hypothesis that gazelles cause economic growth and structural sectoral change in the 42 industries over the 10 years span of their dataset.

Entrepreneurship is endogenous to the institutional context within which it emerges - the set of conditions in which entrepreneurial action is conducted has effects on entrepreneurship itself [3]. It is an "order creation" process that has both bottom-up and top-down causes. Therefore, using analogies or methods from evolutionary biology are not the most suitable ones either. Apart from biological systems, who adapt slowly through mutation, entrepreneurship is a phenomenon of rapid changes in the context of firms [28]. No matter their type, the entrepreneurs are people willing to invest their current resources in return for future benefits. They are motivated by combinations of several things [29] and therefore their incentive are subjective and heterogeneous.

Everything cannot interact with everything else at once, and therefore entrepreneurs are more like "surfers waiting for the big wave" [29], meaning that they cannot control the events and structures, but they can bend them to their purposes, to a certain extent. In his analysis of policy entrepreneurs, Kingdon also brings an important contribution to the understanding of gradual evolution (incrementalism) and punctuated equilibria (sudden changes) for the phenomenon of entrepreneurship. Gradual development is one of the reasons why entrepreneurs work on their plans for a longer period of time and do not invent products instantaneously - this is how they are able to take on an opportunity as soon as it arises; they recombine familiar elements into new ventures.

Runde uses the concepts of morphostasis and morphogenesis and Searle's concept of
agentive functions to point towards aspects of technological change that act as breaks or accelerators of structural change [30]. This is important for what he calls "user-driven innovations", that can take three types of change.

Whereas the academic literature on entrepreneurship does not make a clear distinction between individuals that start new companies and companies/groups of individuals that spot new opportunities and start new companies, the focus usually falls on the phenomenon of entrepreneurship in a broad sense. The framework laid out by Busenitz [31] is an important one to entrepreneurship research because it removes the research from focusing on the firm as a unit of analysis, and shifts entrepreneurship research towards the processes and interactions of multiple stakeholders in a complex environment.

2.4 Entrepreneurs as Procedural Actors

Simon states that, in the real world, the firm has to choose the course of their entrepreneurial action between the "right course" and the "good course", that can only be approximated [28]. The real economic actor is a “satisficer”, not a maximizer.

"With this shift, the theory of the firm becomes a theory of estimation under uncertainty and a theory of computation." [28]

Moreover, he continues, the organization-market boundary is movable. In the face of uncertainty, the entrepreneur uses hierarchy rather than markets in making decisions. Entrepreneurial action lies at the origin of any social form of organization. But an evolutionary framework for studying entrepreneurship is not the most suitable one, since Darwinian evolution is myopic [28]. Entrepreneurs need to exhibit more foresight and fore-thinking in order to spot business opportunities or in order to innovate. The fitness test is on the profitability and growth rate of the firm [32]. Also apart from biological systems, in entrepreneurship one individual can “copy” the successful algorithm of another, and therefore “successful mutations can be transferred between firms” [28]. These transfers are not costless, they imply a "cost of learning". Learning is any change in the system that produces a
more or less permanent change in its capacity for adapting to the environment.

Every year, more people start businesses than get married [16]. Therefore an important question is: who is the entrepreneur?

![Diagram of the stages of an entrepreneur]

Figure 2.1: The Stages of An Entrepreneur.

The environment of the entrepreneur must have 2 properties: 1) morphostatic and 2) morphogenetic [30]. Morphostasis represents the structural stability or organizational form that a company or an industry exhibits with respect to rapid change. This property can be coded as a cost for labor growth or sales growth (the cost of entrepreneurial action) or as a lag in adaptation and coordination (such as a “menu cost” [33]). On another hand, morphogenesis allows development as a consequence of entrepreneurial action, therefore it can be represented as an incentive for employment growth either at the policy level as a parameter for the environment, or at the agent level as an expectation for sales growth.

In “The Illusions of Entrepreneurship“, Scott Shane portrays the American entrepreneur based on several statistical findings [16]. In his research, he is bringing evidence against
some of the common popular conjectures and popular beliefs about what constitutes an entrepreneur. According to Shane, a smaller percentage of population starts businesses today than they did in 1910 in US. *(i.e. US households that own a business decreased from 14.2% (1983) to 11.5% (2004)).*

Entrepreneurs are also more likely to start businesses in less attractive industries, like construction and retail and the most common reason to start a business is to avoid working for others. People who change jobs often, who are unemployed and who make less money or people who make a lot of money are more likely to start their own businesses [1, 16].

The American typical entrepreneur is a white man in his 40s, married with a working spouse, attended college but did not graduate, was born in US and has lived here his entire life, has spent much of his life in the town where he started his business, is just trying to make a living, not trying to build a high-growth company, worked previously in the industry in which he started his company and has no special psychological characteristics [16]. Entrepreneurs dont have a different psychological makeup than the rest of us. An owner-managed firm has a typical revenue of $90,000, while for the self-employed is $183,973 and unemployed turned entrepreneurs are unlikely to become highly successful.

Most start-ups are not innovative. Even among the best 500 start-ups, only 10% offer a product or service that other companies do not offer. Most new businesses do not intend to do something innovative enough to alter the market they are in. Only 2% of new business founders expect their companies to have a substantive effect on the markets in which they operate [34]. Also, only 16.9% of new start-ups employ more than 1 person. The average firm that has any employees has approximately 3.8 of them [16].

Unlike the high-growth companies, the typical start-up is not innovative, has no plans to grow, has one employee and generates at best $100,000 in revenue. Only one third of start-ups survive and operate their business routinely for the next 7 years since founding and they are usually capitalized with $25,000, taken primarily from the founder savings [16]. The "typical" entrepreneur ends up working more hours but earning less than he would have earned had he worked for someone else. Start-ups create fewer jobs than it is
usually understood in the business circles. Only 1% of people work in companies that are less than 2 years old and 60% of people work in companies that are more than 10 years old.

In fact, many of the High-Impact Companies are not start-ups [13].

A study of a representative sample of the founders of new businesses started in 1998 showed that 81% of them had no desire to grow their businesses [35]. Almost 2/3 of firm founders do not expect their new companies to generate more than 2 jobs within the first 5 years [1]. Also, most of them expect to have sales of less than $100,000 within their fifth year of operation. This leads to the following question: Why do some firms perform better than others? A Kauffman firm survey showed that 63% of new firms report they have some sort of competitive advantage [36]. Acs names this competitive advantage ”leverage” [1].

Starting a new business is a process, not a one time event. Writing a business plan actually facilitates and increases the odds of company survival [16]. Approximately1/5 entrepreneurs are involved in the start-up process forever, they don't abandon the effort, but they never complete it either and 53% of the new businesses with multiple founders are started by spouses [16]. Only 10% of new businesses are founded by teams with non relatives. Moreover, 2/3 of team ventures are started by a single founding entrepreneur and the others join later.

New firms are financed by debt or equity in equal proportions. Venture capitalists make investments in approximately 3000 companies a year, out of which only 500 are start-ups. They account for only 1.9% of total small business financing, ad 92% of it goes into healthcare and information technology. Twice of that is provided by angel investors. These are only a part of informal investors. Approximately 1% of US households holds an ownership stake in a private business they do not manage. More than half of these investments are less than $15,000. Only 15% of informal investors invest as part of a group. The typical informal investor invests in a business without expecting financial gain, mostly to help out a friend.

In conclusion, what makes some entrepreneurs more successful than others? According to Shane, industry seems to play an important role: 17% difference in survival rates of 4
years across industries and you are more successful in industries where you can get a barrier to imitation and industries that are technologically intensive [16]. Yet Acs et. al showed that, in the case of High-Impact Companies, industry is not a determinant factor [1]. But Shane and Acs agree on the following traits as being important for entrepreneurial success:

- those who start larger start-ups, with access to capital
- those who capitalize their business with over $100,000 through equity
- those who set up corporations and not sole-proprietorships
- those who work full time on their business, not part time
- those who start in teams
- those who write business plans
- those who seek customers missed by others
- those who sell to businesses and not to individual consumers
- those who start marketing plans sooner
- those who do not compete on price, but on other dimensions (such as quality)
- those who start by focusing on one product or service
- those who organize their business managerially
- those who are college educated (25% more chances of survival)

Entrepreneurs who start companies are also unlikely to get back to employment [16]. Therefore previous entrepreneurs have a higher preference for becoming entrepreneurs again than staying employed and this trait is modeled as learning and transmission of experience from one task to another in the ABM in Chapter 4. Many successful entrepreneurs also show some personal relationships with other people that have entrepreneurial behavior and
the role mentorship plays for the new entrepreneurs [16]. The transmission of learning, either formal or of tacit knowledge, is parametrized in the model and will prove to play an important role for the success of companies. A majority of all the characteristics enumerated above are modeled into the attributes of the agents in Chapter 4, section 4.2.

2.5 Innovation and Entrepreneurial Skills

Until now, entrepreneurship theories have focused on entrepreneur as a major innovator or technological disruptor. However, the facts point towards innovation as being only a secondary factor for the emergence of high-growth entrepreneurship. Similarly, the entrepreneurial psychological skills are only secondary to the emergence of high growth companies.

2.5.1 Innovation

Nordhaus defines Schumpeterian profits as profits from entrepreneurial endeavors who innovate [37]. Specifically, Schumpeterian profits represent a normal return on investment in innovation and have a high depreciation rate. Nordhaus calculated the Schumpeterian profit margin as the ratio of the Schumpeterian profits and total revenues [37]. His results imply that innovators capture 2.2% of total surplus from innovation. The low figure is a result of the easiness of entry-exit on the market and that imitators can follow quickly breakthrough innovators. Technological change can have two types of impact on profits: if knowledge is public, it lowers the prices of products, but if knowledge is private, it leads to Schumpeterian profits. The depreciation rate of Schumpeterian profits comes from that property of information as a good that is expensive to produce, but inexpensive to reproduce. Generally, only 20% of the rate of return on capital is attributable to Schumpeterian profits.

Entrepreneurs are not only founders of firms, they are knowledge seekers in the sense that, as innovators, they are constantly receiving and sending information to the world. The tension in entrepreneurial activity comes when they are seeking profit from their activity of
innovation and information exchange. This points to the question of how do entrepreneurs bring innovations successfully to the market. Prices, in the Hayekian sense, are necessary, but not sufficient [38]. They are surrogate proxies, for many things that are underlying that, such as the production maps and plans that change in the face of the new information from the market.

Policy entrepreneurs bend problems to their solutions [29]. As any entrepreneur, they are looking for opportunities, read the windows of opportunity well and move at the right moments. The tension comes when the windows of opportunity in policy solutions are unpredictable.

The idea of user driven innovations from Runde shows that technological objects have a function and a form and a user can change either the function, either the form or both, thus spurring user-driven innovations. [30] Morphostatic processes play a role in technological transitions: they retain structure, organization and form. Therefore they are characterized by a state of structural stability and efforts to retain the status-quo. On another hand, morphogenesis represents the development of the structure, it is adaptive in response to the changing environment and can be gradual or not. Both morphostatic and morphogenetic social processes act as breaks or accelerators of technological change.

“Technological surprise” or breakthrough innovations are more likely to be found by bringing together pieces of information that are conceptually more distant one from another [39]. This is closer in meaning to the concept of ”creative destruction” [21] than to micro-inventions.

Mokyr distinguishes between micro and macro-inventions [40]. Micro-inventions are considered gradual improvements upon an existing technological basis and they are subject to diminishing marginal returns. Macro-inventions introduce new concepts, appear in clusters and defy diminishing marginal returns in Schumpeterian fashion [21]. This research offers a novel approach that explains temporal clustering of business process innovation. The agent-based model in Chapter 4 will presents some causal factors for business process innovation that, at its core, it is a model of entrepreneurial emergence.
This research does not consider entrepreneurial discovery or innovation to be purely random. The systematic component of entrepreneurial discovery is linked to the capacity of individuals to spot them, which Israel Kirzner [22] has termed as ”alertness”. Alertness to profit opportunities is not randomly distributed across individuals and time. Time is a critical factor for this research. For example, a 19th century innovator such as Edison for example, was not able to conceive of the iPod because 1. the scientific basis did not exist and 2. because he did not grow up within what has been called the ”socio-technological context” [41] preceding the discovery of iPod. Even if Edison had possessed all the relevant scientific information (hard facts) he would still miss the relevant cultural knowledge (soft facts) that would predispose him to make the discovery. Soft facts or ”tacit knowledge” [42], consist of knowledge that cannot be easily captured in mathematical formulations, it is hard to communicate and specific to certain individuals and their particular context. Many forms of cultural norms are examples of tacit knowledge and will be immediately apparent to anybody who has spent a longer duration abroad.

The analysis of innovation tends to be preoccupied with supply side explanations of innovation [43]. However, “production and consumption phenomena are an intertwined network [44]. Entrepreneurial innovators rely on cues that they take from the consumer culture they are familiar with. Geels introduces the term socio-technological systems to incorporate both the supply side and the demand side factors [43].

The concept pertains to the linkages between the technological regime, the user and market regime, the socio-cultural regime, the science regime, and the policy regime. Changes in one regime, such as the socio-cultural one can affect another regime, e.g. the direction of technological innovation.

If Schumpeterian entrepreneurial discovery/innovation were purely randomly distributed the time span between a scientific breakthrough enabling a certain product (e.g. iPod) and its actual creation would be random as well. Treating the discovery process as a random event neglects the entrepreneurial process in a behaviorist fashion, meaning that it does not explain it [45]. Entrepreneurs, seen as ”lucky fools”, somehow spot the connection between
technological feasibility and potential demand. Without assuming this mystical ability, a random discovery process would entail bizarre consequences as there is an infinite amount of products and services that can be created with a given set of inputs. However, only a finite number of projects are taken on by entrepreneurs. It becomes clear that individuals, i.e. entrepreneurs, subjectively decide about the feasibility of these projects based on their subjective knowledge and judgment (soft facts). To the extent that this decision process contains systematic components, social scientists should seek to understand them instead of assuming a mystical black box.

2.5.2 Entrepreneurial skills

Jerry Kaplan, a serial entrepreneur, made a list of skills that are more likely to be found in successful entrepreneurs [46] and explained them in a series of video lectures for Stanford University. Most of these skills are related to communication with teams and team formation abilities: ability to build consensus, ability to keep a clear and consistent message, fast decision-making, knowing when to trust and when to delegate, ability to telescope: focus in on the details and moving back from focus to the bigger picture. Also, he grouped the external skills that are more conducive to entrepreneurial actions: exceptionally good at sales, take on huge risks and resist stress. Another important finding of his is that more experience is learned from mistakes [46].

For example, Sarasvaty attempts to explain the entrepreneurial way of finding and creating opportunities and describes it as a certain type of logic, which she calls "effectuation" [47]. The Kirznerian concept of "alertness" ultimately functions as an on/off switch of entrepreneurial skills and entrepreneurship research should focus on the directedness of the entrepreneurial "gaze" [45]. The individual will therefore see different opportunities depending on his predisposition, which in turn is a result of his unique life history. The new venture creation is not as an automatic maximization response to opportunities [47]. The entrepreneur is not a blank slate profit maximizer but brings his or her own personality to the project. In that respect, a passionate cook is more likely to write a cookbook because
his gaze is directed toward this and other related activities [45]. Despite the connection between tacit knowledge [42] and the directedness of the entrepreneurial gaze, it should not be forgotten that unforeseen events beyond the individual control represent important determinants of success for any venture as well.

Some empirical interesting findings show that immigration is too varied in groups to matter for likelihood of startups. Also, being a good networker does not seem to increase the odds of starting a new business [16]. Women are less likely to be successful entrepreneurs because they are not interested in running their own businesses for financial goals but for flexible schedules [16].

As mentioned in the introduction, the fitness and robustness tests are represented by the profitability and growth rate of the firm [32], and not by survival or reproduction. Apart from biological systems, which adapt slowly through mutation, entrepreneurship is a phenomenon of rapid changes in the context of firms. Also apart from biological systems, one entrepreneur can copy the successful algorithm of another, and therefore successful mutations can be transferred between firms” [28]. These transfers are not costless, they imply a "cost of learning". Learning is represented computationally as any change in the system that produces a more or less permanent change in its capacity of adaptation to the environment.

In this dissertation, I am searching for these causes by creating and growing High-Impact Companies in an artificial world (computer simulation). The results can have numerous implications for measuring entrepreneurial value and success both at the micro-economic and at the macro-economic level.

2.6 “Gazelle” vs. High-Growth vs. High-Impact Companies

Business people and entrepreneurs are aware of the existence of the gazelles, high-growth or High-Impact Companies. While not willing to wait for economists to develop theoretical and empirical understanding of this phenomenon, different consultancy companies have
developed qualitative surveys and methodologies in order to understand the causal factors, economic, psychological, financial or political that are more conducive for starting and growing a company [48]. Canada has identified 21 indicators that lead to the growth of a company. Fourteen (14) of these variables are positive leading indicators (i.e. product life cycle status, new product introductions, emerging technologies, investment in R&D, distribution of R&D expenditures (new product, improved product, or improved production), market, sales, planned expansions, export sales growth, regulatory changes, projected employment needs, projected utility service needs, jobs added, and space added. Seven (7) of the variables are negative. They can reduce a company’s growth potential or distract the continuation of business on a positive growth course (i.e. unions, mismatch between a company’s skill requirements and the available workforce, negative regulatory conditions, competition, merger/acquisition activity in the industry, union activity, changes in base technology, jobs lost, and access to specific critical workforce skills).

2.6.1 High-growth companies

At the formal institutional level, the Organization for Economic Cooperation and Development has taken a step towards gathering data with respect to the existence of "gazelles" and "high-growth enterprises" in the member countries. They have also developed a separate methodology for statistical purposes [49]. OECD uses the terms of "gazelle" and "high-growth" synonymously.

OECD defines high-growth enterprises as follows: *All enterprises with average annualized growth greater than 20% per annum, over a three year period should be considered as high-growth enterprises. Growth can be measured by the number of employees or by turnover.[49,50]*

Gazelles are defined as a subset of high-growth firms: *Gazelles are the subset of high-growth enterprises which are up to five years old. The definition is: All enterprises up to 5 years old with average annualised growth greater than 20% per annum, over a three year period, should be considered as gazelles.[49,50]*
OECD recommends employment and turnover as indicators to measure growth (see Figure 2.2) and, more specifically, the rate of high-growth enterprises and the rate of gazelles among newly born enterprises. The first one represents the number of high-growth enterprises as a percentage of the total population of active enterprises with at least t employees and the second one as the number of gazelles as a percentage of all active enterprises with at least t employees that were born four or five years ago.

![Diagram](image)

Figure 2.2: The 2 dimensions of growth for a gazelle or high-growth company in OECD methodology.

They provide this data aggregated at the country level for 2005-2007 for the following industries: construction, finance, health, horeca, manufacturing, mining, real estate, sales and transportation. As indicator, they provide the rate of high-growth enterprises only for
a few countries, for 2006 or 2007 (see Table 2.1). Interestingly, this rate varies greatly between countries that are part of the European Union, which bears the question whether the emergence of these companies is caused or influenced at all by European level policies and regulations or any external factors. By looking comparatively at the rates in countries like Slovakia and Czech Republic or Latvia and Estonia, countries that share similar business practices and shared similar economic histories for a large amount of time, we can form an intuition that the external factors might not play a very important role.

Table 2.1: The High-Growth Companies Ratio in Selected European Countries. Source: Eurostat - OECD, retrieved August 2011

<table>
<thead>
<tr>
<th>Country/Time</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Czech Republic</td>
<td>0</td>
<td>5.93</td>
</tr>
<tr>
<td>Denmark</td>
<td>0</td>
<td>5.81</td>
</tr>
<tr>
<td>Estonia</td>
<td>0</td>
<td>7.12</td>
</tr>
<tr>
<td>Spain</td>
<td>0</td>
<td>4.62</td>
</tr>
<tr>
<td>Italy</td>
<td>0</td>
<td>4.56</td>
</tr>
<tr>
<td>Latvia</td>
<td>6.84</td>
<td>0.72</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>0</td>
<td>5.8</td>
</tr>
<tr>
<td>Hungary</td>
<td>0</td>
<td>5.99</td>
</tr>
<tr>
<td>Netherlands</td>
<td>4.82</td>
<td>5.84</td>
</tr>
<tr>
<td>Portugal</td>
<td>0</td>
<td>2.91</td>
</tr>
<tr>
<td>Romania</td>
<td>0</td>
<td>1.38</td>
</tr>
<tr>
<td>Slovenia</td>
<td>4.13</td>
<td>7.09</td>
</tr>
<tr>
<td>Slovakia</td>
<td>0</td>
<td>0.87</td>
</tr>
</tbody>
</table>

The European Foundation for Entrepreneurship Research concluded that there is a huge gap between the predominance of these companies in Europe versus the US and that this gap is due to business location/cluster proximity and the role of universities in teaching entrepreneurship [51] (see Figure 2.3).

In the USA, the Kauffman Foundation has been researching or investing in the research of entrepreneurship and business growth for half a century. In one of their latest working
papers, they recommend policy makers to focus on taxation, immigration, access to capital and academic commercialization in order to spur the creation of high-growth companies [36] (see Figure 2.4).

2.6.2 ”Gazelle” companies

Other authors also distinguish between high-growth companies, gazelle companies and High-Impact Companies. These definitions are similar, they usually differ on how they view: 1. employment; 2. the length of time. (see Figure 2.5)

All these definitions have arisen from the original finding of Birch based on Dun&Bradstreet dataset [14]. The term ”gazelle” company was coined in 1979 by David Birch, in a report he had not published at the time. The report portrayed the relationship between the US
firms and job creation based on the Dun and Bradstreet Market Identifier (DMI) [14]. The "gazelles" are not necessarily start-ups. Birch published another study later in which he used the data set from 1990-1994 in order to assess the new job creation in the US economy [52]. He found that the "gazelles" that start from a base of 100 employees are responsible for 53% of the new jobs created in the economy.

A "gazelle" has to grow at least 20% a year for four years, from a base of at least $100,000 in revenues - in effect, at least doubling in size over that four-year period. That means that roughly only 1 in 16 companies in US may qualify as "gazelles". "Gazelles" can be found not only in high-tech industries, they can be found in any area of business: close to 30% of all "gazelles" are in wholesale and retail trade and another roughly 30% are in services (some technological, some not). Only a small number of "gazelles" are funded through venture capital; only 5,000 out of 350,000 gazelles that were start-ups during the dot-com bubble were funded by venture capitalists [53]. They operate mostly in local markets, while only
Figure 2.5: The subset of High-Impact Companies in High Growth Entrepreneurship. This picture shows how different forms of entrepreneurship relate to each other definitionally. High-Impact Companies are the smallest subset, but they can occasionally overlap the "gazelle" companies due to different methodologies.

A very important data find that current literature provides is that the "gazelle" phenomenon is not just a matter of finding the right niche. Although necessary, finding the right niche is not sufficient - companies can start in the same industry at the same time, and one will grow fast while the other stagnates or grows very slowly. This implies that the entrepreneurial "skill", that represents the ability to start and grow a business, is probably more important than entrepreneurial "alertness" as previously described, that represents the identifying of an opportunity. A couple of years ago, there was published an agent-based model, where several levels of alertness and skills for the entrepreneur were tested in order to find out which one is more important [54].

On average, there should be 43 new start-ups to create 1 net job ten years later; this represents a "macroeconomic bad yield" [16] and jobs in new firms are not as high-quality or as well-paid as jobs in average firms.
From 1994 to 1998, one of the economic booms in US, there were 288,636 gazelles (which accounted for about 2% of all U.S. companies at the time) that created more than 8 million new jobs, accounting for 68% of all net new jobs at that time [55]. After the dotcom bubble, the "gazelles" were influencing the U.S. economy (creating more than 200% of American jobs from 1998 to 2002) [55]. Of the nearly 21 million businesses operating in the U.S. today, only around 350,000 are "gazelle” companies. These are the 2 percent of businesses that generate on average 80% to 90% of all employment growth. "Gazelle” companies are also industry innovators, in a Schumpeterian way. These are the companies that spot unique market opportunities and move rapidly to exploit them. In the process, they influence their industries by creating entirely new ways of producing their products or services [55].

The typical "gazelle” has around 47 employees, tends to be found in the more active industries of that specific period in time (after 2000, they were in high tech, today they are mostly in petroleum industry) and generate nearly 60% more revenue per employee than the typical entrepreneur [55]. Typically, a gazelle reaches this high efficiency within the first 5 years of its life.

During the past 10 years, business start-ups have approached approximately 600,000/year (US Small Business Administration). New and smaller firms have been responsible for 55% of the innovations in 362 different industries and 95% of all radical innovations. "Gazelles” produce twice as many product innovations per employee as do larger firms. As employment, they generated approximately as many jobs as the entire US economy (10.7 million from "gazelles” out of 11.1 million total) [10].

On the downside, the gazelle entrepreneurial environment is a highly stressful one. The entrepreneurs starting this type of firms have both high and rapid growth as a goal (both in size and in wealth), while other types of firms are generated merely with the purpose of bringing income.

Both Birch and Lesonsky assert that is not necessary that a company starts as a "gazelle” [52,55]. Some companies become "gazelles” later in their life. Clearly, there is a connection between employment and sales for a "gazelle”: data shows that, on average, a "gazelle”
employee brings 60% more revenue than an average company employee [48, 55]. In other words, these are highly efficient companies.

2.6.3 High-Impact Companies

*High-Impact Companies* are a subset of the "gazelle" companies. As the employment growth is the main indicator for a high-impact company, while the revenue growth is secondary and the age of the company is absent, there are usually less High-Impact Companies in US than gazelles.

Compared to the population of “gazelles”, which roughly accounts for 6.25% of the entire population of companies in US in any given period of time, the population of High-Impact Companies is smaller - 1.5% to 4% of the entire population of companies in US [1]. The procentual difference between high-growth firms, gazelles and high-impact firms comes from the different methodologies of defining the gazelles and High-Impact Companies and from their age differences. Similarly to the gazelles, HICs are age independent. Therefore, high-impact or "gazelle" should be regarded more as phases of growth in the evolution of a company than as a company in itself.

They are measured by using two composite indicators, the revenue and employment growth quantifiers [1], which are being discussed section 3.1. While these measurements are classifying High-Impact Companies, high-impact entrepreneurship is the overall social phenomenon that creates High-Impact Companies and the policy question is how can this phenomenon be spurred in any economy.

The "three pillars" of HIE that Acs proposed [1] are the following:

1. Innovation in the products or business structure.
2. Occupational choice.
3. Access to capital.

These general pillars are fostering HIE, but they are not sufficient ingredients for the creation of High-Impact Companies. The transition from the macro-level phenomenon of high-impact entrepreneurship to the micro-level High-Impact Companies is done in this
research by removing the focus from the external factors to the internal ones.
Chapter 3: The World of High-Impact Companies: Popular Conjectures and Empirical Facts

3.1 The Growth Quantifiers

There are many theories and popular conjectures about entrepreneurship. In this chapter, I am testing some of these theories using 2 comprehensive data sets with respect to High-Impact Companies in the US, data provided by Corporate Research Board. In entrepreneurship research, it is important to have data not only about the past, but also about the change. In the case of High-Impact Companies, this role is fulfilled by the employment and revenue growth quantifiers, which exhibit some very interesting patterns with respect to the evolution of High-Impact Companies.

The growth quantifier is a quantitative measurement of growth proposed by Acs et al. in order to differentiate the world of High-Impact Companies from other types of high-growth entrepreneurship [13]. For a long time, there has been a popular conjecture in entrepreneurship literature that only start-ups are the staple of entrepreneurial activity and that only successful start-ups are leading towards significant growth in employment and revenue. The growth quantifier is a measurement designed to include in the pool of High-Impact Companies both very small and very large companies that have significant growth in employment and revenue [1]. The growth quantifier is the product of change in absolute and percentile (relative) terms. The percentile term includes large companies in the pool, while the absolute term includes the start-ups in the pool.

EGQ or the employment growth quantifier is the measurement that distinguishes the subset of high-impact companies from the "gazelles". Originally, Birch defined the “gazelles” based on rates of revenue growth, without taking the employment growth into account [14].
Acs et al. introduced an alternative definition for High-Impact Companies in order to take into account both sales and revenue growth. [1]

The employment growth quantifier is the product of absolute and percent change in employment over a four-year period of time:

\[ \dot{g} = \frac{(\Delta g)^2}{g^2-100}, \]

where \( g \) is the employment and

\[ \Delta g = g_{t+4} - g_t \text{(the difference in growth between the fourth year and first year of growth)}; \]

(in the case of the revenue growth quantifier, RGQ, \( g \) is the revenue; the same methodology of calculating the quantifier applies)

This research is using two comprehensive data sets provided by Corporate Research Board. The first one comprises the evolution of 164 economic and demographic aggregated indicators for the entire population of High-Impact Companies; these were selected as those companies with a revenue growth quantifier of more than 90% out of the 21,167,830 US companies existing in US in 2010. The second data set comprises the evolution of 10 granular indicators for the entire population of High-Impact Companies in the US during 1995-2006 inclusive. The first data set used for this empirical analysis represents the same data Acs et al. used in their analysis of high-impact entrepreneurship for SBA [13], but it is updated to the 2007-2010 period of time. The second data set is the data of all High-Impact Companies in US from 1995-2006, which comprises their geographical location, employment and revenue in the first year and fourth year of analysis and the employment and growth quantifiers. This data shows there had been 951,189 High-Impact Companies in US during that decade (see Figure 3.1).

\(^1\)“Granular” means yearly data in absolute numbers, not aggregated in ranges over a 4 year period of time.
In 2010, there were 326,509 High-Impact Companies in US [56]. According to Acs, the US High-Impact Company rate has changed from 3.8% (1994-1998), to 2.1% (1998-2002), to 2.2% (2002-2006) respectively. For the current period of analysis (2006-2010), the rate is 1.5%. One of the reasons for the decline in the population of High-Impact Companies could be the economic crisis of 2007, yet overall the High-Impact Companies show substantial new job creation in the economy even during this period of time.

Even during the economic crisis of 2007-2010, the employment growth trend shows that most High-Impact Companies have a quantifier of at least 60% (see Figure 3.2). If the overall high-impact rate has declined for the last 4 year of study, the companies that were high-impact still showed significant job creation and employment growth.

Figure 3.3 shows that actually the rapid period of growth for High-Impact Companies can be even shorter than 4 years. Most of the companies that had high growth in employment or revenue in 2007-2008 had low growth in 2009-2010 and vice-versa. Since the plots depict the same population of companies that is characterized by the same growth in revenue and employment for the aggregated period of time 2007-2010, the companies that had zero or no growth in the first period must have had very high growth in the same period; similarly, the companies that had high-growth in the first period, must have waned in the

<table>
<thead>
<tr>
<th>Table: Descriptive Statistics</th>
<th>N</th>
<th>Range</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales growth rate</td>
<td>951199</td>
<td>6.49709</td>
<td>0.00000</td>
<td>6.49709</td>
<td>4.451050</td>
<td>4.451050</td>
<td>14.427555</td>
<td>1.300614</td>
<td>365.4944</td>
</tr>
<tr>
<td>Employment growth rate</td>
<td>951199</td>
<td>0.352403</td>
<td>0.00000</td>
<td>0.352403</td>
<td>0.352403</td>
<td>1.171109</td>
<td>1.324418</td>
<td>31.932526</td>
<td></td>
</tr>
<tr>
<td>Efficiency ratio Year 4</td>
<td>951199</td>
<td>289.066087</td>
<td>19.00000</td>
<td>289.066087</td>
<td>1289401495</td>
<td>102986.95</td>
<td>5285080.07</td>
<td>2784811</td>
<td>241.3013</td>
</tr>
<tr>
<td>Year 1 Sales</td>
<td>951199</td>
<td>969999000</td>
<td>1000.000</td>
<td>969999000</td>
<td>3.123</td>
<td>2630154.99</td>
<td>6.936419</td>
<td>319.0003</td>
<td>152188894</td>
</tr>
<tr>
<td>Year 1 Employment</td>
<td>951199</td>
<td>16000.00</td>
<td>0.00000</td>
<td>16000.00</td>
<td>6314009</td>
<td>6314009</td>
<td>39764490</td>
<td>164.5500</td>
<td>56822234</td>
</tr>
<tr>
<td>Year 4 Sales</td>
<td>951199</td>
<td>2.111</td>
<td>227.000</td>
<td>2.111</td>
<td>9.512</td>
<td>9556241.30</td>
<td>4.01583</td>
<td>2.3191717</td>
<td>216.0700</td>
</tr>
<tr>
<td>Year 4 Employment</td>
<td>951199</td>
<td>1500.00</td>
<td>0.00000</td>
<td>1500.00</td>
<td>15169724</td>
<td>15169724</td>
<td>22623704</td>
<td>97.1790</td>
<td>9443.7986</td>
</tr>
</tbody>
</table>

Figure 3.1: The basic statistical measurements for HICs in US, 1995-2006 data
second period. The data is not annualized, the specific values for the 2 year quantifiers are taken from the aggregated dataset as such.

Figure 3.4 shows how the quantifiers for the same population of companies relate with each other for the entire 4 year period of time. The relation between the EGQ and RGQ shows that most of the companies have more revenue growth than employment growth overall and that they grow at a higher growth rate. Only a few thousand companies grow more in employment than in revenue, but at a smaller growth rate.

The analysis of the quantifiers depicted in the previous 2 figures above shows that high-impact companies can go out of their high-impact phase sooner than the 4 year standard
period given by the original definition of the "gazelle" and that their employees start to bring more revenue per employee, thus becoming more productive and more efficient as they exit their high-impact phase of growth.

In 2010, there were 66 high-impact companies with sizes in the 20,000 - 360,000 range of employment and 278 companies in the $1-123 billion range of revenue. These high-impact companies were obviously super-performers that are not small or medium size companies. These "outliers" account for roughly the following numbers in job creation: 22,000 jobs created by a super-performant high-impact company compared to only 7 jobs per company created in the overall US economy in 2010. Similarly, the revenue generated by the most prolific High-Impact Companies is of $62 billion/company compared to $0.7 billion/company.
Figure 3.4: The Employment and Revenue Quantifiers for the 326,509 HICs in US, 2007-2010 data

3.2 Power law distribution tests in High-Impact Companies

Previous academic research showed that the distribution of US companies by the size of employment follows a power law[24], and specifically a Zipf distribution. In order to replicate this result for the subset of High-Impact Companies, I am testing the following hypothesis:

**Hypothesis 1.** *High-Impact Companies are power law distributed by employment size.*
The log-log distribution of HICs in US using the mean of the ranges in employment size and frequency shows that the HICs closely follow a power law distribution.

Figure 3.5: The Log-Log plot of employment size of HICs in US, 2007-2010 data.

In this case, the employment size has been binned in unequal bins of 1–4, 5–19, 20–99, 100–249, 250–999, 1000–4999, 5000–19999, 20,000–300,000 employees respectively.

The employment size distribution of High-Impact Companies is not Zipf, but it still follows a power law: \( y = 10^{7.25}x^{-1.3} \) and \( R^2 = 0.94 \). This result confirms the hypothesis.

Power law distributions are a regularity or a pattern observed in the data originally found and researched in physics [57]. In social science, the power law shows a universal character of the phenomenon that can be "removed" from the context and conditions that
social phenomena are usually bounded by. In other words, the social phenomena power law distributions show that the phenomenon is independent of its social domain. Particularly with respect to growth and change, social phenomena are more likely to follow a power law [28].

Figure 3.6: The log-log distribution of HICs on employment size, 1995-2006 data. The data was binned in equal bins of 100. The highest value is 35,000 employees.

The log-log test for the employment size distribution of HICs for the second data set (1995-2006) confirms the hypothesis that HICs follow a power law distribution (see Figure 3.6). The following plot is used for testing the granular data on this data set. The mathematical expression of the distribution is: \( y = k \times x^{-2.44} \), where \( k \) is the intercept. This result is also a power law close to a Zipf distribution, and we can infer from this that the
granular data from the second dataset is more refined for testing and modeling.

The log-log tests for both data sets support the hypothesis that HICs follow a power law distribution by employment size.

As hypothesis 1 is confirmed, the following step is to look at the distribution of HICs on areas of job density. The job density areas are geographical areas separated as such by the average number of employees per company in one MSA [56]. This variable calculates the concentration of employees in all 273 markets throughout the U.S:

\[ e_d = \frac{e}{c_{tot}}, \]

where:

- \( e_d \) is job density of MSA
- \( e \) is employment in MSA
- \( c_{tot} \) is total companies in MSA.

This variable calculates the density of employees in 273 markets throughout the U.S. as defined by the ACSL (American Corporate Statistical Library) [56]. Market job density is separated into 8 ranges of concentration (Highest Concentration, Higher Concentration, High Concentration, Mid-High Concentration, Mid, Low Concentration, Low Concentration, Lower Concentration, Lowest Concentration), ranging from less than 50 to over 800 employees as an average of the employment size per company in one MSA.

Figure 3.7 shows that the number of HICs is, in general, decreasing with the size of job concentration in MSAs. This is a counter-intuitive result, since HICs are expected to be found in areas of high job density in US. On the contrary, this figure shows that actually there are more High-Impact Companies in areas with low job-density and less high-impact companies in job-concentrated areas. The implication is that most of new jobs created by HICs are to be found in areas with few jobs. This could imply that, as job creators, HICs are responsible for most of the jobs in the local areas where they activate.

But, since data with respect to the distribution of all US companies on areas of job
density was not immediately available from the US Census, it is difficult to formulate this empirical finding into a hypothesis and remains as an open question for further explorations.

### 3.3 Geographical Distribution of High-Impact Companies

People tend to start businesses in areas they are already working in and start-ups tend to be similar to the areas companies in concentration in that specific local area [16]. The HIC data decomposed by submarkets and industries will show, in the next sub-chapter, how businesses cluster in the same economic geographic area, supporting this claim. Start-ups are also more common in services than in manufacturing, and in personal services than
professional services. There is no correlation though between start-up emergence and rich areas. On the contrary, money follows the start-up preference and the overall geographical distribution of HIC will show, also in the next chapter, that they are spread all over US. They do tend to accumulate though in areas of capital availability, such as metropolitan areas and access to capital is one of the 3 crucial necessary factors for the growth and success of a company [1].

**Hypothesis 2.** High-Impact Companies are located only in a few specific locations on the territory of US, such as Silicon Valley, major urban areas or the East Coast.

The geographical distribution of the High-Impact Companies shows that they are roughly spread all over the territory of US (*see Figure 3.8 a.*) [62]. Unsurprisingly, most of the locations are in the North East area and Los Angeles area. What is surprising though is the number of locations in the South East area, particularly in Florida (*see Figure 3.8 b.*). The locations represent the submarkets as defined by ACLS and there are 1,776 geographic submarket identifiers all over US.

Submarkets divide markets into more manageable sizes to provide more precise information about the immediate vicinity of an establishment. By understanding the submarket, it is possible to identify opportunity and challenges for the establishments doing business there [56]. Submarkets are not only MSAs or major urban areas, they also include smaller towns and rural areas.

Their density is different though for each of the above locations. While Florida is still among the top of the number of companies in the state, areas like North Virginia, Atlanta GA and Columbus OH show a high density of these companies on their submarkets. The density map on top shows ”clusters” of HICs in different US areas by the location of their headquarters (Alaska, Puerto Rico and Hawaii do not represent such ”clusters” of HICs), while the density map on the bottom shows how many companies are found in one location all over the territory of US.
The geoproximity of HICs to major transportation routes represents an emergent phenomenon that can be compared to the exploitation of common pool resources for human complex ecosystems institutions emerged in Africa [63]. Proximity to connectivity hubs such as highways and to urban areas reduces coordination costs related to logistics for supply
Figure 3.9: Distribution of HICs both by industry and submarkets, 2007-2010 data. The size of the nodes shows the number of companies within the same submarket and industry (i.e. construction - Moultrie, GA; business services - downtown NY), while the colors represent the industry of the company.

and delivery of products (see Figure 3.10).

For example, Columbus, OH appears as one of the areas where High-Impact Companies are dense. It has a high diversification in industry. It is a business friendly and very well connected area [64]. All the surrounding state destinations are within 2 hours drive. Agents in proximity one with another will learn from each other, whether they cooperate or compete. In these highly connected areas - as it is this case of Columbus, OH - the innovative companies have fewer barriers of entry on the market than elsewhere [64].
A study that looked at the relationship between entrepreneurship, employment and creativity showed that the larger the creative employment of a region, the higher the levels of entrepreneurship and regional growth. A region that has more new entrepreneurial firms open in the previous year is more likely to have higher growth in new entrepreneurial firms in the current year [65].

Visually, the geographical distribution of High-Impact Companies practically tells us that they can be found all over the territory of US and not just in certain areas, as Silicon Valley or North East coast. The following subsection tests this hypothesis.
3.3.1 Statistical tests for geographical distribution

The log-log plot of the distribution of High-Impact Companies on zipcodes shows that a few locations foster a large number of these companies, but that does not support the hypothesis that they are present all over the territory of US (see Figure 3.11).

Figure 3.11: The log-log plot of High-Impact Companies distribution on zip codes in US, 1995-2006 data and their comparison with 2010 US Census data for all companies.

*Comparison with the whole US Economy.* The US Census data published in 2010 with respect to ZIP codes and MSAs shows that the US companies are distributed all over the
territory of US (see Figures 3.11, 3.12 and 3.13). All zipcodes have at least one company and all MSAs have at least one company registered (data is by establishments).

Similarly, the log-log plot of the distribution of High-Impact Companies on MSAs shows that a few metropolitan areas foster a large number of these companies, but overall they are present all over the metropolitan areas of US (see Figure 3.12).

Figure 3.12: The log-log plot of High-Impact Companies distribution on MSAs in US, 1995-2006 data and their comparison with 2010 US Census data for all companies.

The following log-log plot of the distribution of High-Impact Companies on county codes

47
shows that a few counties foster a large number of these companies, but overall they are present all over the counties of US (see Figure 3.13).

Figure 3.13: The log-log plot of High-Impact Companies distribution on county codes in US, 1995-2006 data and their comparison with 2010 US Census data for all companies.

The tests above show the independence of the emergence of High-Impact Companies on their absolute geographical location (such as zipcodes or metropolitan areas), supporting the hypothesis that they are distributed all over the territory of US. Therefore hypothesis no. 3 is not verified.
3.4 High-Impact Companies and Industries

Another popular conjecture in entrepreneurship research says that most successful companies are high-tech or are technological innovators. In this subsection I am looking at the distribution of High-Impact Companies across all industries in US by SIC codes. I am testing the following hypothesis:

**Hypothesis 3.** *High-Impact Companies are activating only in certain industries, particularly in high-technology areas.*

As previously mentioned, a very important data find is that the high-impact company phenomenon is not just a matter of finding the right niche. Although necessary, finding the right niche is not sufficient - companies can start in the same industry at the same time, and one will grow fast while the other languishes and that this implies that entrepreneurial skill is probably more important than entrepreneurial alertness. A simple histogram of the distribution of HICs on different industries shows that they are present in all major industries, not only high-tech areas *(see Figure 3.14).*

Approximately 35-40% of new start-ups are in construction, retail, food or services. This represents a 8:1 ratio of other industries against manufacturing start-ups. In the case of High-Impact Companies, the manufacturing industry ratio to others is higher. The original assumption that industry does not seem to matter [13] can actually be revised that most High-Impact Companies are in businesses and services and the least of them are in high-tech manufacturing.

This figure therefore shows that HICs can be found in all industries and that they are actually not very well represented in high-tech manufacturing relative to other groups of industries.

*HICs distributions on geography AND industry.* A combined analysis of industry and geographical distribution gives more insights into the geographical and economic boundaries
Figure 3.14: The distribution of HICs by industry (SIC codes), 2007-2010 data.

(space) of High-Impact Companies. The distribution of HICs on submarkets divided by industry shows that, while New York business professional area has the highest number of HICs, areas like San Francisco high manufacturing are in the tail of the distribution.

Nevertheless, this analysis shows that a finer geographical separation by both industry and location gives some counter-intuitive results. HICs are not located only in Silicon Valley or North East Coast and are not only financial or high-tech companies. These are companies that can be found in construction industry in Florida as well as in manufacturing industry in Texas.

Comparison with the whole US Economy. The frequencies of the distribution of US companies on NAICS codes (2007 data) show that US companies are almost equally distributed
in all industries with an approximately 4.8% frequency (see Figure 3.15).

This shows that, while HICs are consistent with the distribution of all US companies on the territory of US, they are inconsistent with respect to distribution by industry. This analysis above shows that hypothesis no. 4 is not verified.

Table 3.1: The US industries by SIC codes

<table>
<thead>
<tr>
<th>Industry</th>
<th>SICs codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture/Forestry/Fishing/Mining</td>
<td>01-14</td>
</tr>
<tr>
<td>Construction</td>
<td>15-17</td>
</tr>
<tr>
<td>Non High-Tech Manufacturing</td>
<td>20-39 ex High-Tech</td>
</tr>
<tr>
<td>High-Tech Manufacturing</td>
<td>283, 357, 36, 372, 376, 38</td>
</tr>
<tr>
<td>Transportation/Communications/Utilities</td>
<td>40-41, 44-49</td>
</tr>
<tr>
<td>Wholesale Trade/Distribution</td>
<td>42, 50-51</td>
</tr>
<tr>
<td>Retail Trade (excluding Eating/Drinking)</td>
<td>52-59, ex 58</td>
</tr>
<tr>
<td>Eating/Drinking Retail Trade</td>
<td>58</td>
</tr>
<tr>
<td>Finance/Insurance/Real Estate</td>
<td>60-67</td>
</tr>
<tr>
<td>Services (excluding Business/Professional)</td>
<td>70-89 ex Bus/Prof</td>
</tr>
<tr>
<td>Business/Professional Services</td>
<td>73, 80, 81, 87, 89</td>
</tr>
</tbody>
</table>
3.4.1 Occupational distribution in a High-Impact Company

Another interesting popular conjecture that circulates in business circles is that successful or high-growth companies employ only highly educated people from finance and engineering or computer science and that they are mostly focused on financial services or producing a technological product. An analysis of the occupational distribution within HICs will help understand better the inner processes and managerial dynamics within these companies. Therefore, I am testing the following hypothesis:
Hypothesis 4. The High-Impact Companies employ only highly educated people.

The following figure shows which are the predominant labor activities within HICs (see Figure 3.16). This result shows that mostly HICs perform managerial and clerical activities.

Figure 3.16: The dependence of managerial activity on other occupations in HICs. (2007-2010 dataset)

The distribution of jobs within High-Impact Companies shows very little variability. High-Impact Companies show similar percentages of labor distribution within their organization. The occupations given by data are: professional, managerial, sales, clerical, crafts, operatives, services, laborers and other. Each of these is separated into 7 ranges of labor concentration.

This shows that productive activities in High-Impact Companies are similarly distributed within these companies. It also shows that the managerial structure or business process is essentially the same regardless of the industry or geographical location where they perform.
Management concentration in HICs shows peaks in the construction and business/services industries, meaning that high-impact companies have either very small or very large managerial concentration. Guy Kawasaki said that he values start-ups at $100,000 per MBA and $1,000,000 per engineer, while Mark Suster said that the ideal 6 person company has 5 engineers to 1 CEO ratio [66].

The regression of the managerial activity on other occupations above shows that, generally, HICs have low managerial activity relative to their other internal activities. This empirical result is incorporated into the team formation method in the ABM in Chapter 4.
3.5 The Ages of HICs

Another popular conjecture with respect to successful companies is that they are usually start-ups. Acs et al. and Lesonsky already asserted that HICs and "gazelles" respectively are not necessarily start-ups [1, 55].

The analysis of the difference between the founding year and the starting year of the high-impact phase that HICs are not necessarily start-ups. In order to see if HICs are start-ups or young companies, a time series of the founding years in the over 950,000 population of companies from the second data set shows that indeed the company age is not relevant
for the high-impact phase.

<table>
<thead>
<tr>
<th>Correlations</th>
<th>Year Start</th>
<th>Year Founded</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Correlation of Year Start and Year Founded</td>
<td>1.000</td>
<td>0.731</td>
</tr>
<tr>
<td></td>
<td>0.731</td>
<td>1.000</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td>Year Start</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Year Founded</td>
<td>0.000</td>
</tr>
<tr>
<td>N</td>
<td>Year Start</td>
<td>951189</td>
</tr>
<tr>
<td></td>
<td>Year Founded</td>
<td>951189</td>
</tr>
</tbody>
</table>

Figure 3.19: The correlation between founding year and high-impact start year in HICs, 1995-2006 data. Although the correlation coefficient is high, there is no perfect correlation between the starting year of the company and the starting year of the high-impact phase.

The Pearson correlation between the year the companies were founded and the year of the start of their high-impact phase is 0.731, significant on a 2-tailed 0.01 level.

The figure above shows that HICs are age independent. The oldest high-impact company is 357 years old. Out of the 951,189 companies for which there is data with respect to the founding year, there is a significant number of companies - 102,256 - that are older companies. Out of these, more than 65,600 companies are older than 5 years and therefore survived the start-up threshold of success [10,36].

### 3.6 Super High-Impact Companies

Super HICs are those High-Impact Companies with more than 8 years of HIC-type of growth [13]. The data for the super High-Impact Companies comes from the second data set, that comprises all super High-Impact Companies in US for the period of 1995-2006 (over 900,000 companies). With respect to the age of the HICs from the 1995-2007 data set, there were 22% missing values.
companies). A map of the geographical distribution of super High-Impact Companies during this period shows that they are broadly distributed all over the territory of US. There were approximately 20,000 super High-Impact Companies during that decade. (see Figure 3.20).

![Map of the geographical distribution of super High-Impact Companies in US, 1995-2006 data.](image)

Figure 3.20: The Geographical Distribution of Super High-Impact Companies in US, 1995-2006 data.

This map shows that companies that live long enough and are successful enough to qualify as high-impact companies are geographically and timely independent. The Pearson correlation in the following map shows that the employment in the fourth year of the 4 year aggregated period is not correlated with either the zip codes nor the county the company’s headquarter is situated (they are both close to 0). Since a linear regression of the employment size of the company on the zip codes is not significant for showing clustering, the absence of correlation shows that there is no dependence between the employment size
and geographic location. (see Figure 3.21).

<table>
<thead>
<tr>
<th>Correlations</th>
<th>Year 4 Employment</th>
<th>Zipcode</th>
<th>Fips County</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 4 Employment</td>
<td>Pearson Correlation</td>
<td>1</td>
<td>0.003</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.003</td>
<td>1</td>
<td>0.032**</td>
</tr>
<tr>
<td>N</td>
<td>41098</td>
<td>41098</td>
<td>41098</td>
</tr>
<tr>
<td>Zipcode</td>
<td>Pearson Correlation</td>
<td>0.561</td>
<td>0.000</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.006</td>
<td>0.032**</td>
<td>1</td>
</tr>
<tr>
<td>N</td>
<td>41098</td>
<td>41098</td>
<td>41098</td>
</tr>
<tr>
<td>Fips County</td>
<td>Pearson Correlation</td>
<td>0.214</td>
<td>0.000</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.561</td>
<td>0.000</td>
<td>1</td>
</tr>
<tr>
<td>N</td>
<td>41098</td>
<td>41098</td>
<td>41098</td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (2-tailed).

Figure 3.21: (a) The Pearson correlation between employment size and zip codes or FIPS county codes, 1995-2006 data. (b) The Kendall tau-b test and Spearman correlation between employment size and zip codes or FIPS county codes, 1995-2006 data.
A much better test for the null hypothesis is the Kendall’s tau test, which tests the correlation between employment size and zip codes by rank and not linearly. But similarly, the correlation is close to zero. Tau-b statistic, unlike tau-a, makes adjustments for ties and is suitable for square tables. Values of tau-b range from –1 (100% negative association, or perfect inversion) to +1 (100% positive association, or perfect agreement). A value of 0 indicates the absence of association. A Spearman correlation of 0 indicates that there is no tendency for Y to either increase or decrease when X increases. The Spearman correlation increases in magnitude as X and Y become closer to being perfect monotone functions of each other. Unlike the Pearson correlation, the Spearman correlation does not implicitly rely on a linear function between the dependent and the independent variables.

The super High-Impact Companies show a weak correlation between the growth quantifiers. (see Figure 3.22). As expected, for longer period of times, there is more linearity between employment and revenue.

<table>
<thead>
<tr>
<th>Correlations</th>
<th>Sales Growth Quant</th>
<th>Employment Growth Quant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales Growth Quant</td>
<td>Pearson Correlation</td>
<td>1</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>41098</td>
<td>41098</td>
</tr>
<tr>
<td>Employment Growth Quant</td>
<td>Pearson Correlation</td>
<td>.035**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>41098</td>
<td>41098</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 3.22: The correlation between EGQ and RGQ in super High-Impact Companies, 1995-2006 data.

But this can also mean that actually the employees in super High-Impact Companies are not as efficient as the employees in High-Impact Companies; in High-Impact Companies,
the company stops hiring, but brings more and more revenue subsequently.

By industry type, the same statistical tests and correlations show no association between the company employment size and the industry they are performing in. There is a slightly higher correlation between sales in year 4 and industry type, but the difference is insignificant (see Figure 3.23).

3.7 High-Impact Companies and Entrepreneurial Activity

High-Impact Entrepreneurship is fundamentally the study of the actions of individuals responding to market opportunities by bringing inventions to market that can create wealth and growth. [1] Routine entrepreneurship is really only a type of management [67], while HIE rests on 3 pillars: labor markets, technological change and capital markets. [1] These “three pillars” that are necessary conditions for leverage in a new entrepreneurial organization: occupational choice, technological change and financing choice (or access to capital). [1] Risk or risk perception from an entrepreneur is not an important condition for building a growth business and shifting the wealth creation.

To be a leveraged start-up, you have to be interested in selling one thing to a lot of people rather than a lot of different or semi-custom products to individual clients; these are growth businesses and not ”job replacement” businesses. Risk perception is not an issue for the HI entrepreneur, but hoe they align their life and pursue single-mindedly the success of the business is a crucial factor. This is in line with the first pillar of entrepreneurship, occupational choice and why and how people choose to become entrepreneurs, why human capital and education matter. This is subsequently represented by skills and learning in the model.

The second pillar, technological change, is represented by the tasks with high revenue and high labor hours. According to the knowledge spillover theory of entrepreneurship, the new knowledge is exogenous to the model and the agent endogenously engages in a leveraged start-up [1]. The model replicates this by exogenizing the tasks and opportunities and by endogenizing the team formation and completion of the opportunities they act on.
### Correlations

<table>
<thead>
<tr>
<th></th>
<th>Year 4 Employment</th>
<th>SIC</th>
<th>Year 4 Sales</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 4 Employment</td>
<td>Pearson Correlation</td>
<td>1</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.903</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>41098</td>
<td>41098</td>
</tr>
<tr>
<td>SIC</td>
<td>Pearson Correlation</td>
<td>.001</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>41098</td>
<td>41098</td>
</tr>
<tr>
<td>Year 4 Sales</td>
<td>Pearson Correlation</td>
<td>.175</td>
<td>-.009</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.062</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>41098</td>
<td>41098</td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (2-tailed).

### Kendall's tau-b

<table>
<thead>
<tr>
<th></th>
<th>Year 4 Employment</th>
<th>SIC</th>
<th>Year 4 Sales</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kendall's tau_b</td>
<td>Year 4 Employment</td>
<td>1.000</td>
<td>- .054**</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>41099</td>
<td>41099</td>
</tr>
<tr>
<td>SIC</td>
<td>Year 4 Employment</td>
<td>- .054**</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>41099</td>
<td>41099</td>
</tr>
<tr>
<td>Year 4 Sales</td>
<td>Year 4 Employment</td>
<td>.599</td>
<td>- .107**</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>41099</td>
<td>41099</td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (2-tailed).

### Spearman's rho

<table>
<thead>
<tr>
<th></th>
<th>Year 4 Employment</th>
<th>SIC</th>
<th>Year 4 Sales</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman's rho</td>
<td>Year 4 Employment</td>
<td>1.000</td>
<td>- .079**</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>41099</td>
<td>41099</td>
</tr>
<tr>
<td>SIC</td>
<td>Year 4 Employment</td>
<td>- .079**</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>41099</td>
<td>41099</td>
</tr>
<tr>
<td>Year 4 Sales</td>
<td>Year 4 Employment</td>
<td>.759</td>
<td>- .162**</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>41099</td>
<td>41099</td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (2-tailed).

Figure 3.23: (a) Linear correlation between super high impact companies size and industry type and (b) The Kendall tau-b test and Spearman correlation between employment size and SIC codes (1995-2006 data).
Since the firm does not exist exogenously as in most theories of the firm, the team formation and growth of the firm is an endogenous process. An important note to be made here is that these are not necessarily high technology firms. While there is a popular conjecture that most successful start-ups are high-technology firms [16], actually only 10% of the high-impact firms are high-technology firms.

The third pillars of HIE is given by the capital markets. The way the firm is financed is given in the model by scenarios that describe different access to capital and the way this gives a company the "leverage" necessary to become a HI firm.

Data suggests that most people who become productive entrepreneurs are actually employed at the time they make the decision to become an entrepreneur [1,16]. An analysis of the theories and models of why a person decides to become an entrepreneur shows inconclusiveness. Jovanovic proposed a learning model where he concluded that entrepreneurs learn about their abilities, which cannot be known ex-ante. [68]

The work that a researcher conducts for a firm increases both the firm’s stock of innovations and the human capital of the researcher. [1] This is depicted in the model by the learning rate for the teams and employee skills in the company. This implies that researchers with experience are more likely to decide to leave the firm and start a spin-off. The agent based model experiment results show that entrepreneurs with the highest skill levels in the beginning of the company growth are more likely to lead to high impact companies than if the firm starts with lower skills entrepreneurs. After the company grows, the employees become more and more specialized in the skills they are using. The spin-off former employees can give to new innovations in their new firms are facilitated by the nature of knowledge.

The theory of entrepreneurs as generalists is more likely to be applicable to replicating entrepreneurship, while the theory of entrepreneurs as specialists is more likely to be applicable to the "gazelles". On the incentives scheme, Hvide found that small sized companies are more likely to "fine-tune" their wage policies to their workers, while large firms have more rigid policies [69]. In the model presented in the following chapter, this is represented by the wages that are paid according to the level of skills an employee possesses. All these
studies show that, from a behavioral perspective, there is no answer on why entrepreneurs choose to start a new business.

Figure 3.24: The market entrepreneurship indicator for HICs, 2007-2010 data. This indicator shows the distribution of HICs on growing and dynamic markets

Innovative entrepreneurial firms, even those in the high-tech industry or those that are very young, have a balanced capital structure, close to 50% between debt and equity, similar to all US businesses. Private equity dominates the early growth stages of the company, while debt financing becomes more and more important as the company gets ready to become public. [1]. Also, when it comes to access to capital through external funds, the external debt is preferred to external equity. Agency theory has emerged as the dominant theoretical perspective for studying the relationship between the venture capitalist and the entrepreneur [70]. Acs notices that the practices of venture capitalist investment have little
changed over time (last 500 years, more specifically): practically, the venture capitalists employ the following methods to monitor their investments: staged financing, preferred stock, board seats, negative covenants and specific exit rights. [1,16]. The staged infusions of capital are as important as the intensive monitoring for the venture capitalists.

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid Above Aver</td>
<td>80</td>
<td>19.9</td>
<td>19.9</td>
<td>19.9</td>
</tr>
<tr>
<td>Average</td>
<td>82</td>
<td>20.3</td>
<td>20.3</td>
<td>40.2</td>
</tr>
<tr>
<td>Below Aver</td>
<td>80</td>
<td>19.9</td>
<td>19.9</td>
<td>60.0</td>
</tr>
<tr>
<td>High</td>
<td>85</td>
<td>21.1</td>
<td>21.1</td>
<td>81.1</td>
</tr>
<tr>
<td>Low</td>
<td>76</td>
<td>18.9</td>
<td>18.9</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>403</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3.25: The distribution of HICs in frequency and percentages on different job dynamics, 2007-2010 data.

The market entrepreneurial indicator (see Figure 3.24) shows growth through newly created businesses because of emerging opportunities. It is quantified from 1 to 5 (low to high) based on the departure from the national mean of new start-up jobs created (within 2%, between 2 - 4% or less or higher than 4%).

Figure 3.25 below shows that High-Impact Companies are approximately equally distributed (within 20% each) on areas of new jobs created by start-ups in US. This means that High-Impact Companies do not depend on how dynamic the local job markets are.
Chapter 4: An Agent-Based Model of Company Growth

4.1 Conceptual Model and Scope

There have been a few agent-based models published to date with respect to firm growth [71], entrepreneurship [45, 54] or team formation [72, 73].

In the model of firm formation based on endogenous teams and self-organizing employees, the employees are utility maximizers and adjust their efforts according to an equally distributed output between team members [71]. There is labor turnover and the success of the company is based on the ability of a company to hire and retain productive workers, therefore eliminating "free-riders". In the model presented in this chapter, the employees are perfectly incentivized according to their skills and are not utility maximizers. As in Axtell’s model, the employees are being hired and fired, but not based on their free-riding nature, but according to the level of skills required by the task the firm has to successfully complete. Also, the success of a company in this model is represented by the level of revenue a company can acquire during a simulation. As in Axtell’s model, this model here is based on self-organizing teams and employees and shows the life-cycle of a company based on employment size, but apart from the original model, this model is looking for specific conditions of maximum shifts in employment size.

Black, Oliver and Paris [54] provide an interesting agent based model for entrepreneurial action. In this model, entrepreneurial "conation"\(^1\) represents what theoreticians have named entrepreneurial "alertness" or other authors named it "gaze" [45]. The entrepreneurs are heterogeneous, with different skills and conation profiles, and the model demonstrates that entrepreneurs with one high, but not both, of these profiles, are more likely to commit

\(^1\)"Conation", as defined and parametrized by the authors [54] can be interpreted as the innate, genetic entrepreneurial character of a person.
to the opportunities they spot. The model shows that either of the two will commit to the opportunities they spot. Another important conclusion they make is that the high skilled entrepreneurs reach their opportunities faster than the other ones. Berea and Runst presented another conceptual model of entrepreneurship based on opportunities and entrepreneurial alertness [45], that showed how entrepreneurial characteristics are evolving inter-generationally.

Distinctively from these models of entrepreneurship, the model of firm growth presented here does not take into account the role of the entrepreneur based on some innate skills. Any employee can start a firm and can be fired from the firm later on if his skills are no longer useful. On another hand, this model retains the exogenous nature of the tasks or opportunities and the matching of the tasks with the skills of the employees or entrepreneurs.

The models that describe team formation [72] and team coordination [73] are not models of firm growth. These models describe how teams form in a combinatorial fashion due to different activation regimes (Padgett) or in an optimizing fashion due to coordination costs (Rojas-Villafane).

Padgett’s model of hypercycles of team formation based on skills shows that non-used skills are forgotten, while often-used skills will reproduce. Team formation is activated based on the source skill, the target skill or both. [72] Padgett’s model shows that, depending on the activation of skills, the team is self-organizing or hierarchical. As in Padgett’s model, this model introduces learning at the firm level, only that learning is not based on the products that pass through a firm, but on the skills and team work that pass through the firm. As in Padgett’s model, firms and teams are collections of skills. The products in Padgett’s model are looking for compatible skills to activate; similarly in this model, the tasks are activating the employees based on their skills that would be compatible with that specific task. But individual learning behavior is different than Padgett’s model, since individual learning is dependent on the other teammates skills and not solely on the use of the skill. Also different than Padgett’s model, there is no individual decay rate, since mathematically it represents only a multiplier of the current skill. While Padgett’s model
is spatially distributed and the behavior of the firms is influenced by their neighbors, this model is not - here, the firms learn only endogenously from the tasks and the employees they complete and use. But the internal learning behavior is similar in the two models: the skills that are used are passed on from one time tick to another, while skills that are not used are forgotten (fired). As in Padgett’s model, a firm dies when it does not have any more skills (employee population = 0) to work with. Apart from Padgett’s model, where the population of skills is held constant (by analogy with chemistry), in this model learning and skills value can be infinite, as in social systems there is no ”constant” population of skills, they change and diversify continuously. In this model, skills are not substitutable and, as in Padgett’s model, activation is based on the compatibility of capacities with the requirements of the tasks.

The Team Coordination Model (TCM) proposed by Rojas-Villafane is an interesting agent-based model of team formation that introduces the concept of ”coordination cost” [73]. While the model is looking for an optimizing solution to the performance of teams taking into consideration a roughly parameterized coordination cost (generally equivalent to time and communication costs) spent between employees in order to complete a task, it does not show how teams form or evolve or how the firm lifecycle emerges. But TCM does not include individual skill heterogeneity apart from 3 levels of expertise (low, medium and high) and it does not include learning either for the individual or the teams.

Generally speaking, in an agent-based model, the skills that lead to entrepreneurial decision-making can be either parameterized or the model can be populated with cognitive agents (such as ACT-R) that exhibit learning behavior. In this model the methods of interaction are simple though and do not imply any decision behavior. But, in an agent-based model, the collective work of the employees as a team can be successfully represented as more than the sum of its parts - an emergent behavior. The representation of skill learning and team formation first needs to answer the fundamental question: how does a team of agents evolve/perform based on the evolving skills of individuals? What exactly is a skill? Skills are attributes of the individual, employee or entrepreneur and it represents
the ability to complete a certain action $i$ or is an input to complete another action $j$, which requires more than one skill or one type of skills. In the literature it has been represented as traits and features [74], arraylists (NetMason, Padgett) [75, 76], Cobb-Douglas (endogenous growth models)[77], GA (computational organizations, resource-based), game-theoretic (cooperative/competitive behavior), tech production recipes (Auerswald/Kauffman)[78], information/beliefs diffusion models [79].

The model developed and presented here has borrowed different aspects from the models above. Overall, this model has borrowed some concepts and methods as well as proposed new ones in order to preserve the nature of the firm as both an organization and a process.

Also, since High-Impact Companies are not dependent on location, age or industry (see Chapter 3), this computational effort is modeling the internal business processes of a company, particularly how teams form and coordinate based on the skills of the employees. Aldrich [80] stated that there are two theories of team formation in entrepreneurship research - one that is based on the complementarity of skills and work experience and another one that is based on the fit between employees and the "smooth functioning of the team members". Similarly, this model is based on team formation starting with the skills of the employees (that include the work experience) and the smooth functioning of the teams (coordination of the inter-team work).

The "inner working" of a company, namely the managerial structure or the business process lays at the conceptual core of this model. The scope of the model is to find and replicate "anomalies" or drastic changes in employment growth. One of the fundamental questions with respect to the conceptual development of this model was whether the entrepreneurial skills and managerial skills should be separated, as in real life many entrepreneurs employ other people to manage their business. Evidence shows that for "gazelles", the entrepreneurs remain the managers of their companies in the high majority of the cases [50]. Therefore this model considers the original employees as "founders" or "entrepreneurs" and does not model specific "entrepreneurial" skills, which are very loosely defined conceptually by the literature (see Chapter 2).
This agent-based model estimates the growth of the companies in revenue and employment based on team formation and coordination, a series of exogenous projects and the skills of the employees. Essentially, this is an ABM of team formation and coordination. Teams represent the mezzo level of a company and approximate the managerial structure (see Figure 4.1). The conceptual development of the model starts with identifying agents and agents’ behavior and ends with the implementation of the model in MASON, a discrete multi-agent simulation based on JAVA platform [81].

The output represents a company (maximum time span) or a phase of evolution of a company (flexible time span), as it can grow and decline both in employment and revenue.

The model is initialized with a population of employees that are required to perform a
The task is a combination of revenue and labor hours and there is one task performed at each time tick.

The agents are heterogeneous both at the individual level (skills and wages) and at the mezzo level (team size and the coordination and learning parameters) of the company. The two layers of heterogeneity at the intra-firm and inter-team levels are a better representation of the internal business process than the ones advanced by network theories and applications [75, 82], since teams work as units and learning is transmitted through a organizational "transactive" memory [83]. The conceptualization of team formation in this model is a combination of spontaneous and "by design" formation: the leaders and employees are not chosen randomly, but they are randomly initialized with their pairwise skills.

Conceptually, this model also has to fulfill some necessary and sufficiency conditions in order to be anchored in the reality described by the previous 2 chapters, namely:

*Necessary conditions*: The model has to be general enough to cover any industries, any geographical area, any company age (not just start-ups), to use longitudinal and not relational data.

*Sufficiency conditions*: The model has to be specific enough to distinguish between other types of activities and management, to distinguish between importance of logistical and business processes activities and production activities, and to show the co-evolution of employment and revenue.

### 4.2 The Agents

There are 2 types of agents in this model: the employees and the teams *(see Figure 4.2)*. The agents are paid wages according to the skills they use at each time tick and the teams coordinate with each other in order to collaboratively perform the task.
Figure 4.2: Team formation in the ABM.

The agents employees have an endowment of 2 types of skills: managerial (or agency) skills and production skills. The reason for this dychotomy is that, in a company, there are roughly these two types of tasks that any employee has to perform - the actual work and other administrative procedures. Managerial or agency skills can also be viewed as a proxy for inter-personal relations with team members. Additionally, any manager has different skills for delegating work. Managerial skills are important for the business process or organizational innovation, while production skills are important for product innovation.

The agents are endowed with heterogeneous incentives: each agent receives a unique wage $w_i$, based on the skills they are using at each time tick. All agents use their skills for interacting with the team members, within the team; they do not use their skills for interacting with outside team members at one time tick, only teams do (through the coordination cost) [84].
The formation of teams is efficient: the task is fully completed with the skills of the existing employees and/or the new hired employees. In this fashion, all simulated companies have the highest capital/labor (labor productivity) efficiency. Even under these conditions, the typical simulations show that efficiency is not sufficient for the growth of the companies, as companies with high labor productivity can still fail. Which are the particulars of team formation in a firm? Teams do not form spontaneously in a firm, since there are delegation/agency problems that a firm has to address, but teams are not entirely designed either. Individuals learn and forget, teams learn and forget, but differently [83,85]. Formal education and tacit knowledge are difficult to separate within a business process. Teams also learn and they transmit their knowledge through a learning parameter, $l_w$.

If the individual is a collection of skills, the team is a collection of individuals and the firm is a collection of teams. At any time tick, the multiple logical relations of one-many; many-one and many-many between skills-employees-teams-company are being fulfilled.

The entrepreneurship literature provides sufficient reasons for generalizing entrepreneurial skills into the mix of managerial and production skills for this model. Many researchers have tried to make all sorts of lists of psychological characteristics associated with entrepreneurs, such as: perseverance, commitment, problem-solving skills, leadership, risk taking, goal setting and many more. But any attempts to accurately ”profile” an entrepreneur have failed. The decision to behave entrepreneurial is a mix of environmental factors (social and business networks) and preferences that currently is unclear for researchers how it leads to decision making. But there are 3 skills that most researchers seem to agree on as being very important for the birth and development of a ”gazelle”: managerial skills, risk preference and “high-stress” resistance. These skills can be parametrized; particularly in the case of the managerial skills, the business literature asserts that there are currently not sufficiently good qualitative or quantitative methods to explain what is the managerial skill of an entrepreneur who creates a ”gazelle” [48].

Also, approximately 92% of new startups have an owner who has experience in that industry, typically of 10 years, and 55.9% of new firm founders attribute their new ideas to
their experience [16]. The interactions of the entrepreneurs with their customers at their
previous jobs is listed as a primary source of new business idea (30.9%). Therefore 61% of
new businesses serve the same customers and markets as their previous employers do. An
interesting find is that people who dealt drugs as teenagers are 11-21% more likely than
others to start their own businesses in adulthood [86]. Also, going to school increases the
likelihood of becoming an entrepreneur, just not the graduate degrees. People who have
more managerial experience are more likely than others to start their own business; also,
people who are in the business world are more likely to start their own business than people
in government, education or healthcare [87].

4.3 The Base Case of the Model

The model of company growth based on team formation is simple: it has 2 types of agents
(employees and teams), 7 rules of behavior (fire, hire, work, pick manager, pick employees,
pay wages, learn) and a few parameters (see Table 4.2).

Based on these parameters, the model is computing the following resulting endogenous
variables:

The model starts with the initialization of a heterogeneous population of employees and
of a random task. The task is a combination of revenue and labor hours and there is one
task performed at each time tick:

\[ T = (R, h), \]  

where

\( T = \) current task or project;
\( R = \) current project revenue;
\( h = \) current project labor hours.

The initial population of employees are the founders (if the company is a start-up) or
<table>
<thead>
<tr>
<th>Parameters</th>
<th>Unit</th>
<th>Representation</th>
<th>Base Case</th>
<th>Maximum Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employee Size</td>
<td>int</td>
<td>Initial population of employees</td>
<td>1</td>
<td>300,000</td>
</tr>
<tr>
<td>Learning Weight</td>
<td>double</td>
<td>Managerial skill learning</td>
<td>0.1</td>
<td>1</td>
</tr>
<tr>
<td>Coordination Weight</td>
<td>double</td>
<td>Time lag in inter-team work</td>
<td>0.1</td>
<td>1</td>
</tr>
<tr>
<td>Skills</td>
<td>double</td>
<td>Managerial and Production skills</td>
<td>U [0,1]</td>
<td>pcap=mcap=100000</td>
</tr>
<tr>
<td>Task revenue (R)</td>
<td>double</td>
<td>Expected net revenue from a project after fixed and investment costs</td>
<td>U [0,1]</td>
<td>n/a</td>
</tr>
<tr>
<td>Task labor hours (h)</td>
<td>int</td>
<td>Expected labor hours required for the completion of the project</td>
<td>U [0,1]</td>
<td>n/a</td>
</tr>
<tr>
<td>Probability of firing</td>
<td>double</td>
<td>Probability for an employee to be fired even if s/he does not have the lowest level of skills</td>
<td>0.1</td>
<td>not varied</td>
</tr>
<tr>
<td>Run time</td>
<td>time tick</td>
<td>Event; Time to complete a task</td>
<td>100</td>
<td>not varied</td>
</tr>
</tbody>
</table>

The existing work force in the company at time 0 (if the simulation represents only a phase in the company life cycle).

The employees are randomly endowed with a set of 2 skills, production and managerial skills, with values expressed by real numbers. The task is also randomly initialized with a number of labor hours required (expressed by an integer value) and an expected revenue (expressed by a real value):

The motivation for randomizing these initializations (skills and tasks) is in accordance with the entrepreneurship literature and the generalized conditions for a model that needs to
Table 4.2: The Variables computed in the model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unit</th>
<th>Representation</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employment</td>
<td>int</td>
<td>Company employment after each task event</td>
<td>&gt;= 1</td>
</tr>
<tr>
<td>Employment changes</td>
<td>int</td>
<td>Discreet marginal employment after each task event</td>
<td>&lt; 0, 0 or &gt; 0</td>
</tr>
<tr>
<td>Net revenue</td>
<td>double</td>
<td>The revenue the company retains after completion of task</td>
<td>&lt; 0, 0 or &gt; 0</td>
</tr>
<tr>
<td>No. failed projects</td>
<td>int</td>
<td>A project that has a net revenue &lt; 0</td>
<td>&gt;= 0</td>
</tr>
<tr>
<td>Wage</td>
<td>double</td>
<td>The wage paid to employee according to the skill s/he uses</td>
<td>= managerialskill or productskill</td>
</tr>
</tbody>
</table>

Initialize  employee population size  
Initialize  task labor hours and task revenue  
Record  employee skills (managerialskill and productskill)  
Record  task revenue

Figure 4.3: The pseudo-code for the initialization of the model.

cover different industries, products and educational and experience specialties. At any time, a company performs not only productive tasks (with the purpose to simultaneously deliver
several products or services or both), but also many administrative tasks. Any company has its specific procedures for dealing with customers, suppliers, for managing its employees, but also for fulfilling law abiding requirements such as filing takes. Even the most specialized company has a multitude and diverse pool of tasks that needs to perform simultaneously or maybe only once (i.e. filing for bankruptcy). Also, even in the most unspecialized company, the employees differ in their skills and aptitudes required to perform a task. While the tasks are more or less easy to represent quantitatively in terms of revenue, costs and work input, the skills of the employees are subjective and difficult to quantify. The overall skills of an employee are a mix of education, experience and psychological characteristics such as intuition or attitude, that are unique for each employee and difficult to express quantitatively. Therefore this subjectivity and heterogeneity of the skills is expressed in the model by unique real values for each employee. In this way, the model retains the diversity and variety of projects and employees that can be found in the world at any point.

A task that has values for revenue and labor hours that are low (closer to 0) can be interpreted as a project that is not very laborious and does not bring too much revenue (such as a laundry place). A task that has "high" values for labor hours and revenue can be interpreted as a complex task such as building a highway or luxury brand (high here means the difference in value of this attributes comparatively to other tasks). A task that has high revenue and low labor hours can be approximated for a successful consultancy company, while a task that has low revenue and high labor hours can be interpreted as a craftsmanship or a restaurant or any activity that is laborious and not very successful overall (i.e. agriculture in the 19th Century).

In this way, the model allows for a large variety of inputs. After the task "comes", it is broken down into subtasks. These subtasks are assigned to spontaneously formed teams and the existing employees are staffed into these teams in order to complete the subtasks. The completion of any subtask is done by matching the value of the added skills of the employees to the labor hours value of the subtask.
Divide task labor hours into subtasks

Number of subtasks = Number of teams

Sort existing employees by production skill values
Assign employees to subtasks based on their production skill
Record the number of subtasks involved in this task labor hours

Figure 4.4: The formation of teams in the model.

The employees are staffed in the decreasing order of their production skills. The employee with the highest value of the managerial skill within his team is picked as the team leader. In this way, the best employees are the ones first retained by the company, whether they are specialists or managers:

Sort employees by managerial skill
pick the team managers (max managerial skill)

Figure 4.5: The assignment of the managers of the teams.

The company hires employees based on the availability of current employees and on the labor hours of the task and on the labor costs incurred by the currently hired employees:
If not enough managers to staff every team
    then hire a new employee

Get an unassigned employee
    add him to the team

Calculate labor cost for company (sum of all wages)
If labor cost < task labor hours
    then hire a new employee

Figure 4.6: The hiring of the employees.

If an employee does not work (i.e. is not assigned to a team) he has a 10% probability of getting fired from the team.

The 10% probability can be changed in the model if the model needs to be adapted to another environment. There is no quantitative research with respect to how much is the value of this variable - the probability of firing the "wrong" employee from the job or for retaining the "wrong" employee that is not competitive enough for the job is difficult to quantify as are the skills of the employee. But this variable is equivalent to the cost of imperfect monitoring of employees or to the informational asymmetries existing in any company [88].
If employee e is assigned continue
else fire him with 0.1 probability

Figure 4.7: The firing of the employees.

This model does not attempt to reproduce the Stiglitz and Shapiro model though, it only justifies the imperfect firing method, as the decision making process and monitoring against shirking is imperfect in the real world.

Each employee is paid a wage according to the skill he is using. Team leaders are paid a wage equal to their managerial skills and the other employees are paid a wage equal to their production skill. The agents are therefore perfectly incentivized and not utility maximizers.

The work is performed by perfectly fulfilling the task at each time tick with hiring/firing the necessary employees into teams, paying the wages and adding the coordination costs to the total labor cost of the company:

\[
\begin{align*}
\text{hours worked by teams} & = \text{task labor hours} / \text{team size} \\
\text{wage} & = \text{product skill} \\
\text{wage} & = \text{managerial skill} \\
\text{total cost} & = \text{wages} \times \text{hours worked by teams};
\end{align*}
\]

Figure 4.8: The company works by completing the task and paying the labor costs.

The company calculates the net revenue as the expected revenue from the Task less
the sum of all wages paid in order to complete that task. This represents the profit of the company. As in real life there are many other costs associated with the business process of a company (such as fixed costs or operational costs or investment costs), none of these costs are directly a function of the individual labor of the employees. In finance, the cash flow is represented by the NPV (net present value) formula [89]. This formula is a summation of all the expected cash flows discounted to the present value. Therefore all other costs associated with a project are included in the expected value of the revenue and from the point of view of this model, only labor costs are a variable.

The coordination cost is linearly dependent on the coordination weight parameter and the size of the team. Coordination weight is the same for the entire company, yet the coordination cost is different for each team. The bigger the team, the higher the coordination cost. At the company level, it increases the labor/work requirements of the company:

\[
\text{coordinationcost} = \text{coordinationweight} \times \text{teamsize}
\]

\[
\text{labor} = \text{managerialskills} \times (\text{totalproductionskills} - \text{coordinationcost})
\]

\( (\text{coordinationcost} > \text{totalproductionskills}, \text{as total production skills} > 1 \text{ and coordination cost} < 1. \) \)

Learning weight is the most important parameter for the transmission of skills. The employees learn at each time tick in two ways: the production skills of each employee are updated with the average of the production skills of the co-memembers of the team discounted with the organizational learning weight; the managerial skills are updated based on a function of the team size discounted with the learning weight. The managerial skill increase with the learning weight and decreases with team size. More employees are more difficult to manage. Recent studies on team size performance and management showed that individuals in larger teams perform worse and that relational loss in larger teams, defined as a form of individual level process loss, predict poor individual performance [90].
\[
\text{productionskill} = \text{learningweight} \times \text{average}
\]
\[
\text{managerialskill} = \text{learningweight} \times \text{teamsize}
\]

Both these skills cannot exceed a predetermined cap (\textit{pcap and mcap respectively}), as learning cannot be infinite. But how much a person can learn either from education or from experience is open for debates and the learning curves concepts in organizational theory [85] showed that learning depends on a series of factors, such as industry, business disruptions and interaction with the market and the environment. This cap can be subsequently removed or changed in the model.

\[
\text{managerialskill} = \min(\text{managerialskill, mcap})
\]
\[
\text{productionskill} = \min(\text{productionskill, pcap})
\]

At a first glance view of a few randomly picked simulations, the values of the skills of the remaining employees in the company after 100 ticks did not reach or get close to the current value of 100,000, but a more thorough analysis of the cap effects is probably required.
net revenue = task revenue - (sum of all wages)
company revenue = sum of all net revenues

company employment = employee population size + employees hired - employees fired

if company hires
    delta employment = no. employees hired
else
    delta employment = no. employees fired

if task revenue < 0
    add number of failures
record number of failures
record net revenue
record employment
record delta employment

Figure 4.9: Company internal accounting.

During one run (simulation), the employment size can grow or decline or it can remain constant. The changes in employment size do not seem to be linearly dependent on the values of the project revenues. This makes sense since the company fires or hires employees according to their skills, the employees are not homogeneous agents (see Figure 4.10).

The following plots show how companies evolve keeping all the parameters constant at minimum in the base case. They don't show any "anomalies" or interesting behavior.

As the figure above shows, there seems to be no relation between the project revenue and the changes in employment. Moreover, the company does not seem to be continuously growing in employment, as the employment change can take negative values.
Figure 4.10: A typical run of the base model. This is a simulation with the main parameters kept constant at the minimum values in the base model. (runtime = 100). The y coordinate on the right is the scale for revenue and the y coordinate on the left is the scale for employment.

Other simulations show that a company grows or declines in employment (see Figure 4.12).

By varying the initial employment size, while keeping everything else constant, the simulation did not show any particular anomalies or regularities in employment behavior (see Figure 4.13). It is just noise. Similarly, by varying only the coordination parameter, the model did not show any interesting behavior. On another hand, though, when the number of founders or the learning weight (lw) were varied (i.e. founders between 3 and 30 and lw between really high (> 0.7) and really low (< 0.3) values), the simulation showed interesting behavior with respect to employment changes (see Figure 4.13).

By varying the parameters at the maximum values (1) in the case of cw and lw and a relatively high number of founders (4), there are no particular behaviors that can be
Figure 4.11: Other typical runs of the base case of the model (runtime = 100). The y coordinate on the right is the scale for revenue and the y coordinate on the left is the scale for employment. These plots show no interesting behavior.

Figure 4.12: The simulation of a company that grows (a) or not (b) (runtime = 100). The y coordinate on the right is the scale for revenue and the y coordinate on the left is the scale for employment.

observed other that the changes in employment seem to be declining in magnitude "sooner" (during the first 30 time ticks) than in the base case (decline after the first 50 steps) (see Figure 4.14).

The main parameters in this model are the initial population size, the inter-team coordination and the team learning weight.
1. $e = \text{an initial employee population size, that can range from 1 up; the company growth can be initialized from start-up or from a later stage}$

   $$e \geq 1 \quad (4.2)$$

2. $cw = \text{a coordination parameter for inter-team work at each time tick}$

   $$cw \in (0, 1), \quad (4.3)$$

3. $lw = \text{a learning weight for intra-team work at each time tick}$

   $$lw \in (0, 1) \quad (4.4)$$

### 4.3.1 The parameters - the learning weight

The *team learning weight* or parameter $lw$ approximates the ”transactive” or shared memory in close relationships that Wegner et al. described in an influential paper [83]. This
Figure 4.14: A simulation with maximum values of \( lw \) and \( cc \) and high number of founders.

The parameter allows for organizational learning at the company level. Wegner concluded that: “The existence of an organized system for knowledge in the couple holds with it the potential for disorganization.” [83]. The employees that remain hired retain the learning from the previous teams and co-workers, while the newly hired employees start to learn from scratch. At the same time, with the employees that are fired the company loses both the production and the agency skills these acquired while working in the company teams. This type of learning allows for memory transition in relationships [83]. Even more so, in relationships that are impromptu (such as team formation in this model), the memory of the tasks that require assigned expertise is better recalled and stored by individuals.

### 4.3.2 The parameters - the coordination

Coordination \( cw \) essentially represents a time lag, it is a time parameter that denotes how long it takes for several subtasks to be sequenced or simultaneously executed. A
coordination cost of 0 means that all subtasks are executed simultaneously and there is no waiting time from one team to another. Coordination cost higher than 0 represents a waiting time between the completion of the subtask that each team has to perform at one time tick. In the Team Coordination agent-based Model (TCM) proposed by Rojas-Villafane, coordination mechanisms are described as team communication (time spent in information sharing and decision making), task organization (the use of formal methods) and a shared mental model (overlapping knowledge between teammates) [73].

In this model, the overlapping knowledge is represented by the learning weight of the team, while the task organization is perfect and efficient. Management is an expression of internal transaction costs and coordination costs and "good" management reduces these costs. Management is the most desirable product for the company if it reduces transaction costs [73].

4.3.3 The parameters - the initial company size

The model allows for any initial company size $e$. A company can be simulated starting with either 1 employee or as much as 100,000 employees. This parameter allows therefore not just for simulation of start-ups, but for simulation of company evolution at later stages. In this way, the later phases of high-impact growth in older companies can be included in the analysis.

4.4 The Environment

This model is about growing a firm with multiple, simultaneous teams from any non null number of founders up. The firm hires when there are not enough employees to fulfill the tasks at one time tick and fires the employees with the lowest skills, with a probability parameter (as in real life, the firing decision between one employee and another can be subjective and sometimes erroneous).

In reality, the market or a industry poses 3 major problems at the microeconomic level:
coordination problems (temporal and spatial); incentive problems (such as contestable markets \[91\] or fiscal policies) and information problems (asymmetries) \[92\]. For entrepreneurs and employees as procedural actors, the feedback with the environment has central importance. They will have to choose a combinatorial “mix” of information they receive either from the environment (industry), either from the employees (user-driven innovations) either from their personal motivation to innovate. Based on the information they receive and the expectations they have, they will adjust sales and employment for the following time tick.

The asymmetric information environment is represented by the probability of firing as well as by the coordination cost. The power to delegate of the leaders of each team, which normally would imply principal-agent issues \[92\], is done here efficiently and with perfect information - each team successfully completes the assigned subtasks. Therefore this model the agents have perfect information at the individual level and imperfect information at the mezzo level.

The difficulty of endogenizing or quantitatively defining the \textit{Tasks} (entrepreneurial opportunities, niches) comes from two facts: the market is essentially a highly uncertain environment and the entrepreneurs are highly subjective with respect to their actions. The market is characterized by ”the invisible hand” and it continuously escapes modeling efforts \[15,93\] and cannot be quantified from the top-down; it is macro-emergent by definition. On another hand, the incentives and risk profiles of the entrepreneurs are also highly heterogeneous and in fact, there is no equal matching between an entrepreneur profile \(i\) and a specific business opportunity \(j\). Many founders start a company before identifying the business idea. Only 21\% identify an idea and start a business at the same time \[16\] and only 33.2\% of new founders engage systematically in finding new business ideas, because they did not develop a business idea in a breakthrough moment, but it came over time \[16\].

In this model, there is no labor movement between entrepreneurs during runtime, as in Axtell’s model \[71\], but there is an infinite labor supply for the simulated company.

In the base model, the access to capital is also non-restrictive. While the access to capital is one of the 3 pillars of high-impact entrepreneurship \[1\], the motivation behind
not relaxing this assumption yet has some empirical support [16]. The personal wealth of the entrepreneur or receiving a windfall such as an inheritance does not increase the odds of starting a business. This means that, in general, overcoming the lack of initial capital is actually more of a perception than a real fact and the wealth effect is very small increasing the wealth 3 times increases the odds of a start-up by only 10% [16]. Only 1/3 of entrepreneurs ask for external funding when they start. Because they dont really grow their business, they dont need to raise capital from external sources. External sources and business growth are interlinked, as investors put money into businesses that grow: have cash-flows, sales and employees. The access to capital allows for companies to use either loans or investment to grow in sales and employment. The view on the development of the firm is closely related to an investment strategy - entrepreneurs that are "set" for growth rather than for generating income have already a business model for the expected growth of their company. The relationship between the entrepreneur and the way they envision the development of the business represents the high-impact phase of growth.

In many cases, the exit strategy of a company can be expensive and therefore the "exit" is done by continuously downsizing labor. Downsizing can be performed in 2 ways: by firing employees or by reducing their wage or time of working. In this model, the company "exits" by downsizing back to 1 employee. In reality, the exit strategy is a function of size - big companies that go bankrupt or go on IPO sale have exits that are costlier than small companies that file for bankruptcy or go into acquisition by another company. Bankruptcy is also costlier when liquidating than when debt reorganizing. In this model, while there is infinite access to capital, downsizing as an exit strategy is a function of the company size and current undertaken projects.

Competition and market are absent in this model of company growth. The base case model shows that both growth and no growth can be achieved in the absence of markets and competition. Competition does not seem to matter for start-ups; they tend to form in industries with a lot of existing firms [16]. Entrepreneurs rarely interact one with each other directly in the real life. In the same industry, they compete, in related industries
(by factors of production), they cooperate. They pay attention to price signals and other sources of information from the market (trends or "environment") [38] in order to adjust their employment and sales and in order to innovate. Industries that have the highest firm failure rates are also the ones that have the highest firm start-up rates and industries where it is easier to start a business are more prone to failure than others [16]. But an extension of the base case model with market and competition, restricted access to capital and costly labor turnover is desirable and represents one of the immediate next research steps.

4.5 Time tick/ Simulation and Activation Regime

In this model, time is not chronological, but a series of events and the activation regime is random. All the simulations described and analyzed here have a runtime of $t = 100$ time ticks. Each time tick represents an event for the company, such as completion of a project or completion of an administrative task.

Timestep represents one event. The time lags between time ticks are represented by coordination. The correspondence between time ticks and the real world could be done if we had a recording of the how much it takes for any company to complete a task and how much dead time there is between tasks is. This convergence time between tasks (coordination) is highly dependent on the managerial structure of the company, and thus on the team size (see the base case model above). Another difficulty in anchoring the virtual time into real time comes from the fact that the employees are paid by their skills, not by the hour.

For the calibration or sensitivity analysis of the model, there are two major difficulties posed by non-chronological time ticks: one is associated with time calibration, and another one is associated with phenomenological calibration [94]. The time calibration difficulty refers to the correspondence of the iterations with the real, physical time when the social phenomena to be analyzed is cyclical or punctuated (and basically non-linear). In this model, the time ticks are events, not time units (i.e. hours, months, ...) and the coordination costs are time lags incurred by teams working simultaneously in order to complete the task,
regardless of the duration in the real time as chronological (days, months) or dynamic (frequency).

The phenomenological difficulty comes from the non-dimensionality of some empirical-observations (magnitudes, intensities, territorial size) and from the correlation of the phenomenon with the time dimensions. This can be overcome by empirical statistics of the correlated phenomena [94]. In the case of this model, this is overcome through cross-validation (see section 4.7) and through the labor hours required to complete a task. As the tasks are randomly initialized with an expected labor hours requirement, the labor hours a better approximation of the real time necessary to complete a task than the time-steps are.

4.6 Input and Output: Preliminary Results

The model generates as output the company total revenue and the absolute employment growth at each time tick. Employment grows in steps, a phenomenon observed in the reality, while revenue grows continuously, which is also observed in reality. Only revenue growth describes the ”gazelles”, while both revenue growth and employment growth describe a high-impact company. In order to validate the model, a description of the types of output the model can potentially generate is helpful. The literature describes 4 levels of output in ABMs [95]:

- Level 0: The model is a caricature of reality, as established through the use of simple graphical devices (e.g., allowing visualization of agent motion);
- Level 1: The model is in qualitative agreement with empirical macro-structures, as established by plotting, say, distributional properties of the agent population;
- Level 2: The model produces quantitative agreement with empirical macrostructures, as established through on-board statistical estimation routines;
- Level 3: The model exhibits quantitative agreement with empirical microstructures, as determined from cross-sectional and longitudinal analysis of the agent population.
The agent-based model of High-Impact Companies based on team formation achieved the levels 1 and 2 of output.

**Level 1 Output:**

A recent report from Berkley and Stanford University that analyzed over 650 Internet start-ups in Silicon Valley measured "thresholds" and "milestones" important for the success of companies and innovations, as well as to propose a different method for assessing entrepreneurial performance [96]. Among their main findings are the following: that "founders that learn are more successful", that team of 2 founders grow or "scale-up" more quickly than one founder companies, and that teams that focus on business process than product are more likely to successfully scale up. Scaling and reproducibility in management or business processes are also important for Internet start-ups. [96]

![Figure 4.15: Scaling (employment growth stages) in Internet Start-ups, 2011](image-url)
Some first original results of the ABM showed that, independent of the team size, hiring is a step function and revenue is a continuous function.

Figure 4.16: The output in employment and revenue of one company the simulation that shows high-impact phase. In this case here, lw = 0.7, cc = 0.5, e = 3.
The figure above shows how a "typical" company generated during one simulation can replicate the 4 stages of scale described empirically by the Genome Project report. This output level is not satisfactory though for validating the model and replicating the growth of High-Impact Companies.

**Level 2 Output:**

The model also shows dynamically, during the 100 steps of the simulation, that some companies from 100 randomly generated can achieve spikes in employment growth at later stages of their evolution. There are 2% of companies that exhibit important changes in employment growth.

Figure 4.17: The employment change in 100 time ticks for 100 computer generated companies. (for these companies, the initial parameters were arbitrarily varied)
As found in the real data [1], these results also shows that the simulation can generate the high-impact phase independent of the original company age or size. This result is particularly encouraging in entrepreneurship research, where these is still persists the popular conjecture that only start-up companies can be high-growth, High-Impact Companies.

This level 2 output results are not calibrated though and it does not explain the role of the parameters in the evolution of these companies. It also does not have a sensitivity analysis to back up the results. Unless there is correspondence between the emergent patterns observed in the model and the processes which have led to the formation of these patterns, the model can be refuted. Only people can do pattern recognition and analogical reasoning [97], computers cannot, and only by ex post verification, validation and calibration the researcher can interpret and reason about what the computer shows. Therefore, the following verification and validation steps are performed (see section 4.7.).

4.6.1 Calibration

Calibration is a comparison between measurements in order to test for the correctness of results; it corresponds to the sensitivity analysis. Calibration has to be consistent and systematic, otherwise the exceptions due to randomness and multiple runs might be misinterpreted as coding artifacts. Calibration shows how much dependency there is between the output results and the initial conditions.

In order to calibrate the model, the model is run in independent simulations by manually tweaking the 3 main parameters \((e, cw, lw)\) for any values between \((0, 1)\), with the exception of the initial employee population size, where the ceiling was set to 30 employees (values between \(1 - 30\)). The experiments are run with 30 runs on each cell (same combination of parametric values) over 100 time ticks each.

While there is no correspondence between the simulation time ticks and the real time of a company life, although the period of 4-year specific for high-impact phase can be approximated by 30 consecutive time ticks. The average life of a company in US is 13 years [99], and the corresponding inference from here is that the 30 time ticks high-impact phase
in the simulation represents approximately 1/3 of the average lifespan of a firm.

From an output of 550 companies individually generated, 9 companies exhibited simultaneous growth in total revenue and employment size for a period of more than 30 consecutive time ticks. This represents 1.6% ratio of high-impact phase in the total population of companies generated, which closely replicates the high-impact rate provided by the real data (see section 3.1.).

From a larger output of 12,177 companies, 8,901 did not grow at all (73%) and the top 4% (486 companies) were responsible for 43% of employment growth in the entire population of companies and the 2% top companies were responsible for 35% of employment growth.

The 2% most successful companies at one time tick are plotted in order to look for the sensitivity to the initial conditions:

Figure 4.18: Sensitivity analysis. The maximum employment/tick as a function of lw and project failure for the top 2% performers. The companies on the x axis are sorted in the increasing order of their maximum employment growth/tick.
As project failure and lw are not a good predictor for maximum changes in employment independently, the regression analysis on the main parameters e, lw and project failure shows the dependance of the maximum employment values on these initial conditions:

![Figure 4.19: Regression coefficients of the maximum employment/tick on learning, founders size and failure](image)

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
</tr>
<tr>
<td>1</td>
<td>-1.371</td>
<td>9.133</td>
</tr>
<tr>
<td>No failed projects</td>
<td>3.327</td>
<td>6.23</td>
</tr>
<tr>
<td>lw</td>
<td>81.304</td>
<td>12.392</td>
</tr>
<tr>
<td>e</td>
<td>20.265</td>
<td>2.215</td>
</tr>
</tbody>
</table>

a. Dependent Variable: maximum employment

In order to test if there is a correlation between the learning weight and the number of failed projects, a parameter sweep in increments of 0.1 over the employee size, coordination cost and learning weight is performed, while keeping the probability of firing constant to 0.1 and the skill caps to 100,000. The data produced by the parameter sweep indicates how the variables of maximum changes in employment, learning weight, project failure and number of founders relate to each other and which are the conditions that are more likely to lead to maximum employment growth.

### 4.7 Verification and Validation

#### 4.7.1 Verification

Specifically for simulations and agent-based models, verification is a debugging step, that ensures that the model is correctly implemented and working as intended [98]. On another
hand, validation ensures that the behavior of the model does correspond to the behavior of the target [98]. Verification checks the internal validity of the actual program: does the program do what it is designed to do? The strategy used for verifying that this implemented model does what the conceptual model says it does is the following: debugging, verifying the random number generator and the activation regime. Particularly the activation regime (with each time tick) posed problems due to random number generators - the model initially generated only negative revenue tasks. In the final version of the model, the activation of the task is pegged to the size of the company and to the 40 hour work week. The tasks are not generated on a certain distribution, since there is no empirical data with respect to entrepreneurial opportunities. They are as random as the random number generators allows them to be.

4.7.2 Internal validity

Conceptual model validation determines that the theories and assumptions upon which the model was built are correct and the model representation of the problem entity is adequate for the purpose of the study [73]. Operational validation determines if the model output has sufficient accuracy to use the model for its intended purpose. The Levels 1 and 2 of output show that there is internal validity in the model.

4.8 External Validation

Validity is concerned with the correspondence between the model and the reality: is this a good model, does it correspond with the empirical data? Exact correspondence is not expected, as real social phenomena are uncertain and unpredictable, and the agent-based models may have stochastic processes or randomized variables, and this can pose a serious difficulty for the model. In open systems, model results are always underdetermined by the available data. [100] In this model, the external validation is done by comparing the base case model outputs with the real data, namely: the number of founders, the power law distribution of employment size and the Laplace distribution of the growth rates.
Empirical literature on the "desirable" team of founders that is more likely to lead to company growth and success points towards teams of 2-3 founders, that usually have complementary skills (such as one is a manager, another one is an engineer) [96, 101, 102]. This model replicates these survey findings (see Figure 4.20).

Figure 4.20: The number of founders that are more likely to lead to maximum employment changes and maximum employment size in the simulation.

The parameter sweep original employment size shows that simulate companies that start with 2-4 employees are more likely to lead to spikes in employment change, to high average employment growth rates as well as to high employment size at the end of the simulation. The model also replicates the power law distribution of a population of over 12,000 companies. The employment sizes of the simulated companies follow a distribution
of the following shape (see Figure 4.21):

Figure 4.21: The log-log plot of the frequency of generated companies and employment size in the agent-based model (binned in 100).

As shown in Chapter 3, the Laplace distribution tests, usually performed for the average growth of a company, cannot be performed in the case of High-Impact Companies. As seen in the chapter above, the employment growth quantifier is not an accurate measurement of the average growth and cannot be fitted on a Laplace distribution.
4.9 The Constructive Failure Hypothesis and Other Simulation Results

The goal of this dissertation is to provide an answer to the following major research question: What causes the high-impact phase of growth in certain companies? The analysis from the previous chapters points towards the following main hypothesis, labeled as the constructive failure hypothesis:

**Hypothesis 5.** The high-impact phase of growth of a company depends on both the organizational learning and rate of failed projects simultaneously.

While there is some "common sense" in the business circles and also some mass-media accounts about the importance of failure and learning from failure, there is virtually no academic literature with respect to failure in entrepreneurship. In order to test the hypothesis above, I am generating more than 4000 companies that are sensitive to the main parameters: $cw$, $lw$, $e$.

The study of the organizational learning curves [85] showed that the conventional form of a learning curve, $y = ax^{-b}$ can vary according to organizational "forgetting", labor turnover, transfer of knowledge and economies of scale. Regardless of the industry, the learning curve in an organization still falls into the 80-20 % ratio; organizational "forgetting" is particularly important for learning by doing and for the changes in the learning curves and it can be due to disruptions in productions (such as strikes, natural events or changes in products and processes).

Another important factor for organizational learning is the labor turnover in companies where labor is not standardized and cannot be easily transmitted from one person to another. But even in organizations with the same internal learning rates, the organizational learning rates might be different due to the transfer of knowledge from the outside (i.e. the market) [85].

101
Cumulative output is usually used for measuring learning by doing. Although learning curves have been found in many organizations, there is great variation in the rate at which organizations learn, due to forgetting, labor turnover, incomplete transfer within organizations and knowledge transfer across products and organizations. This ABM takes into account disruptions in the business processes (high coordination costs, failed projects) and has organizational learning rates.

The results of the statistical analysis of the growth of these companies are here described (see Figure 4.22)

Figure 4.22: The scatterplots of learning, failed projects and maximum changes in employment in 4000 generated companies. The variable Maximum represents the maximum employment/tick.
This figure shows some expected results: that companies that learn more grow more, a high rate of project failures does not lead to growth, but it also shows an unexpected result: that companies that had high employment changes also have a medium number of failed projects and a high learning rate. After a string of losses, most companies take on successful projects, which means that they learned fast enough from their initial failures. But companies that had only high learning rates or only a small number of failed projects did not significantly grow in employment. Learning weight alone or the number of failed projects alone are not predictors for company growth.

The following analysis shows how the founders size and learning or failure rate are conditions for high employment growth.

The constructive failure (or, in "common" terms, how much should you fail and learn before becoming successful) here means that the company growth is dependent on both the number of failed projects and the learning rate (see Figure 4.23).

These partial correlations show how strong the dependance on the number of founders is. They also show that, while learning and failure are inversely correlated, when they meet somewhere in the "middle", they are more likely to lead to the growth of a company. This is also visually depicted in Figure 4.22.

This analysis shows how the maximum employment changes in the simulation are dependent on the correlation between learning and failure. While this is not a perfect correlation or strong dependence, the results point towards verifying and not dismissing the main hypothesis advanced by this research.

As employees improve their skills with each time tick (individual learning) and teams have their own learning weights and coordination costs, the ultimate result of a company sustainable growth depends on the magnitude of the task it has to perform at each time tick. Even with infinite availability of human capital (skills) and continuous learning rate, the company that takes too much loss in consecutive projects will eventually fail. These are the companies that don’t learn from their failures. On another hand, companies that have a high learning rate and the a certain amount of failed projects are the ones that succeed.
Figure 4.23: The box plot (top) and heat map (bottom) above show that 1 founder is more likely to have more failed projects, while 2-3 founders are more likely to achieve growth in employment with a medium number of failed projects. The number of founders and the learning parameter are independent of each other.

and eventually stabilize as big companies that keep on generating revenue.

The constructive failure hypothesis is therefore sustained by this simple model. While
Figure 4.24: The correlation coefficients between learning, failure and maximum employment changes.

It is not a "black and white" answer to the question of how much learning from failure leads to success, it provides a likelihood or statistical answer to why this idea is important for the growth of successful companies. This hypothesis undoubtedly needs more testing, and some possible research extensions are presented in the following chapter.
4.9.1 Cross-validation

One possible proxy for overcoming validation problems is through cross-validation [103]. Essentially, cross-validation involves grounding both the assumptions, specifically the micro-behavior of the agents, and the observations of system macro-behavior, in real social systems. The skills and incentives of the agents and the team formation in this model are modeled according to the empirical facts described in the Chapter 2 and Chapter 3. On another hand, the emergence of the companies replicates the high-impact rate and the 2-3 founders rate given by the aggregated data with respect to high-growth companies. Both in the real world and in the model, the chance of an HIC versus a non-HIC is around 2%, depending on the parameters.

The survival rate of the simulated companies is of 38.625%. The rate of survival of newly formed businesses past 4 years in approximately 44%, across all industries. [104,105] But the average failure rate in the model versus the average growth rate in the model is 44% rate of failure of projects. In the model, 6.6% is the employment growth rate in the population of simulated companies, while in US the employment average growth rate is.

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td>-7.5563</td>
<td></td>
<td>-2.256</td>
<td>.141</td>
</tr>
<tr>
<td>FailedProjects</td>
<td>5.927</td>
<td>1.513</td>
<td>3.720</td>
<td>.000</td>
</tr>
<tr>
<td>LearningWeight</td>
<td>165.995</td>
<td>29.993</td>
<td>4.244</td>
<td>.000</td>
</tr>
</tbody>
</table>

a. Dependent Variable: MaxEmp1Change

Figure 4.25: Statistical analysis of employment spikes as dependent on learning and failure.
Chapter 5: Conclusions

“When you’re in a startup, the first ten people will determine whether the company succeeds or not.”

Steve Jobs, excerpt from "In the Company of Giants" [106]

5.1 Lessons Learned and Contributions to Entrepreneurship

This research tests the main hypothesis that companies that learn more from failed projects are more likely to become high-impact, high-growth companies. It also employs novel methods into entrepreneurship research, that are more suitable to understand quantitatively a social phenomenon that is, first and foremost, conceptually unclear across researchers and, secondly, has been researched mostly qualitatively. Innovation and production are only one aspect of growth and success for companies and not the most important one. While the production process is mostly about efficiency and optimization, scaling and growth are concerned with everything else in the company: the business process, the skills of the employees, the organizational learning, a.s.o. There is a performance versus adaptation trade-off [28] in complex social phenomena, and successful companies such as high-impact ones are both performing and adapting, just better than the rest of the companies. Innovation is not only about product and producing, innovation is also about management and team organization.

The research on education and leadership points towards the importance of learning from antecedents in order to replicate success [107]. The author introduces the Leadership for Learning Framework (see Figure 5.1), asserting that leaders that focus exclusively on results fail to understand their actions and cannot replicate their success; these are the “lucky” leaders. On another hand, leaders that understand their actions from past failures and achievements are the ones that become “leading” leaders and can replicate their success.
The entrepreneurs are not only "lucky fools", as described in the beginning of this research. They are learning and, more importantly, a few of them are leading their fields or industries.

Hierarchy alone will never make systemic change work [107] and organizational learning is crucial for how internal processes help a firm change and adapt. The question is how to move beyond islands of excellence, how to move from learning to leading [107] and to replicate success. The high-impact phase of growth is not sustainable. As this research showed, most of the companies grow rapidly only for a couple of years and out of 950,000 High-Impact Companies, only approximately 20,000 can sustain this growth (super HICs). This is a 2% rate out of a 2-4 % high-impact rate. Roughly, that is 0.04% rate of sustainable growth in employment and revenue at the macroeconomic level.

Peter Drucker intuitively answered in an older interview that the answer to this question
depends on adaptation, managers and team formation: the flexibility of team formation for outsourcing and for managers that cooperate, not control. [108]

For example, one form of organizational learning is changing the business model and re-branding the company. In this respect, organizational learning is a form of adaptation. On another hand, mass media has pointed towards new type of leaders as those who facilitate team formation. They are endowed with collaborative leadership and act as strategic facilitators for the overall success of a team.

5.2 Organizational Learning, Team Formation and Coordination Costs

Simon’s coordination costs [28] and Coase’s transaction costs [20] are conceptually similar: in order to achieve an agreement between two individuals or entities, there is a time lag or some sort of friction between agents before they align for that common, simultaneous goal. Particularly in organizations and groups of individuals, constant simultaneity is not possible. In essence, this is the departure from General Equilibrium theories, where agents perfectly coordinate. Even with perfect efficiency in the working environment, this research shows that departures from perfect coordination lead to asymmetric growth for the microeconomic agents.

One of the most important questions in economics and business literature is: what is the right combination of incentives and learning in order to achieve growth? [109] This research points towards incentivized employees that learn and contribute to organizational learning as well.

The basic socio-cognitive mechanisms, when carried out across and within large numbers of individuals, results in the co-evolution of social structure and culture. At the organizational level, the duality between action and knowledge leads to a co-evolution of the organizations form and the underlying knowledge networks, distributed cognition and transactive memory. [110] She also refers to an ecology of networks, which is a collection of
knowledge networks within and between agents, formed of interactions and by joint decisions. These interactions have also been studied broadly by Hannan and Freeman (1989), within their work on organizational ecology. They argued that in a social network, selection is more important than adaptation and that there is a high dependency on the network density and resource partitioning.

Friedkins study on informational flows [84] shows how important the ties between closed networks are. But the most important aspect that is neglected in networks research on teams is that networks often actively react to the ongoing process [82]. Also, if the information is valuable, individuals start searching through the network specifically or create new links and the problem of how agents and networks adjust to diffusion is still open [82]. Watts showed that 2 affiliation networks make them more searchable and closer to the small worlds phenomenon [111]. This implies that there are always information cascades around breakthrough innovations. Well connected networks have higher thresholds for new adoptions.

Teamwork simulation models can be divided in two main categories according to the purpose of the model: one that seek to develop artificial intelligence agent to interact with or substitute for human teams, and the second type of models seek to simulate human teams with the purpose of analyzing or improving them. [73] Particularly the second type of models, they contribute by simulating tasks and dependencies not in a static fashion, but dynamically and stochastically.

This dissertation proposes the idea of constructive failure as an equivalent of the "creative destruction" [21] concept at the firm level. The process of team formation and destruction within a company points towards the idea of clusters that predict high-impact events. At the macro level, there are economic crises and booms, at the micro-level there are high-impact phases and bankruptcies.
5.3 Future Research and Additional Considerations

While this research shows what causes High-Impact Companies to grow, it does not show what causes them to fail. An interesting future research can look into the causes of failure of high-impact growth and the unsustainability of their growth or, at least, what causes them to leave the high-impact growth. To date, there has not been a research done on what happens to High-Impact Companies after they leave their high-impact phase: do they become "elephants", as mentioned in the introduction?

At this stage, the model does not take into consideration finer aspects such as restriction of access to capital, the risk profile of the entrepreneur, market and competition and other external factors. Taking these into consideration may be a first step into extending this research.

Another possible future work would be to pool SBIR data together in order to find the learning weight and failed projects correlations. This could answer another important question, which is: do companies learn faster than individuals from their failures? An interesting derivation from this research might be found regarding a consistent relationship between the serial entrepreneurs and gazelle companies they founded: was this phenomenon reproducible in real life?

Also, a methodological question with respect to the difference between failure to failure and constructive failure to success can be answered by looking at this data. This could potentially provide an answer to the threshold of failure: how much failure and how soon?

Which are the demands associated with growth? Which are the costs these companies pay in order to scale up to high-impact phase of growth?

A more refined question with respect to organizational learning from teams concerns creativity. How does creativity emerge in teams and what constitutes the branding value of a company?

In the end, I believe that the research on High-Impact Companies points towards a method or methodology for measuring of entrepreneurial success. What distinguishes failure from success? How can success be replicated and what entrepreneurial endeavors do we
know for sure that will not work?

5.4 Final Conclusions

The study of entrepreneurship is currently one of the highly interdisciplinary areas of research in social science. Nevertheless, there has not been any clear agent based model of entrepreneurship published to date and neither an accurate analysis of who drives who: entrepreneurship causes economic growth or vice-versa? This research tackles a co-evolutionary approach entrepreneurs-economic growth by restricting the analysis to High-Impact Companies and employment growth. The main hypothesis that my dissertation tests is how the high-impact entrepreneurship is emerging and evolving and gives an answer to the common thought in the business circles that organizations that learn more from failures are more likely to succeed. By creating an agent based model of team formation and coordination, I hope that this dissertation achieved in the end a better understanding of what is the “essence” of high-impact entrepreneurship.
Bibliography
Bibliography


[66] Sheel, “feefighters.com.”


[99] C. Kennedy, “50 and still nifty: The average lifespan of a company today has been put at 13 years. who would have given much for the chances of those launched in the dark days of 1947?” The Director, 1997.


[109] A. Schleifer, “Seven things i learned about transition from communism.”


Anamaria Berea received her PhD from the Department of Computational Social Science, Krasnow Institute for Advanced Study at George Mason University. Her research interests have been focused on computational economics and philosophy of science, complexity science methods, asymmetric information networks, the emergence of atypical organizations such as hawala networks or the stateless tribes of AfPak area and on high-impact entrepreneurship. Currently she is working on applying social network analysis, agent-based modeling and GIS analysis to prediction markets and on experiments involving Bayesian network decompositions.

She also has a PhD, an MA and a BA in International Business and Economics from the Academy of Economic Studies in Romania and she is a member of the Romanian Society of Economics (SOREC) and the Center for Complexity Studies – UNESCO Center, Bucharest. Her publications span from policy issues related to European Union Accession to modeling conflict and social networks and she has teaching experience in comparative economics and economies in transition to junior and senior undergraduates in US.

She is an INET Young Scholar (2012), a proposal reviewer for INET and a paper reviewer for a couple of journals and conferences. She is also a recipient of the Phi Beta Delta International Honor Society for scholarly achievement in international education and a past recipient of Earhart Fellowship in Economics.

Simultaneously to her early academic experience, she has gained 5 years of business experience in an international environment. She is fluent in English, Romanian and French and has computer-programming skills in several languages.