Entrepreneurial Attractiveness: Amazon, Google, and the Search for Innovative Hot Spots

Anna Katherine Kropf

Dr. Charles Becker, Department of Economics, Faculty Advisor

Honors thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College of Duke University

Duke University
Durham, North Carolina
2018

Acknowledgements

I would like to express my gratitude to Dr. Charles Becker for the excellent guidance he gave in the production of this paper. His time, advice and profound interest over the entire course of the research are greatly appreciated.

Abstract

Recent economic literature suggests that entrepreneurship in technological fields can spur economic growth, making it a popular topic for city development officials. Yet, this increasingly popular phenomenon is met by many economic questions. One of those questions is which characteristics of metropolitan areas are attractive to entrepreneurs. To answer the question of attractiveness on both the small business and corporate levels, I compare across two case studies: Amazon's search for a second headquarters and Google's tech hub network. Using principal component analysis, I statistically deduce seven components of attractiveness from an original 34 variables. These components are then weighted using three methods—a case study, a survey, and an empirical method—to produce comparable indices of attractiveness. Generally, I find that sizeable population and healthy economy are the strongest components. However, the statistically insignificant components that can change an urban area's ranking considerably are talent and geographic network effects. Ultimately, creating policy to maximize these aspects can change a city's innovative trajectory.

JEL Classification: O, O3; R, R1, R11

Keywords: Innovation, Entrepreneurship; Growth; Index of Entrepreneurial Attractiveness; Principal Component Analysis; Undergraduate Survey; Metropolitan Statistical Area

I. Introduction

The heterogeneity amongst cities poses a problem for economists studying the characteristics that make urban areas entrepreneurially attractive. According to the Global Entrepreneurship Development Institute (GEDI), "entrepreneurial success does not take place in a vacuum. Entrepreneurs exist in the context of their particular geography – be that their local, national, or even supranational economy and society" (GEDI, 2017). The question remains as to which types of magnets—government, private sector, social, economic, and demographic—attract innovators to cities. Further, are the magnets homogeneous across cities even while the cities themselves are not? In other words, I seek to determine if there exists a "formula" for entrepreneurial attractiveness despite the varying strengths and weaknesses across metropolitan areas. While there is no "perfectly entrepreneurial" city, I argue that there is indeed a set of characteristics that equips cities to use entrepreneurship effectively, generating an ecosystem for entrepreneurship. The ecosystem then attracts future growth to the location.

For the purpose of this paper, entrepreneurship is defined as the production and distribution of new and incremental technologies in all types of markets. New and incremental technologies are effective when they spur economic growth and development through job growth, technological improvements, and productivity growth. Therefore, the term "entrepreneur" refers to any party conducting a new venture, whether a small business or large corporation. I use two case studies in this paper, which discuss Google and Amazon as they look for cities in which to invest substantial resources.

The choice of two large corporations over individual entrepreneurial ventures is deliberate. I have identified Google and Amazon as two companies that look for cities based on their innovative and entrepreneurial attractiveness, but with two different goals in mind. First, Google invests resources in startup communities across the globe, looking specifically for innovative potential.

Second, Amazon's search for a new headquarters illustrates the characteristics relevant to its corporate decision, which involves looking for innovative power. In other words, while Google pursues an initiative adjunct and distinct from its core business, Amazon wants to best position its core business. Therefore, Google's choices serve as a proxy for individual entrepreneurial ventures, because it is difficult to track individual entrepreneurs' locational decisions; Amazon is a proxy for corporate decision-making. Using case study discussions of these two corporations and various statistical tools, I produce a combination of city characteristics that predicts attractiveness. Ultimately, I observe the difference between small-scale entrepreneurial and large-scale corporate innovative attractiveness.

There is one caveat concerning corporations in the context of entrepreneurship: whether or not the negatives of their arrival outweigh the benefits. There are potential downsides for the city in which Amazon chooses to locate, some more detrimental than others. These concerns are popular in the media given Amazon's looming choice, but in reality these concerns apply to all corporate geographical moves. First, in "re-emerging" urban areas, one of which Amazon is likely to choose, there is great concern over gentrification and soaring housing costs. According to Leana Garfield of Business Insider, Amazon's move would detonate what she refers to as a "prosperity bomb" in the housing market (Garfield, 2018). The shift of young millennials—especially those that Amazon attracts and hires—to the urban core would increase the magnitude and speed of gentrification. Further, new development and higher incomes would increase housing and rent prices in general for all demographics. Other externalities include increased traffic, construction, and congestion. Seattle has seen each of these effects—increased construction, traffic, and housing prices—since Amazon's arrival, so it is expected that Amazon's new city would as well (Garfield, 2018). Ultimately innovation is important to economic growth, but the method of its advancement can bring negative economic externalities.

Several entrepreneurship indices have been published across various units of analysis in an attempt to compare entrepreneurial ecosystems. Yet, without a framework defined by transparent metrics applicable to cities across the world, extensive city-level analysis remains incomplete. In this paper, I provide a framework for analyzing why cities are entrepreneurially attractive to two technologically innovative corporations. This framework then leads to a weighted indexed ranking of the cities targeted in my research. In addition, the paper compares several weighting methods to provide final conclusions. First, examples of three different cities—discussed as part of the Amazon and Google case studies—are targeted in statistical analysis to show which strengths cities possess that affect their ranking. Second, because companies are made up of employees, it is important to understand employee locational preferences. Surveys of soon-to-be college graduates determine which qualities of cities are most attractive to the employees themselves. I then provide an empirical analysis of the data, using regression results to establish a final ranking.

The following section provides background on the topic of my research. I also review previous literature concerning entrepreneurship, including indexed rankings that apply different data techniques from those used in this paper. Sections III and IV analyze Google's tech hub network and Amazon's search for a new headquarters, respectively. As innovative corporations, understanding which city characteristics they find attractive provides insight as a foundation for a broader look at entrepreneurship. Section V then discusses the variables included in an index of entrepreneurial attractiveness, synthesizing past research, case studies, and data analysis. In section VI, I present a statistical model, using principal component analysis (PCA), that delineates the variables and provides numeric terms. Section VII discusses the results of PCA, analyzes the weights used to produce an index, and presents the framework for an objective ranking of metropolitan areas across the United States. Finally, Section VIII discusses geographical effects as an important finding in the research. Section IX concludes.

II. Background and Literature Review

Across the globe, economic growth often relies on the institutional, political, and social policies of the nation in question. Yet, it is generally agreed that the growth of national income and the subsequent increase in quality of life is favorable to citizens and politicians. Government policymakers should seek economic policies that maximize their cities' growth potential, particularly through the pursuit of effective innovation and entrepreneurship. Modern research suggests that entrepreneurship can cause beneficial outcomes, including increased productivity, job growth, innovation, overall economic growth, and other positive externalities (GEDI, 2017). Therefore, optimizing city-level factors to create an attractive climate should be conducive to new ventures. By city-level factors, I refer to the aggregated business, government, and demographic data that point to successful innovation. New ventures in the technological fields particularly contribute to accelerated economic growth trends.

Entrepreneurship studies span microeconomic and macroeconomic models. At the microeconomic level, businesses seek to maximize profit according to underlying demand and cost functions. Location of the business affects costs, the labor market, and synergies in ways that affect profit maximization. Governments, households, and firms in the aggregate maximize their respective goals, inducing macroeconomic responses. The market responds with changes in investment, employment, inflation, and aggregate output. For macroeconomists the goal is long-run growth, especially significant when translated into growth of per capita income. According to macroeconomic growth models, overall growth can only be supported in the long run by technological advancement paired with investment, not by capital accumulation in the economy alone. Hence, innovation plays an important role in both the micro and macro economies. Applying elements of these fields to individual cities provides insight into the locational aspects of economic growth.

These broad economic models provide the foundation on which to examine how cities choose a policy mix to support entrepreneurship. Important factors to observe include human capital, which grows primarily in the form of university education; technological progress, which contributes to local, national, and global progress; and output, which increases as an economy grows and part of which is reinvested into further growth. Ultimately, establishing the ways cities can maximize these factors is important to civic development. It is by using this analysis that governments at all levels can apply appropriate policy and incentives for future growth that ultimately result in higher standards of living for their population.

Past research provides strong support of the importance of continued entrepreneurial research due to its supposed effect on the growth of an economy. There currently exist several entrepreneurship indices that compare the entrepreneurial environments across cities, states, and countries. First, the Kauffman Foundation has produced the Kauffman Index to rank 50 states and 40 metropolitan areas across the United States in three categories: growth entrepreneurship, startup activity, and main street entrepreneurship (The Kauffman Foundation, 2016). In the 2016 iteration of the growth entrepreneurship index, Washington, DC ranked first, followed by Austin, TX, Columbus, OH, Nashville, TN, and Atlanta, GA. Their 2016 startup activity index ranks Miami, FL first, followed by Austin, TX, Los Angeles, CA, San Diego, CA, and Las Vegas, NV. Lastly, their 2016 main street entrepreneurship index ranks the top five as Pittsburgh, PA, Boston, MA, Portland, OR, San Francisco, CA, and Washington, DC. A second index of importance is the Global Entrepreneurship Index (GEI) produced by the Global Entrepreneurship Development Institute. It measures the health of each of 137 countries' entrepreneurship "ecosystems," using measures such as attitudes, abilities, and infrastructure (GEDI, 2017). The 2018 iteration of the GEI ranked the United States first, followed by Switzerland, Canada, the United Kingdom, and Australia.

Along with the data provided by these foundations, the literature provides insights concerning the characteristics of cities most important to economic development, entrepreneurial effectiveness, and business performance. Having analyzed the effect of various economic, demographic, and institutional qualities of cities on their overall economic environment, these sources serve to provide direction as I consider the most important characteristics. While the studies included here are not exhaustive, they paint a picture of a successful entrepreneurial environment.

First, Etzkowitz (2008) proposed that three independent entities—universities, government, and industry—function interdependently in the form of a three-fold force to affect a city's innovative entrepreneurial environment. This three-fold force is modeled by a triple helix. Government affects the regulatory environment. Universities produce knowledge, human capital, and networks. Then, industry provides the businesses that invest financial capital and ultimately produce output in the markets. In a highly complex relationship, each of these three entities interacts with the others to influence innovative effectiveness. According to Steiber and Alänge (2013), there appears to exist a collaborative combination that produces the most effectiveness, but it is difficult to measure due to the numerous actions involved within each of the branches.

Cities stand to benefit from corporate involvement, as well as public and private sector collaborations. For example, during the 1950s three universities in North Carolina—Duke University in Durham, North Carolina State University in Raleigh, and University of North Carolina at Chapel Hill—ventured together into the creation of a research and technology park, called the Research Triangle Park, to spur growth and development of North Carolina's economy (Cirillo, 2013). Geographic proximity between universities and companies brings opportunities for shared resources and ideas. These ideas then transition more easily into the private market (Wessner, 2009). Further, bringing international companies to the area, such as IBM, put the Research Triangle on the map as an innovation center. Ultimately, attracting and supporting enterprise—whether from small,

large, international or domestic companies—plays a major role in investment and growth trajectories.

Each of the triple helix components is discussed in depth by other researchers. First, Delgado, Porter and Stern (2010) discussed the vitality of agglomeration. Second, Porter (1990) discussed the importance of concentrated knowledge and resources. Third, Audretsch, Lehmann, and Warning (2005) discussed the role of universities. Porter (1990) reported that while cities are an important unit of analysis, many entrepreneurial decisions affecting these environments stem from the national level. At the national level, the innovative environment is created, with a look toward international business and relations. Because of a national tendency toward business abroad, cities must also enable their companies to use international business as part of their business strategy. Business abroad refers to accessing both international consumer markets and supply chains. Therefore, cities must provide access to the larger global markets for production, distribution and information attainment. Delgado, Porter, and Stern (2010) emphasized the importance of cluster effects in the entrepreneurial environment due to "the presence of a strong cluster environment, which reduces barriers to entry and growth and enhances regional comparative advantage" (p. 498). Increased resource efficiency, returns on investment, and access to capital are several of the benefits of agglomeration. Audretsch, Lehmann, and Warning (2005) stressed the importance of university effects within a city's limits due to the concrete benefits of proximity to a university's research efforts, money, and personnel. University effects are bounded geographically; the cost of accessing spillover effects increases as distance from the university increases. Therefore, it is important for universities to be proximal to—or better yet, involved with—the city.

Finally, Reyes, Roberts, and Xu (2017) found there are factors that weigh directly on the firm in positive or negative ways, while there are others that have no effect, despite assumptions suggesting otherwise. The study regressed observed factors against job creation and labor

productivity. They find that agglomeration, access to capital, and infrastructure produce some of the most prominent effects on the dependent variables, while regulations and tax code have little to no effect. These observations are important to consider when choosing determinants of entrepreneurial attractiveness.

In order to synthesize the determinants collected in this research, several common statistical techniques are applied, namely principal component analysis (PCA) and subsequent ranking of the scores. Ye and Mikic (2016) used these methods to produce an index of Asia-Pacific Regional Integration Effort. The integration index consists of a social sub-index and an economic sub-index, designed to be uncorrelated with each other. PCA determines the weights of each input, which then assigns an index score to each country. Vyas and Kumaranayake (2006) also constructed a socio-economic index using similar techniques.

III. Google's Tech Hub Network

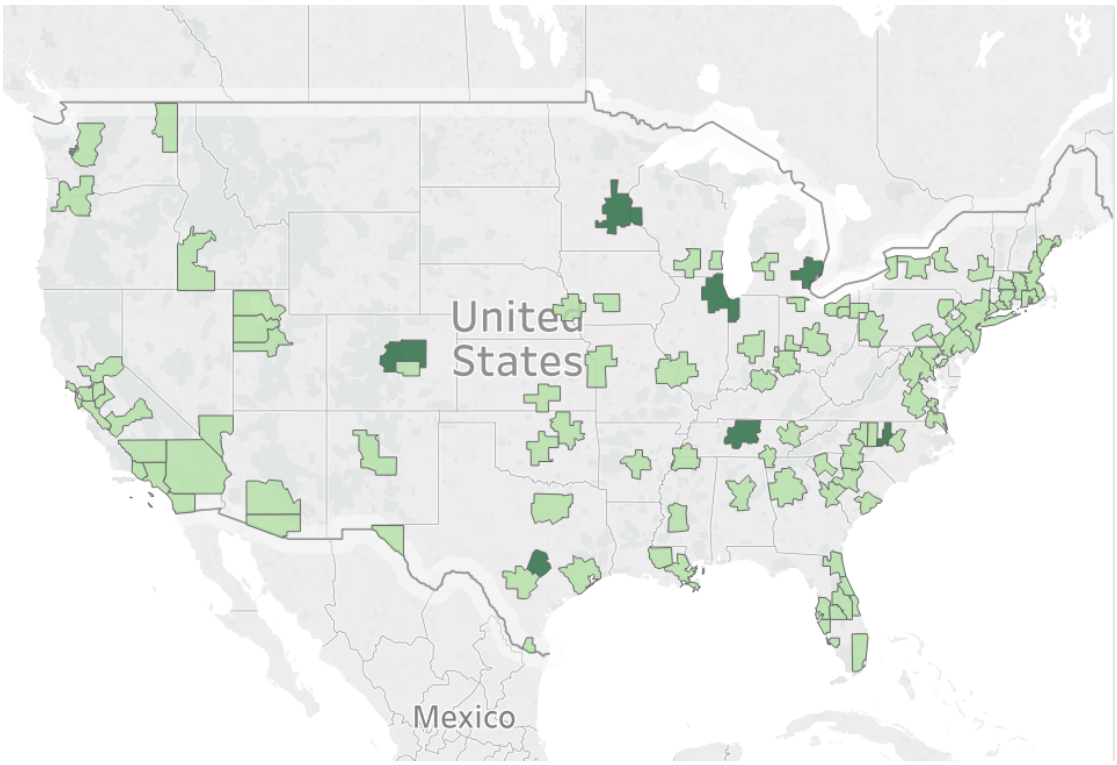
Google's entrepreneurial initiative, Google for Entrepreneurs (G4E), was established in 2011 as a way to expand technological development in regions outside of Google's backyard in Silicon Valley (Google Inc., 2011). Parallel with its company tradition and mission, Google focuses primarily on technological growth in its community outreach. While many cities across the nation and the world are engaging with entrepreneurship in a beneficial way, Google's immense resources greatly benefit the tech startup communities within reach of its influence. According to Mary Grove, Head of Global Entrepreneurship Outreach at Google, Google for Entrepreneurs plans to create a web of partnerships that can foster strong entrepreneurial connections across the world, provide programs that can train entrepreneurial minds, and introduce Google products as accessible tools for entrepreneurs (Mulvey, 2012). In return, the company receives the marketing benefits of guiding these regions to a prospering small business and tech scene (Procopio, 2013). While the extent to

which Google's entrepreneurship "goals" are executed is unproven, the identification of tech hubs across the U.S.—and the world—is significant to a city's entrepreneurial attractiveness. Inclusive communities, global networks, learning programs, and open office spaces are four foci implemented by its city support efforts.

One of Google's specific networking programs is the North America Tech Hub Network, inaugurated in 2013. This network consists of Google-sponsored startup communities across the United States and Canada that form an interconnected, thriving entrepreneurial community. Google provides the cities with events, training opportunities, and technical infrastructure to minimize costs and maximize resources for new companies. The ten tech hubs include American Underground in Durham, North Carolina; The Capital Factory in Austin, Texas; The Entrepreneur Center in Nashville, Tennessee; COCO in Minneapolis, Minnesota; Communitech in Waterloo, Ontario; Galvanize in Denver, Colorado; Grand Circus in Detroit, Michigan; Notman House in Montreal, Quebec; 1871 in Chicago, Illinois; and Centraal in Mexico City, Mexico. Each of these cities was chosen specifically by Google as a location worthy of resources to support their entrepreneurship economies. As is the common thread among the cities, Google made their choices based on the existence of already existing and growing tech networks. These cities would benefit from resources to advance innovation further. Google then benefits from their continued success by sharing in the publicity or having direct access to buy out a start up's idea. The metropolitan areas where these members of the tech hub network reside are displayed in Figure 1.[1]

[1] Canada and Mexico locations are not presented in the map because they are not included in the data portion of this research. The darkly shaded areas indicate Google's tech hub network locations, while the lightly shaded locations designate all other metropolitan areas included in this research.
All graphs and figures in this paper are produced using Tableau®.

Figure 1. Map of Google's Tech Hub Network



Each of the cities on Google's list has a strong network of entrepreneurs. Google desires to be a part of these communities through mentoring programs, global outreach, individual startup success, or training a new generation of tech geniuses. According to Adam Klein of American Underground—one of the Google for Entrepreneurs tech hubs located in Durham, NC—Google searches for the "best in class tech hubs doing work outside of the Bay and NYC (usual tech hot spots) that they could support" (Klein, personal communication, 2017). This qualification results in the choice of cities primarily on reputational characteristics, where Google may be attracted due to some of a city's specific strengths, rather than generalized metrics. Several of these qualities include strength of the creative and tech class, social network amongst entrepreneurs, and existence of entrepreneurial community. Granovetter (2005) applied social network theory to competitive advantage, highlighting the strength of weak ties and the power of social connections in fostering new innovations, formations and growth. These phenomena are at work in entrepreneurship-prone

cities, where breadth, ability to connect, and sharing of resources amongst a community of entrepreneurs fosters beneficial economic effects. Cities attractive to entrepreneurs have not only the financial capital resources, but also the human and social capital resources to place their startups in the best position for success. It is important to understand these characteristics because a look at each city in detail provides insight into the observable characteristics important to city attractiveness. The following subsections look at a few of the Google tech hub cities, and what makes them stand out as entrepreneurial.

III.1. Durham, North Carolina

Durham has a university culture intertwined with the success of the city, enabling it to succeed through highly concentrated knowledge and resources. Not only does Duke University own or contribute to the majority of office space and real estate development in the city, but it also conducted research and development (R&D) expenditures of nearly $1.1 billion in 2016, ranked tenth in the nation, with another $1.0 billion by the eleventh ranked University of North Carolina-Chapel Hill only a few miles away (National Science Foundation, 2015). Further, Durham's share of creative class employment is ranked first at 48.4% (Florida, 2012). These are the people driving the downtown culture, especially through the interaction of the entrepreneurial community with its surrounding restaurant and entertainment scene.

With Research Triangle Park (RTP) in close proximity, large amounts of capital are invested into research-specific ventures. These investments have been traditionally corporate, but more recently have spilled over into smaller entrepreneurial ventures. The presence of incubators and accelerators in both Durham and RTP facilitates the movement of resources. In the span of five years, the startup community, American Underground, has grown from 90 startups to 275 (Klein, 2017). In terms of the entrepreneurial ecosystem, Durham has the necessary elements: university

involvement, strong enterprise, and a creative network of small businesses.

III.2. Denver-Boulder, Colorado

Denver, Colorado hosts a high-tech training facility for aspiring coders, computer science hobbyists, and data engineers—a haven called Galvanize. The facility is essentially its own college, producing graduates in high-tech fields. It is also a place for these techies to launch their creations. Though it began in Denver, the Galvanize community extends beyond the city, into Boulder and distant entrepreneurial cities like Austin and San Francisco. Boulder's innovative technology and presence of tech companies has earned it one of Google's "eCities" awards (Galvanize, 2017). Along with over 100 startups in the community, it is home to the Denver branches of large tech companies like IBM, Uber and Box. Further, 200 of Google's own employees work out of Boulder, Colorado.

Denver, commonly recognized for its outdoorsy-ness, attracts the millennial interested in a nature-based lifestyle that other cities cannot offer. Boulder has a "bike-to-work" mentality common among these millennial entrepreneurs. The amenities paired with an advanced tech scene provide an attractive and unique lifestyle for young, creative techies.

III.3. Austin, Texas

Austin is another example of an attractive entrepreneurial ecosystem. The Capital Factory in Austin, Texas has strong ties to the city's government, Google, and the University of Texas at Austin. Austin's entrepreneurial community receives support from its own Mayor and a recently founded Innovation Office. The Innovation Office runs government sponsored programs including accelerators, workshops, and consulting (The City of Austin, 2018). While the initiative is primarily to increase civic innovation, it stimulates communication between the government and business. The connection has allowed the mix of private and public sector resources to promote Austin's own

local startups. This relationship has also been infiltrated by Google itself over the years. For example, the establishment of Google Fiber was a joint effort of government and startups forming a relationship with Google (The Capital Factory, 2017).

The culture of Austin is also conducive to an innovative environment. According to Kevin Johns, the Director of the City of Austin's Economic Development Department, the city has created for itself a community of eager and interconnected entrepreneurs (Pollard and Frauenheim, 2016). In such a community, sharing resources is commonplace as opposed to stealing resources. The city has a young population, also flooded by students of UT Austin. The university's highly regarded graduate programs include the 17th ranked McCombs School of Business and the Cockrell School of Engineering (U.S. News, 2018). Ultimately, Austin's attractiveness boils down to talent and culture.

Google's tech hub network is more homogenous than the cities on its list appear to be. The theme among these tech communities is young, talent-infused millennial populations who enjoy a lifestyle where live, work, and play become one. Creative industries, including technology industries, require communities where the restaurant and entertainment scene is intertwined with the business culture. The cities on Google's list, and the tech hubs they host, are changing the business environment from a corporate-dominated model to an interactive model between start-up, corporation, and city.

IV. Amazon's Search for HQ2

In September 2017, Amazon, Inc. announced its plans to invest over $5 billion into a second headquarters—Amazon HQ2—to complement its current Seattle headquarters (Amazon.com Inc., 2018). Requirements for the location of Amazon's new headquarters include size and resources, talent, and infrastructure. The announcement has received great attention and has spurred a

competition between cities across the United States and Canada to prove their worth. For cities, Amazon's announcement is significant; a headquarters for a company that has the magnitude and innovative power of Amazon is expected to bring thousands of jobs, billions in investment, and a plethora of opportunities to wherever they choose to locate. The official decision is expected mid-2018.

According to Amazon's announcement, HQ2 hopes to positively affect the employment opportunities, the community, the environment, and the entrepreneurship prospects of the city it chooses. Through its own product, Amazon believes they "empower small businesses to grow and seize the opportunities of the digital economy" (Amazon.com Inc., 2018). In Seattle's community, Amazon has spent billions in capital investment, employee compensation, hotel and travel expenditures, public transportation, and local retail that stimulate and grow the economy (Amazon.com Inc., 2018). Indirectly, Amazon claims Seattle has gained approximately 53,000 non-Amazon jobs, $38 billion in investment, and $17 billion in local income increases as a result of their presence (Amazon.com Inc., 2018). It boasts that the same benefits will impact the location of HQ2, including added benefits from construction, hiring, and publicity.

Using unconventional methods to find their new city, Amazon has asked city representatives across the nation to submit proposals for Amazon to consider, bringing a public interest and excitement to the decision. Calling it "The Project," HQ2 has an initial timeline planned for over 15 years, but Amazon expects to bring more than 50,000 jobs to the chosen city with an average employee compensation of over $100,000 (Amazon.com Inc., 2018). While Amazon's description of its project, preferences and goals is extensive, several of Amazon's requirements include:

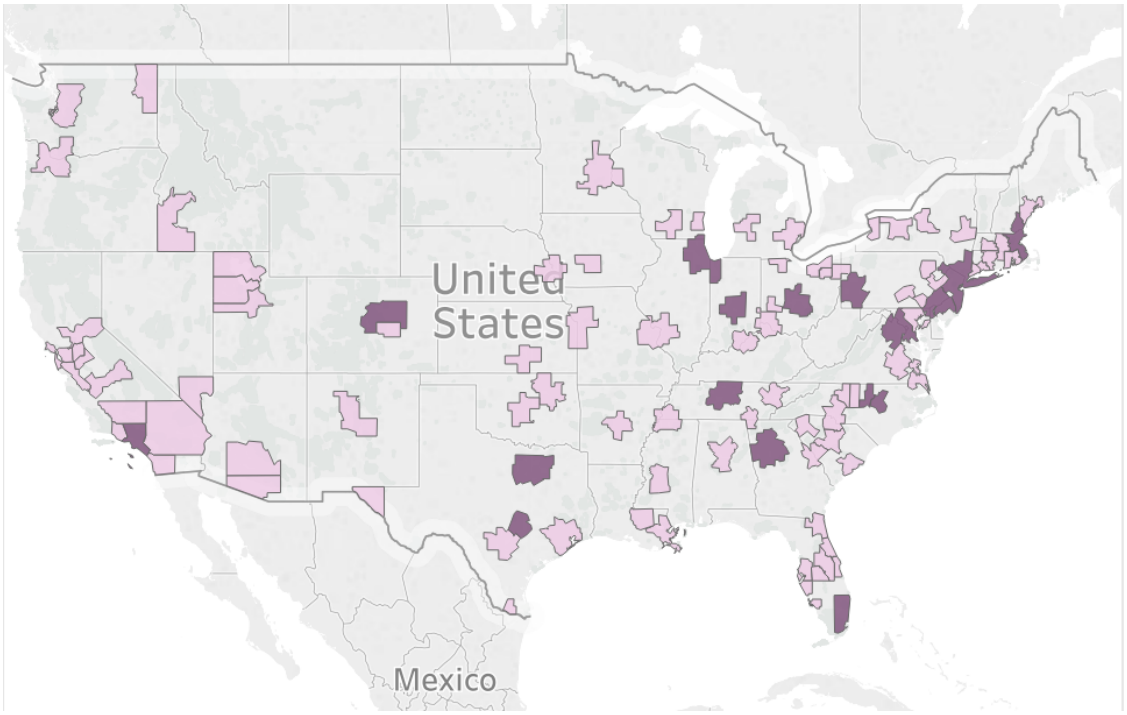
1. A metropolitan population greater than one million;
2. A business-friendly and economically-stable environment;
3. Strong local talent, including a strong university system; and

4. Creative real estate options – specifically, located within 30 miles from the urban center, 45 minutes from an airport, 2 miles from major highways, and with over 500,000 square feet of available space, but room to expand to near 8,000,000 square feet (Amazon.com Inc., 2018).

On January 18th, 2018, Amazon released the names of 20 cities that would continue to vie for the headquarters, cutting the list down from 238. These cities are Atlanta, Georgia; Austin, Texas; Boston, Massachusetts; Chicago, Illinois; Columbus, Ohio; Dallas, Texas; Denver, Colorado; Indianapolis, Indiana; Los Angeles, California; Miami, Florida; Montgomery County, Maryland; Nashville, Tennessee; Newark, New Jersey; New York, New York; Northern Virginia; Philadelphia, Pennsylvania; Pittsburgh, Pennsylvania; Raleigh, North Carolina; Toronto, Ontario, Canada; and Washington, District of Columbia. For the purpose of this study, the list translates into 16 distinct metropolitan areas, which are shown in Figure 2.[2] Newark and New York City combine into one metropolitan area: New York-Newark-Jersey City, NY-NJ-PA. Northern Virginia, Montgomery County, and Washington DC also combine into one metropolitan area: Washington-Arlington-Alexandria, DC-VA-MD-WV. Toronto is left out of the study due to data discrepancies.

[2] Toronto is not presented in the map because it is not included in the data portion of this research. The darkly shaded areas indicate Amazon's short list locations, while the lightly shaded locations designate all other metropolitan areas included in this research.

Figure 2. Map of Amazon's HQ2 Short List



Looking at the characteristics, successes and dynamics of metropolitan areas when weighing attractiveness, there is complexity underlining locational decisions made by companies like Amazon; while one city offers an attractive quality, another city offers a different yet equally attractive quality. Therefore, Amazon must decide which characteristics are most important, weighing the opportunity costs. To understand the dynamics in these decisions qualitatively, the next three subsections discuss three cities that are possible options for HQ2.

IV.1. Philadelphia, PA

The metropolitan area of Philadelphia has a population of over six million, with a median age of 38.6 and a majority white demographic (U.S. Census Bureau, 2017). In terms of education, 36.7% of its population holds a Bachelor's degree or higher, compared to the United States metropolitan average of 32.5%, and 44.5% of the Bachelor's degrees are in science-related fields

(U.S. Census Bureau, 2017).[3] The city's GDP per capita is above average at approximately $63,000 (U.S. Census Bureau, 2017).

Philadelphia is attractive for several reasons. First, it is a city with a strong network of entrepreneurs, but even more prominently, the city has a large and diverse workforce. It is home to a variety of Fortune 500 companies including Aramark and Comcast, as well as multiple pharmaceutical, financial, and information technology corporations. Many domestic and international companies, such as IKEA, Boeing, ING, SAP, and PECO, have headquartered important divisions of their business in Philadelphia. The university system is also prominent, housing a top research Ivy League university, University of Pennsylvania, along with Drexel University, Temple University, and Villanova University. Philadelphia has a weak public sector, most prominently struggling fiscally for several years due to its problematic budget and tax structure (Mooney, 2016). Further, its technology sector is lacking compared to other metros like Durham and Austin. Nevertheless, its strong business environment excites the potential for growth booms in the near future.

Second, Philadelphia is a transit hub, located optimally near other populous cities in the Northeast; for example, the city is less than 100 miles from New York City and 140 miles from Washington DC (Google Maps, 2017). Such proximity to other major commerce hubs offers a highly versatile business environment. Finally, the cost of living in Philadelphia is approximately 35% less than that of living in New York City (Numbeo, 2017).

IV.2. Austin, TX

Amazon has chosen Austin, Texas for similar reasons that Google did, as mentioned previously. Austin is an up and coming technological hub, with a bustling web of accelerators,

[3] The average metropolitan educational attainment of 32.5 is based only on cities with populations greater than 500,000.

incubators, and entrepreneurs. With a fairly young—median age of 34.4—and diverse population of over two million, the city is dominated by a culture that enables a strong innovative environment to emerge (U.S. Census Bureau, 2017). To match Texas' traditionally friendly business environment, Austin's public sector is highly supportive of small business ventures. Further, University of Texas at Austin is ranked highly among universities, offering highly skilled personnel to contribute to the tech environment. Nearly 43% of the population holds at least a bachelor's degree, of which 44.2% hold a bachelor's degree in science-related fields (U.S. Census Bureau, 2017).

Austin's population enjoys a high quality of life, primarily due to a low of cost of living. The city's cost of living including rent is approximately 40% lower than that of New York (Numbeo, 2017). Cost of purchasing daily goods is also 40% lower than that of New York (Numbeo, 2017). The GDP per capita is approximately $61,000 and growing (U.S. Census Bureau, 2017). One negative is that Austin lacks the close physical proximity to other major business hubs that other cities boast, particularly Northeastern cities like Philadelphia.

IV.3. Research Triangle, NC

Compared to Google's focus on Durham alone for its tech hub, Amazon prefers to consider the entire area encompassing three cities: Durham, Chapel Hill, and Raleigh, NC.[4] This area—called the Research Triangle—is home to the Research Triangle Park (RTP), a large tech and science research park. Uniquely, each of the three cities has a strong university that collaborates with the other two. All three universities—Duke, UNC-Chapel Hill, and NC State—are historically ranked among the top 50 universities in the United States (Collier, 2017). Therefore, locating within or near the Research Triangle offers the benefits of a large network of university and business resources, with a prime geographic location among other fast-growing and innovative east coast cities. The

[4] Durham-Chapel Hill, NC is one metropolitan area, and is used as such in this research, rather than as two separate cities as noted here for descriptive purposes.

Research Triangle Park alone is home to IBM Corporation, Cisco Systems, Fidelity Investments, Biogen, Lenovo, and RTI International, among others (The Research Triangle Park, 2017). Further, biotechnology and information technology are fast-growing industries in the region, attracting many brilliant minds. Approximately 47% of both Raleigh's and Durham's populations hold Bachelor's degrees; in Durham, 53.8% of those degrees are in science-related fields, and in Raleigh, 50.1% of them are science-related (U.S. Census Bureau, 2017). While their populations combined are only around two million, the tech and university talent is significant. The small and less dense populations reduce access to the cultural and recreational activities that would interest young professionals, though these lifestyle aspects are improving rapidly.

Added benefits of this region include a growing per capita GDP of nearly $69,000 in Durham—only $54,000 in Raleigh—and an overall low cost of living (U.S. Census Bureau, 2017). Including rent, the cost of living in both Durham and Raleigh is 45% less than that of New York (Numbeo, 2017).

V. Determinants of Entrepreneurial Attractiveness

Despite the individual needs companies and entrepreneurs consider when weighing cities in terms of attractiveness for their new ventures, the characteristics important to the entrepreneurial environment are quasi-universal. Using prior literature, the Google and Amazon case studies, and other sources of insight, I translate the research into a list of quantitative variables. The categories of variables used in this research are demographics, accessibility, quality of life, and performance, business environment. These groups are determined prior to PCA, only separated into categories to indicate the variety of variables used in the statistical analysis. The eventual PCA components do not reflect these specific categories.

First, demographics tell a story about the city and its economy. I use eight demographic inputs: population, median age, diversity, race segments, poverty, unemployment, millennial educational attainment, and technical educational attainment. Large populations are valuable to entrepreneurs as they host "the scale and diverse array of offerings needed to attract talent" (Florida, 2014). This includes recreational and cultural activities that increase quality of life. Essentially, the larger the city the more there is to do. Diversity and median age are also important indicators of the culture within a city. Entrepreneurs benefit from a combination of college graduates' knowledge and the varying perspectives that come from a diverse population. Diversity is measured by percentage of population that does not identify as only Caucasian. Also included are measures of racial constitution. Though companies are advocating minority representation, employees prefer communities homogenous to their selves. Both Google and Amazon are majority Caucasian and Asian, so a measure of the Asian population in a given city is included. Further, it benefits the company to locate in a city where diversity is high while keeping in mind that a downtrodden inner city minority population will not benefit the company's workforce. Unemployment and poverty rates, also economic performance metrics, are entered to control for economic heterogeneity and social problems. Educational attainment, which is measured by percentage of the population aged 25-34 with bachelor's degrees, is an important indicator of the talent within a city. Millennial education is applied instead of general education because tech companies are not concerned with the older population's degrees. It is the recent college graduates and young adults whom Amazon is hiring. The number of degrees in the science and technology fields further represents educational attainment.

Second, three variables are important indicators of accessibility and connectedness. Airport size and amount of airport traffic indicate the level of access to the city by plane. Airport size is measured by the Federal Aviation Administration's (FAA) hub size descriptions—small, medium,

and large—then translated into indicator variables. Airport traffic is measured by commercial enplanement data also from the FAA. The third variable is a simple measure of a city's location within a geographic network, called GEONETWORK. Two factors enter into the measure: distance to other metropolitan cities and strength of the nearest metropolitan cities. The variable is calculated according to Equation 1 below. The calculation involves a regression that determines initial ranks for each city within the network, *a*, denoted by RANK. Initial rank is determined according to a regression of all metros against Amazon's short list. I use the Amazon qualification as opposed to a variable that combines both Google and Amazon under the assumption that Google's tech hubs are decided on a single city's talent rather than the potential benefits of spillover effects. The city in question is referred to as *m*. DISTANCE refers to the distance between the two cities in miles; only distances of less than 150 miles are entered into the calculation. One hundred fifty miles is the maximum distance for what can be considered a manageable "day trip" where the time to travel to a city is approximately three hours. A trip of any longer would require a flight, where ultimately the distance is of little consequence.

$$GEONETWORK_m = \frac{RANK_{a1}}{\ln DISTANCE_{a1}} + \frac{RANK_{a2}}{\ln DISTANCE_{a2}} + \cdots + \frac{RANK_{an}}{\ln DISTANCE_{an}} \quad (1)$$

The third determinant of entrepreneurial attractiveness is quality of life. Lifestyle amenities are an important part of the urban lifestyle, as are cost, safety, climate and traffic. Recreational activities are indirectly captured by the population variable where size implies a greater number of various events, entertainment opportunities, and communities. The quality of life inputs include rental rate, price to income, general quality of life, commute, climate, and crime. The rental rate per square foot of a median single-family residence is one representation of the cost of living. I use rental rates as opposed to home values assuming that many of Google and Amazon's employees will

be renting, as is the trend among young graduates. Nonetheless, rental rates and home values are highly correlated so only one is necessary as an input. The price to income ratio represents rental affordability relative to income. It is calculated as average US apartment square footage of a one-bedroom apartment—750 square feet—multiplied by the rental rate and divided by per capita income (Otet, 2016). Interestingly, the rent to income ratio is insignificant in the data, indicating that cost of living is often built into the salary. General quality of life aims to capture details that influence quality of life and is therefore an indexed measure that includes purchasing power, cost public good availability, pollution, and safety. Commute time is the average number of minutes it takes for workers over the age of 16 to commute to their place of work. Climate is measured by an index of climate likability that takes into account temperatures, humidity, and extreme weather. Crime is measured by an index that takes into account crimes committed and feelings of safety.

Fourth, I measure performance, in terms of city talent and innovation capabilities. Also relevant to performance is education and business environment, which are discussed in other paragraphs. These measures include number of technology conferences the city hosts within a given year, patent filings, gross domestic product per capita, and gross domestic product growth. The patent variable is a per capita variable in order to show the relative amount of innovation taking place without population considerations. It is measured as number of patents filed during a given year, per capita. Gross domestic product growth is a ten-year growth rate from 2006 to 2016, to eliminate the effects of business cycle. Gross domestic product per capita, measured in US dollars ($), is adjusted for purchasing power parity and inflation.

The fifth determinant includes the size and quality of the business environment and is represented by 13 inputs. Number of establishments, amount of annual payroll, and total employment for the technology sector from the US Department of Labor's Business Dynamics

Statistics constitute three of the variables.[5] Three more variables are the ratio of the tech sector's establishments, payroll, and employment to the whole economy. These measures represent how prominent the tech sector is in a given city's economy. Next, four variables represent employment movement. Both change in technology employment and overall employment are included as 10-year growth rates, from 2005 to 2015. Further, job creation is an important indicator of the health of the business environment, where total job creation and per capita job creation are included for a 10-year period from 2005-2015. The number of per capita establishment entries for a given city in a one-year period represents the general entrepreneurship happening within the city. The rate of firm startup and survival is also representative of entrepreneurial activity. High firm entries suggest a startup-heavy environment, and death rates can provide insight into the progressive and regressive elements of entrepreneurship. Therefore, the ratio of firm births to deaths is a twelfth variable. Lastly, clustering effects are a key for the success of business ventures of all sizes. As discussed previously, a high concentration of industries—though not too high so as to drown out small businesses—is a welcome source of resources. One way to measure these effects is by the level of interconnectedness and linkages between industry, which displays how highly valued are collaboration and shared resources. The U.S. Cluster Mapping Project provides percentage of business interconnectedness—how often industries rely on one another—as a metric of business cluster effects. The U.S. Cluster Mapping Project is a collection of open data led by the Harvard Business School and funded by the U.S. Economic Development Administration.

[5] The "technology sector" refers to NAICS Sector 54: Professional, Scientific, and Technical Services.

VI. Data, Methodology, and Model

VI.1. Data Set

The data set includes cross-sectional data on 65 variables from 2016. These include 54 numeric variables, six descriptive variables, five binary variables, and one indicator variable. Thirty-four of the 65 total variables are used in the statistical analysis.[6] Several of the variables are sourced earlier than 2016 due to lack of up-to-date data, including patent data from 2013, cluster strength data from 2015, business statistics data from 2015, and establishment data from 2014. The set consists of each of the 104 metropolitan areas in the United States with populations greater than 500,000.[7] Metropolitan areas are used rather than urban areas or municipalities because the entire suburban area surrounding an urban area is vital to its functionality and livability. The demographics and businesses existing in the outskirts of the city area are significant to its success. Two extra entries are included to reach a total of 106: one is Denver without Boulder and another is the combined Raleigh-Durham to represent the Research Triangle. These two extra cities add further analysis into city proximity synergies, which are discussed in Section VIII. The metropolitan statistical area raw data come from the US Census and their American Community Survey, US Gazetteer, US Bureau of Economic Analysis, Business Dynamics Statistics, the Federal Aviation Administration, US Department of Justice, and the US Bureau of Labor Statistics. Other sources include Zillow Research, Harvard Business School's Cluster Mapping Project, Open Data Network, Google Maps, and Numbeo Indices. See Appendix 1 for an overview of the variables. Using several data sources that are inconsistent with the time constraint of 2016 does not affect the results of the data in terms of selection bias, due to the consistency of using same years for each variable across all

[6] Nearly half of the original 65 variables are excluded from statistical analysis because they describe the same condition as another variable in the data set. For example, size is described by two variables, *size* and *sizekm*, but they are simple linear transformations of one another.

[7] There were 104 cities with 2016 populations greater than 500,000 at the time of data collection. There may be a fewer or greater number of cities with populations greater than 500,000 in the present-day and future population counts.

metropolitan records. Further, using different data sources that use different methodologies also does not introduce bias.

The natural logarithms of *population*, *sizekm*, *annualpayroll_all*, *annualpayroll_tech*, *employment_all*, *employment_tech, establishments_all*, *establishments_tech, estentries*, and *estexits* are taken prior to the statistical analysis. The logarithm is a more appropriate representation of the effect of these variables. Assuming the distribution is homoskedastic in a linear multivariate regression when it is actually heteroskedastic violates the normal distribution assumption. Therefore, I adjust several skewed variables using logarithms to approach homoskedastic errors. For example, when considering population, a difference of tens of thousands will weigh more on an entrepreneur's decision when he or she is comparing cities with populations less than one million than those with more. This relationship is not linear, and therefore would weigh incorrectly on the regression's error term if not appropriately transformed.

VI.2. Methodology

Principal component analysis (PCA) is used to reduce the data set of characteristics of metropolitan cities to several more easily identifiable dimensions.[8] It is particularly important in reducing multicollinearity, given the high correlations between variables in the data set. Each principal component, PC, is created according to a weighted linear combination of the variables, X_i. This initial set of variables involves *n* correlated variables—X_1 to X_n—that PCA combines and reduces to uncorrelated components. The weights, w, are determined according to the explanation

[8] Principle component analysis (PCA) is a statistical technique that decomposes the correlation matrix of a series of variables in a data set. PCA transforms the set of observations using linear transformations into a set of principal components. These components are now linearly uncorrelated. The components with the greatest eigenvalues indicate where there is the greatest amount of variance in the data set. In the case of this data set, PCA reduces 34 dimensions to seven significant components with eigenvalues greater than one.

of the variance, where w_{mn} represents the weight for the *m*th PC and *n*th variable. Equations 2 and 3 describe mathematically the principal component model.

$$PC_1 = w_{11}X_1 + w_{12}X_2 + \cdots + w_{1n}X_n \quad (2)$$

$$\vdots$$

$$PC_m = w_{m1}X_1 + w_{m2}X_2 + \cdots + w_{mn}X_n \quad (3)$$

Because of the high correlations amongst the components, and that multiple variables are used for multiple components, an oblique rotation is used (Ye and Mikic, 2016). The seven components that explain the most variance are kept while the other 27 are discarded. The produced components are then scored individually.

Using a logistic regression of the seven components on a dependent variable, I determine which components are statistically significant in determining whether a city is considered as a potential tech hub or headquarters. Two binary variables indicating Google's tech hubs and Amazon's short list are analyzed, as well as one that combines the two overall. The significant components from the regression are assigned weights that sum to one according to their relative strength, and then applied to establish a weighted index.

After weighting the components according to the regression results and several other methods that will be discussed later in the paper, composite scores are produced for individual metropolitan areas. These composite scores are then ranked to present the potential entrepreneurial attractiveness of each included metropolitan area relative to the others. The variables, which were described in detail previously, describe the demographics, economic health, infrastructure, cost of living, and business environment of each of the 104 metropolitan areas.

VI.3. Missing Values

Due to missing values of several cities, the number of observations is reduced to 99 in total. In this study, metropolitan areas with missing values are excluded; because there are so few total observations, using means for the missing data has the potential to skew the results. These metropolitan areas are Albuquerque, NM, Greensboro-High Point, NC, Lakeland-Winter Haven, FL, North Port-Sarasota-Bradenton, FL, and Youngstown-Warren-Boardman, OH-PA.

VI.4. Preliminary Statistics

The preliminary data summaries are included in the Appendix. Appendix 2 presents a summary of the descriptive statistics for the list of variables. Appendix 3 presents the correlation coefficients to describe the relationships among the variables. The correlation matrix is instrumental in the determination of the PCA eigenvalues, and correlation is used rather than covariance due to the lack of standardization of units.

VI.5. Limitations

One of the most prominent limitations of this study is the inability to describe quantitatively each of the factors that contributes to entrepreneurially attractiveness. The 34 variables are comprehensive and cover the majority of characteristics relevant to entrepreneurial attractiveness. Nevertheless, attempting to thoroughly describe a phenomenon such as entrepreneurship through quantitative data is troublesome given the extensive qualitative aspect, including individual entrepreneur subjectivity and preference. Data often do not describe precisely the conditions necessary for entrepreneurship, nor do variables for these certain conditions exist comprehensively over the set of cities under examination. For example, airport data do not discriminate whether people are visiting for tourist or business purposes. This is where future research can expand on the

current problem, in order to more comprehensively characterize entrepreneurship. Numbers only tell a fraction of the story when it comes to describing a city.

A second limitation of the study is the scope of the data set. A sample size of only 99 cities is a small *n* by statistical standards. Unfortunately, the low sample is unsolvable. I have chosen not to include cities with populations less than 500,000 because these are not in the running for Amazon's headquarters. Including them would skew the results. Though the number of data points is limiting, the comparison does span all cities that are necessary for the analysis. Given my statistical goals, this number of cities is therefore sufficient for accuracy.

VII. Empirical Results and Discussion

I use principal component analysis (PCA) to reduce the set of variables to principal components. The technique returns seven components as significant, and together they capture 80.62% of the variance. The components are described in Table 1, including their statistical descriptions. Appendix 4 lists the component correlations according to the variables and the calculated weights determined by PCA. Oblique promax rotation with Kaiser-normalization is applied due to the high correlations among the variables in the data set and their contributions to the components (Ye and Mikic, 2016). The use of PCA is suitable due to the high chi^2 value of 4445.93 and p-value approaching zero for the test of sphericity.

The determination of the components suggests several relationships among the data. First, component one contains all variables correlated with population, including airport traffic, payroll and employment data, tech conferences, and commute times. Also correlated are per capita GDP, educational attainment, and diversity. Interestingly, high rental prices are correlated with these populous cities, but this does not also indicate a high rental to income ratio. Presumably, income levels are satisfactory given the rental rate. Component two contains significant variables that

suggest a crowded and dynamic economy that can be described in part by a tourist economy. It is positively related to large population, large amount of commercial enplanements, and large number of tech conferences hosted by the city. Also, quality of life is low for residents, including high amount of crime, long commutes, low GDP growth, and high unemployment. Compared to component one, there is a high rent to income ratio, which indicates low income overall despite the high prices. Component three describes the effects of talent infrastructure and networks. This component is scored according to correlations between geographic network location, educational attainment, GDP, and age. Talent in this case is associated with affluence. Negatively correlated with those variables are change in employment, job creation, and establishment entries, implying that some of the enterprise associated with resident talent is not retained by the city.

For component four, which describes a skilled workforce, percentage of science degrees in the city is high, despite low population and low number of total establishments. Another relevant variable captured by component two is a high Asian population. Further, the city has high patent output as well as high interaction between industries indicated by cluster effects. Component five describes inaccessible cities that are unable to drive growth. Economic growth, job creation, patent output, and education are low. Component six describes cities in an economic downturn, where crime is high, prices are low, and establishment entries are low. These cities also have high diversity, which could indicate an urban core with suffering minorities. Finally, component seven indicates cities with entrepreneurial tech sectors, but that are highly isolated from other cities.

Table 1. Description of Components

	Descriptive Name**	Top 5 Metro Areas	Lowest 5 Metro Areas	Eigenvalue*	Percent of Variance*
Component 1	Size	New York, Los Angeles, Chicago,Washington DC, San Francsisco	McAllen, Winston-Salem, Modesto, Augusta, Palm Bay	7.34	21.58%
Component 2	Tourism Economy	San Jose, San Francisco, Honolulu, Boston, New York	Cape Coral, Deltona, Winston-Salem, Augusta, McAllen	4.29	12.63%
Component 3	Talent Infrastructure and Network Effects	Washington DC, San Jose, Durham, Raleigh, Baltimore	McAllen, Memphis, Fresno, Toledo, Grand Rapids	4.18	12.30%
Component 4	Skilled Workforce	Austin, Houston, Provo, Dallas, Salt Lake City	Detroit, Dayton, Cleveland, Springfield, Birmingham	3.81	11.19%
Component 5	Inability to Drive Innovation and Growth	McAllen, Fresno, Stockton, Bakersfield, Memphis	Madison, Denver, Portland, Boston, Albany	2.82	8.29%
Component 6	Economic Downturn and Lack of Revitalization Efforts	Deltona, Stockton, Cape Coral, Riverside, New York	Omaha, Tulsa, Madison, Des Moines, Durham	2.81	8.27%
Component 7	Isolated Tech Growth	Miami, Palm Bay, Denver, Boise City, Las Vegas	Hartford, Springfield, Worcester, Philadelphia, Baltimore	2.16	6.36%

*The values for Eigenvalue and Percent of Variance are after Kaiser rotation
**The Descriptive Name is based on the significant variable components

Before establishing a final index, I analyze how an assumed model responds to various cities, and how choosing a city to be ranked first effects the changing of the weights. Using weights determined through trial and error experimentation, I attempt to bring the three top choices for Amazon's headquarters discussed in Section IV—Philadelphia, Austin, and the Research Triangle (represented by Durham)—to the top of the list.[9] I choose weights for each of the components based on the component scores of city in question. For example, because Durham scores high in talent and skilled workforce, I weight this component higher than the others. The weights offer a perception of the targeted city based on the component strengths as well as the cities surrounding it at the top of the ranks. It is also clear from these three cases that the weights given to the components severely alter the results. Therefore, it is important that the weights are as accurate as

[9] Raleigh, NC and Durham-Chapel Hill, NC are two separate cities in our data set, so Durham represents Research Triangle Park in this exercise for several reasons. First, for this portion of the research it is not plausible to use the "Research Triangle" metro that was included in the data set; using Durham allows for consistency and comparability throughout this section. Second, Durham's ranking is more significant to the research because Google has also chosen it as a tech hub. Further, because of their similarities, when Durham is ranked first on the list, Raleigh is ranked in the top five.

possible as we solve for an overall index. Equations 4, 5, and 6 are the mathematical descriptions of each index composition for each city.

$$\text{Index}_{\text{Durham}} = -1.8C_1 - 0.2C_2 + 3.0C_3 - 0C_4 + 0C_5 - 0.2C_6 + 0.2C_7 \quad (4)$$

$$\text{Index}_{\text{Austin}} = 0.2C_1 + 0.2C_2 + 0.2C_3 + 0.8C_4 - 0.4C_5 - 0.8C_6 + 0.8C_7 \quad (5)$$

$$\text{Index}_{\text{Philadelphia}} = 1.2C_1 - 0.2C_2 - 0.2C_3 - 0.2C_4 + 1.2C_5 - 0.6C_6 - 2.2C_7 \quad (6)$$

For Durham, a first place ranking requires talent infrastructure and networks at 300% but size of city at negative 180%. Further, tourism economy is weighted at negative 20%, as is economic downturn. Isolated tech growth, though Durham is not isolated in this sense, is weighted positively because Washington, DC's geographic network is extremely strong. It is weighted positively solely to eliminate Washington, DC's advantage in locational strength. Given these weights, in the top five behind Durham are Washington, DC, Raleigh, NC, San Jose, CA, and Palm Bay, FL.

For Austin to be ranked number one, its two most prominent qualities are weighted the highest. These are tech workforce at 80% and isolated tech growth at 80%. The remaining weights are assigned as follows: talent infrastructure and networks at 20%, tourism economy at 20%, size at 20%, inability to drive growth at negative 40%, and economic downturn at negative 80%. Denver, CO, San Jose, Dallas, TX, and Salt Lake City, UT follow Austin in the top five.

Philadelphia requires significantly different weights than Durham and Austin to reach a number one ranking, showing how different the city is from the two up and coming tech forces. Philadelphia requires an isolated tech growth ranking of negative 220%, because its most significant strength is its location. Its distance to major economic centers including New York and Washington DC allow it to score high in geographic network. Both talent networks and tech workforce is weighted at negative 20%, implying Philadelphia's lack of a tech sector. Size is weighted 120%, but

tourism is weighted negative 20%. Even though Philadelphia scores very low for inability to drive growth, it is weighted a positive 120% to combat New York's even lower score. Lastly, economic downturn is weighted negatively at 60%. New York, NY, Hartford, CT, Chicago, IL, and Washington, DC rank below Philadelphia with these weights.

In constructing an overall index, I make several hypotheses based on research and analysis concerning the component weights. These assumptions apply to all forms of entrepreneurial attractiveness, not distinguished between Amazon and Google. First is the importance of relatively inexpensive living for the younger population, which includes both high incomes and low rental prices. Interestingly, PCA revealed that a measure of rental price as a percentage of income is insignificant to the components. Nevertheless, high prices are still significantly correlated with large cities. While size and dynamism of an economy cannot be discredited fully, the high rental prices, low purchasing power, and long commutes associated with these large economies is a negative factor that likely drives entrepreneurs elsewhere. Second, talent and skilled labor force is characterized by high per capita income, high educational attainment, and high presence of a technology and professional sector relative to the whole economy. These qualities are captured in components three and four. Third, economic health is important to the development of innovation. Finally, I assume that geographic networks are an important consideration, particularly for Amazon. According to these assumptions, the following Equation 7 is hypothesized.

$$\text{Index}_{\text{hypothesized}} = 0.5C_1 + 0C_2 + 1.2C_3 + 0.5C_4 - 0.3C_5 - 0.4C_6 - 0.5C_7 \qquad (7)$$

Size is given a weight of 50%, tourism a weight of 0%, talent networks a weight of 120%, and tech workforce a weight of 50%. The negative components are inability to drive growth at 30%,

economic downturn at 40% and isolation at 50%. Table 2 shows the rankings according to these weights, and indicates which cities are Google and Amazon cities.

Table 2. Hypothesized Indexed Ranking of Metropolitan Areas

1	Washington-Arlington-Alexandria, DC-VA-MD-WV	*Amazon*
2	Austin-Round Rock, TX	*Amazon, Google*
3	Baltimore-Columbia-Towson, MD	
4	Houston-The Woodlands-Sugar Land, TX	
5	San Jose-Sunnyvale-Santa Clara, CA	
6	Durham-Chapel Hill, NC	*Amazon, Google*
7	Raleigh, NC	*Amazon*
8	Denver-Boulder, CO	*Amazon, Google*
9	New York-Newark-Jersey City, NY-NJ-PA	*Amazon*
10	Boston-Cambridge-Newton, MA-NH	*Amazon*
11	Dallas-Fort Worth-Arlington, TX	*Amazon*
12	Atlanta-Sandy Springs-Roswell, GA	*Amazon*
13	San Francisco-Oakland-Hayward, CA	
14	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	*Amazon*
15	Charlotte-Concord-Gastonia, NC-SC	
16	Salt Lake City, UT	
17	Chicago-Naperville-Elgin, IL-IN-WI	*Amazon, Google*
18	San Diego-Carlsbad, CA	
19	Seattle-Tacoma-Bellevue, WA	
20	Madison, WI	
21	Los Angeles-Long Beach-Anaheim, CA	*Amazon*
22	Tampa-St. Petersburg-Clearwater, FL	
23	Hartford-West Hartford-East Hartford, CT	
24	Minneapolis-St. Paul-Bloomington, MN-WI	*Google*
25	Colorado Springs, CO	
26	Harrisburg-Carlisle, PA	
27	Richmond, VA	
28	Indianapolis-Carmel-Anderson, IN	*Amazon*
29	Miami-Fort Lauderdale-West Palm Beach, FL	*Amazon*
30	Bridgeport-Stamford-Norwalk, CT	
31	Portland-Vancouver-Hillsboro, OR-WA	
32	Baton Rouge, LA	
33	San Antonio-New Braunfels, TX	
34	Phoenix-Mesa-Scottsdale, AZ	
35	Sacramento-Roseville-Arden-Arcade, CA	

36	Oxnard-Thousand Oaks-Ventura, CA	
37	Columbus, OH	
38	Kansas City, MO-KS	
39	Orlando-Kissimmee-Sanford, FL	
40	Albany-Schenectady-Troy, NY	
41	Provo-Orem, UT	
42	Pittsburgh, PA	*Amazon*
43	Worcester, MA-CT	
44	Detroit-Warren-Dearborn, MI	*Google*
45	Cincinnati, OH-KY-IN	
46	Virginia Beach-Norfolk-Newport News, VA	
47	Charleston-North Charleston, SC	
48	Ogden-Clearfield, UT	
49	Des Moines-West Des Moines, IA	
50	Nashville-Davidson-Murfreesboro-Franklin, TN	*Amazon, Google*

Next, I assign weights according to the opinions of undergraduate students in science and technology fields. I have conducted a survey-based study of 51 Duke undergraduate and graduate students from Pratt School of Engineering and Trinity College of Arts and Sciences. The study was designed to receive direct feedback from these students as to how each component should be weighted. It is relevant because current college students in certain fields are the future of companies such as Amazon, as well as future entrepreneurs in tech and innovation. The survey used for this study is presented in Appendix 5. The participants of the study were 53% female and 47% male, as shown in Figure 3. Further, Figure 4 presents the races of the students that participated. While 62% of the participants are Caucasian, minority representation includes Asian participants at 18% and mixed race participants at 14%.

Figure 3. Ratio of Male to Female, Entrepreneurial Attractiveness Survey

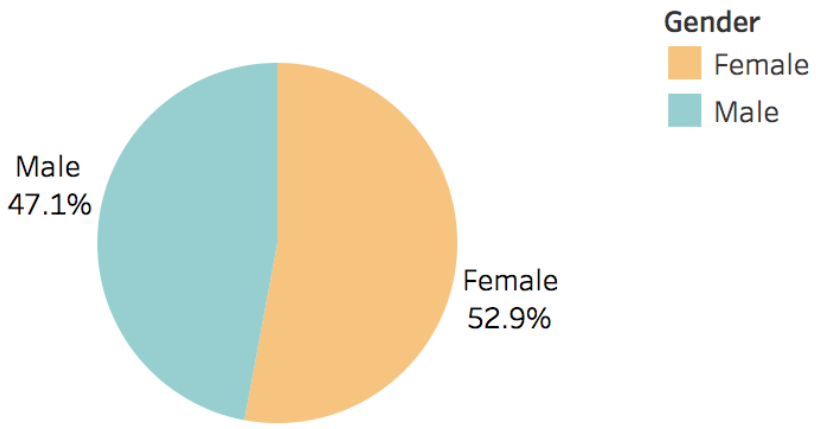


Figure 4. Racial Representation, Entrepreneurial Attractiveness Survey

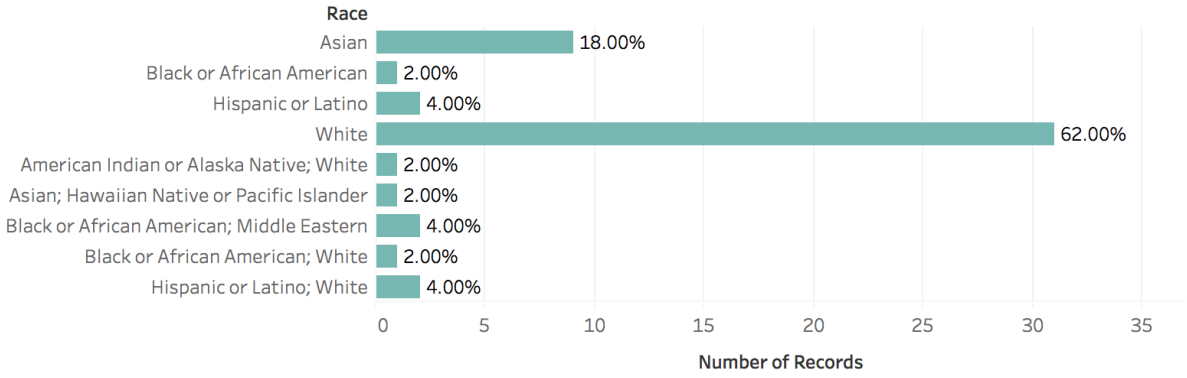


Figure 5 presents the results for whether or not students involved in the study can see entrepreneurship as part of their future careers. The high rate of entrepreneurial-minded students involved in this study—nearly 60%—adds value to the study's results. The participants also come from several different academic backgrounds. The majors represented include Psychology, Economics, Political Science, Public Policy, International Comparative Studies, Cultural Anthropology, Linguistics, History, Visual Media, Environmental Studies, Evolutionary Anthropology, Biology, Neuroscience, Electrical Engineering, Civil Engineering, Biomedical Engineering, Mechanical Engineering, Computer Science, and Engineering Management. These

majors have been further categorized by field, into Engineering, Hard Sciences, Social Sciences, or Visual Arts. The distribution of fields is presented in Figure 6.

Figure 5. Response to Possible Entrepreneurial Ventures, Entrepreneurial Attractiveness Survey

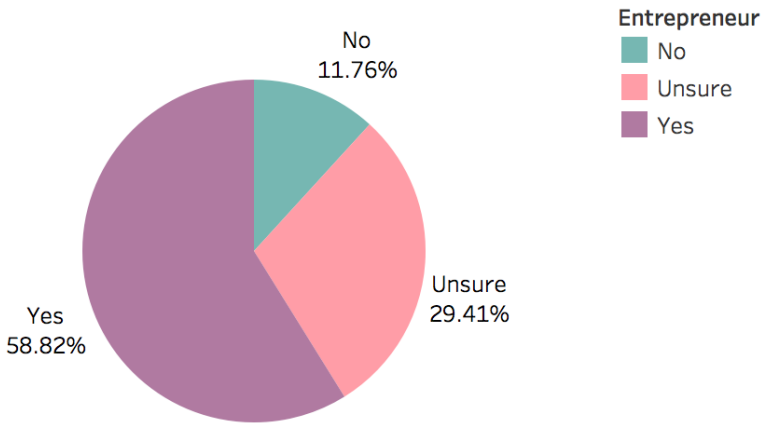
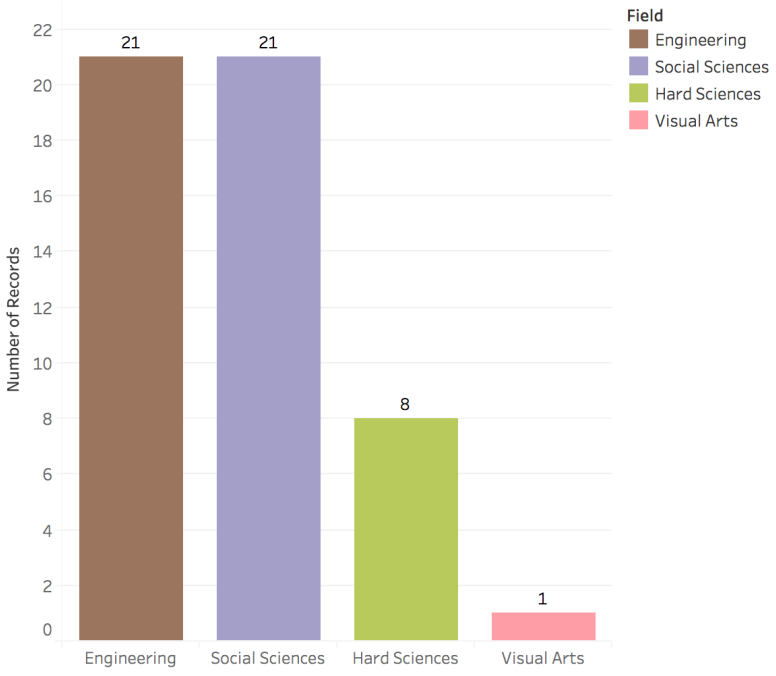


Figure 6. Distribution of Academic Fields, Entrepreneurial Attractiveness Survey



To determine the weights from the study, I averaged the responses of the students for each component. Students were asked to weigh each component—for which detailed descriptions were given—according to their personal preferences, with scores that add up to 100. Along with an aggregate weight, several other demographics' weights were averaged separately to present a more detailed view of preference for certain demographics. The study revealed the weights presented in Table 3.

Table 3. Weights Determined by the Survey of Entrepreneurial Attractiveness

	Size	*Tourism*	*Talent*	*Workforce*	*Small Tech*	*Economy*	*Isolation*
Total	**33.03**	**14.25**	**30.33**	**33.17**	**-2.85**	**-6.02**	**-1.91**
Female	31.85	11.26	27.22	31.78	2.48	-1.07	-3.52
Asian	27.16	21.50	21.40	30.50	1.00	-5.56	4.00
Minority	31.76	19.18	36.38	31.08	-3.38	-10.94	-4.08
Science and Engineering	26.61	24.10	25.66	27.00	-2.00	-2.74	1.37
Social Sciences	42.52	0.86	37.28	41.80	-4.15	-11.30	-7.01

For the aggregate of the study, workforce and size weigh the highest with talent close behind. Tourism is slightly positive. Inability to drive innovation, economic downturn, and isolation each have negative weights. Looking at each of the demographic-specific weights, most are similar in magnitude and direction to the aggregate. For the social science participants, size and workforce are weighted more than 10 percentage points greater than each of the other divisions. On the other hand, science and engineering participants put the least emphasis on these two components relative to other groups. After producing rankings from the weights for each division in Table 3, it is evident that a difference around 10 percentage points does not significantly affect the rankings. In fact, the ranking of the top 50 cities only differed by an average of five cities between each division. For the aggregate, as well as female and social sciences divisions, Washington DC ranks first. For the Asian, minority, and hard science demographics, San Jose ranks first, followed by Washington DC in

second place. Table 4 presents the ranking of metropolitan areas according to the weights from the entrepreneurial attractiveness survey.

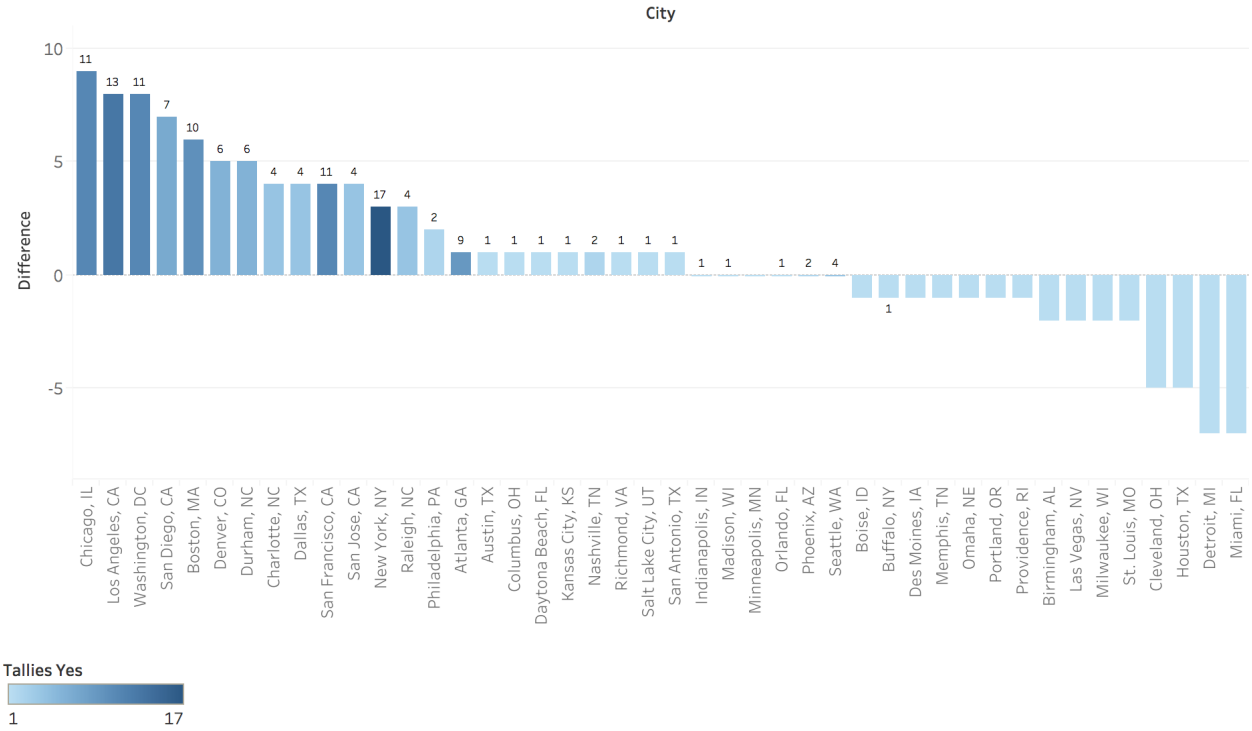
Along with the ranking of cities, the survey asked students to name their most preferred and least preferred locations when finding a job. Scores for metropolitan areas—calculated as the difference between tallies for most preferred and least preferred—are presented in Figure 7. New York, NY scored the highest for both most preferred and least preferred, giving it an overall score of three. Yet, New York ranks third highest according to the weighted index results. Chicago, Los Angeles, and DC scored highest with overall scores of nine, eight, and eight respectively. Miami scored the lowest score at -7, with zero "yes" tallies. Miami ranks 17th according to the weighted ranking, ahead of five other of Amazon's short list cities. Miami and Indianapolis were the only two cities from Amazon's list that ranked at or below zero according to the chart. The study's participants did not mention Pittsburgh at all.

Table 4. Indexed Ranking of Metropolitan Areas According to Survey Results

1	Washington-Arlington-Alexandria, DC-VA-MD-WV	*Amazon*
2	San Jose-Sunnyvale-Santa Clara, CA	
3	New York-Newark-Jersey City, NY-NJ-PA	*Amazon*
4	San Francisco-Oakland-Hayward, CA	
5	Houston-The Woodlands-Sugar Land, TX	
6	Austin-Round Rock, TX	*Amazon, Google*
7	Los Angeles-Long Beach-Anaheim, CA	*Amazon*
8	Dallas-Fort Worth-Arlington, TX	*Amazon*
9	Boston-Cambridge-Newton, MA-NH	*Amazon*
10	Denver-Boulder, CO	*Amazon, Google*
11	Seattle-Tacoma-Bellevue, WA	
12	Baltimore-Columbia-Towson, MD	
13	San Diego-Carlsbad, CA	
14	Raleigh, NC	*Amazon*
15	Chicago-Naperville-Elgin, IL-IN-WI	*Amazon, Google*
16	Atlanta-Sandy Springs-Roswell, GA	*Amazon*

17	Miami-Fort Lauderdale-West Palm Beach, FL	*Amazon*
18	Salt Lake City, UT	
19	Charlotte-Concord-Gastonia, NC-SC	
20	Durham-Chapel Hill, NC	*Amazon, Google*
21	Minneapolis-St. Paul-Bloomington, MN-WI	*Google*
22	Portland-Vancouver-Hillsboro, OR-WA	
23	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	*Amazon*
24	Tampa-St. Petersburg-Clearwater, FL	
25	San Antonio-New Braunfels, TX	
26	Orlando-Kissimmee-Sanford, FL	
27	Madison, WI	
28	Phoenix-Mesa-Scottsdale, AZ	
29	Provo-Orem, UT	
30	Sacramento-Roseville-Arden-Arcade, CA	
31	Kansas City, MO-KS	
32	Colorado Springs, CO	
33	Nashville-Davidson-Murfreesboro-Franklin, TN	*Amazon, Google*
34	Pittsburgh, PA	*Amazon*
35	Las Vegas-Henderson-Paradise, NV	
36	Honolulu, HI	
37	Oxnard-Thousand Oaks-Ventura, CA	
38	Indianapolis-Carmel-Anderson, IN	*Amazon*
39	Baton Rouge, LA	
40	Columbus, OH	
41	Charleston-North Charleston, SC	
42	Richmond, VA	
43	Detroit-Warren-Dearborn, MI	*Google*
44	Bridgeport-Stamford-Norwalk, CT	
45	Oklahoma City, OK	
46	Hartford-West Hartford-East Hartford, CT	
47	Riverside-San Bernardino-Ontario, CA	
48	Ogden-Clearfield, UT	
49	Albany-Schenectady-Troy, NY	
50	Worcester, MA-CT	

Figure 7. Preferred and Least Preferred City Response Scores, Entrepreneurial Attractiveness Survey



To create a viable empirical model for entrepreneurial attractiveness, regression analysis determines the weights of each component. In order to represent the differences between Google and Amazon's interpretations of attractiveness as they align with their goals, two separate regressions are run. The two companies are then combined in a final regression that represents their combined interests. While it is assumed ex ante Amazon's goals vary greatly from Google's, combining the two in a combined regression will confirm or disprove such a hypothesis. Table 5 presents the logistic regression results of the components on the dependent variables representing Amazon's short list, Google's tech hub network, and their choices combined. Appendix 6 presents the specific estimates of the binary logistic models for the regressions in Table 5.

Table 5. Regression Analysis

	(1) Logistic: Google	(2) Logistic: Amazon	(3) Logistic: Combined
constant	-4.41** (0.63)	-3.76** (0.69)	-3.94** (0.82)
C1: Size	0.62** (0.29)	1.07** (0.19)	1.28** (0.24)
C2: Tourism	-0.28 (0.13)	-0.39* (0.26)	0.26 (0.27)
C3: Talent Networks	0.22 (0.36)	0.49 (0.40)	0.43 (0.35)
C4: Skilled Workforce	-0.21 (0.33)	0.07 (0.20)	-0.19 (0.23)
C5: Inability to Drive Growth	-0.30 (0.38)	-0.65* (0.35)	-0.53 (0.42)
C6: Economic Downturn	-1.11** (0.35)	-0.88** (0.40)	-1.26** (0.50)
C7: Isolated Tech Growth	0.70** (0.34)	-0.03 (0.31)	0.24 (0.39)
Number of Observations	99	99	99
Log-Likelihood	-17.70	-22.34	-19.48
Wald chi^2(7)	21.91	39.05	46.02
Prob > chi^2	0.0026	0.0000	0.0000
Pseudo R^2	0.3003	0.5080	0.6285

**Significant at a 95% confidence level

*Significant at a 85% confidence level

The regression analysis shows multiple similarities between Amazon and Google. Amazon's emphasis on city size and population is greater than Google's, though both are significant at a 95% confidence level. Amazon prefers cities with a negative tendency toward tourism, but its confidence level is low enough to suggest that some cities might fall into this category. Google also has a

negative, but insignificant tendency toward tourism. Neither of the components describing talent—three and four—is significant in the regression for Google or Amazon. For Amazon, component five, which describes an inability to drive growth and innovation, is significant at the 85% level in the negative direction. Component six—economic downturn—is has a significant negative coefficient for both Amazon and Google. This indicates an emphasis on economic health of the city. For both Amazon and Google combined, the two significant components—both at 95% confidence level or higher—are a positive coefficient for size and a negative coefficient for economic downturn.

The one difference between the two companies is the geographical placement of the cities that each prefers. Component seven is significant at the 95% level for Google, but is not significant for Amazon. Initially, as Google chose its tech hub locations, it preferred a network spread across the country. Amazon, on the other hand, is choosing only one city, so spreading out is not a concern. A positive coefficient for component seven means the cities are generally more remote geographically. With the exception of location, the results show Amazon and Google's preferences are more similar than expected. They both prefer larger cities that are economically successful and focus on business culture rather than tourism culture. Talent is also positive but insignificant for both cities. This suggests that innovation—and the people behind it—on both the corporate and small business level prefers similar city environments when considering locational attractiveness.

The regression coefficients produce the following weights for each component. I have included all components, even those that are not statistically significant. Eliminating the insignificant components from the analysis would be inappropriate because they capture elements of the data that are vital to the index despite their insignificance. Equation 8 represents the indexed scoring resulting from a regression on Amazon's short list. Equation 9 represents the indexed scoring resulting from a regression on Google's tech hubs. Table 6 lists the top and bottom ten ranked cities

according to Equation 8's results on Amazon. Table 7 lists the top and bottom ten ranked cities according to Equation 9's results on Google. The full list of index scores and rankings is in Appendix 7.

$$\text{Index}_{\text{Amazon}} = 3.37C_1 - 1.25C_2 + 1.56C_3 + 0.22C_4 - 2.04C_5 - 2.77C_6 - 0.09C_7 \quad (8)$$

$$\text{Index}_{\text{Google}} = 6.00C_1 - 0.27C_2 + 2.18C_3 - 2.02C_4 - 2.91C_5 - 10.80C_6 + 6.82C_7 \quad (9)$$

Table 6. Amazon's Ranked Metropolitan Areas in the United States

Top Ten	Bottom Ten
Washington-Arlington-Alexandria, DC-VA-MD-WV	Riverside-San Bernardino-Ontario, CA
New York-Newark-Jersey City, NY-NJ-PA	Springfield, MA
Chicago-Naperville-Elgin, IL-IN-WI	El Paso, TX
Denver-Boulder, CO	Honolulu, HI
Atlanta-Sandy Springs-Roswell, GA	Deltona-Daytona Beach-Ormond Beach, FL
Dallas-Fort Worth-Arlington, TX	Bakersfield, CA
Boston-Cambridge-Newton, MA-NH	Fresno, CA
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	Modesto, CA
Baltimore-Columbia-Towson, MD	Stockton-Lodi, CA
Austin-Round Rock, TX	McAllen-Edinburg-Mission, TX

Table 7. Google's Ranked Metropolitan Areas in the United States

Top Ten	Bottom Ten
Denver-Boulder, CO	Worcester, MA-CT
Washington-Arlington-Alexandria, DC-VA-MD-WV	El Paso, TX
Kansas City, MO-KS	Fresno, CA
Dallas-Fort Worth-Arlington, TX	Springfield, MA
Minneapolis-St. Paul-Bloomington, MN-WI	Riverside-San Bernardino-Ontario, CA
Seattle-Tacoma-Bellevue, WA	Bakersfield, CA
St. Louis, MO-IL	Deltona-Daytona Beach-Ormond Beach, FL
Omaha-Council Bluffs, NE	McAllen-Edinburg-Mission, TX
Austin-Round Rock, TX	Modesto, CA
Atlanta-Sandy Springs-Roswell, GA	Stockton-Lodi, CA

Because of its high level of talent paired with a sizable economy, Washington D.C. ranks at the top of the list of attractiveness for Amazon's considerations. It also ranks top of the growth entrepreneurship index by the Kauffman Foundation, which was mentioned previously. The other cities are notable for their size, their talent, or both. Further, these highly ranked cities host some of the most renowned universities in the country, which produce considerable talent particularly in the fields of technology.

In terms of Google's considerations, Denver, CO tops the list of entrepreneurial attractiveness. Each of the cities is notable for an entrepreneurial community, captured by intra-city talent networks. Interestingly, though, the cities do not exist in geographic networks. With the exceptions of Washington, DC and Denver, CO, most of the cities score zero in the geographic network variable described in Section V. Those that are not scoring zero still score below the average of 29.20.

Washington, DC's size, talent, and economic health score higher than those of each of the other cities. These qualities single it out consistently as an attractive place for entrepreneurial innovation to emerge from both corporations and individuals. For the remaining metropolitan areas in the top ten, and even top twenty, there are unique characteristics that score high and contribute to their ranking. Denver may not be optimally located for all in terms of network, but the combination of sizable economy and strength of talent places it high on both lists. Austin is another similar city, which is fairly isolated but has a strong community of techies and growing economy. Such a combination is attractive to both Google and Amazon's goals. Philadelphia's tech community is lacking compared to other frontrunners and would not suffice for Google, but its size and location are important to large companies like Amazon. Durham and its counterpart Raleigh fail to make the top ten for either company according to these results. Despite having one of the most tech savvy and talented workforces in the country, their sizes are not near populous enough to break into the

top ten. Nevertheless, Durham remains in the top twenty for Google and top twenty-five for Amazon while Raleigh sits in the top twenty for both.

Before concluding the discussion of empirical results, it is necessary to address the bottom tier of metropolitan cities. Specifically, I analyze the bottom cities' shortfalls and what improvements can be made to improve their ranking. Using three cities as examples—Winston-Salem, NC, Memphis, TN, and Providence, RI—I effectually improve the score of each city's most failing component, according to Amazon's coefficient parameters. For all cities, component seven is ignored because a city's geographic network location cannot be improved endogenously.

Winston-Salem's primary shortfall is its population. At 662,000 people, it is well below Amazon's requirements. Hence, the number of businesses and employment opportunities are minimal. Further, its commercial airport services only limited east coast and Midwest destinations, which can hinder efficient business travel to and from the city. Other aspects correlated with population that are lacking include per capita output and educational attainment; with a lack of job opportunities, no new talent or income is moving to the city. Per capita output is $37, 398, which is well below the average of $50,000 for the data set. Adjusting Winston Salem's score for size of -3.62 to Washington DC's score of 5.81, Winston Salem improves from a rank of 89 to an astonishing third place, scoring 18.69. While it is unlikely Winston-Salem would reach Washington DC's population and enterprise strength, attracting small population and enterprise growth would bolster its attractiveness.

Memphis' major shortfalls include a lack of talent and an inability to drive innovation and growth, indicated by below average GDP with negative growth, low educational attainment, low patent output, and low job creation. Improving Memphis' talent pool would likely drive innovation. Therefore, I adjust Memphis' talent network component score from -2.84 to 8.72. This change would improve Memphis' ranking from 86 to second place. Improving innovation and growth alone

raises Memphis' ranking to 51. Providence's primary struggle is the health of the economy. Most of the components score at or below average as well, causing it to be ranked 77 originally. Improving its economy to the level of DC would improve its ranking to 51. On the other hand, adjusting Providence's economy by an even higher score, like Dallas', would improve its ranking even more, to 44.

Ultimately, the weight of the component matters greatly when considering how certain improvements might affect a city's ranking. Improving any city's size and talent would have large effects on the ranking. No city can improve a component instantaneously as in this simulation. Nevertheless, incremental improvements to certain components can benefit even the cities currently ranked in the bottom 25. For these cities, incrementally attracting talent, population growth, and enterprise would improve their long run trajectories.

VIII. Discussion of Geographic Network Effects

In this section, the geographic network effects of city location are discussed in detail. Simply, geographic network effects refer to the positive effects that locating near other successful cities has on the target city's attractiveness. These effects are captured several different ways in this research. Most obviously, the geographic network variable discussed in Section V captures the location of a city relative to other cities. It scores geographic location based on nearest city ranking as well, to indicate that locating near other successful cities is more valuable than locating near unattractive cities. The other way this effect is captured is by looking at cities with nearby counterparts that often function better when considered as one synergic area. Three cities are used to analyze these effects in further detail: Philadelphia, PA, Denver, CO, and Research Triangle, NC.

Philadelphia derives its positive network effects from its location in the Eastern corridor among powerhouses such as New York, NY and Washington, DC. According to the geographic

network variable, Philadelphia ranks fifth in network location, behind Hartford, CT, Research Triangle, NC, New Haven, CT, and Springfield, MA. Philadelphia's ranking prior to including the geographic network variable is nine. After including the variable, Philadelphia rises to an eighth-place ranking.[10]

To analyze Denver and Research Triangle, a different approach is used. Rather than looking at the geographic network variable alone, the cities are adjusted prior to a second iteration of statistical analysis. The metropolitan area Denver-Aurora-Lakewood, CO replaces the combination of Denver-Boulder that was previously used in the study. Research Triangle is a new observation that combines the two metropolitan areas Raleigh, NC and Durham-Chapel Hill, NC. These cities are analyzed using the same statistical methodology applied previously, where regression analysis determines the weights for each component.

The effects analyzed using Research Triangle and Denver represent how cities benefit from the synergies of city collaboration. Given the Denver case, the geographic variable is not relevant to the analysis. Instead, Denver is analyzed solely in the context of Boulder's proximity, rather than a web of multiple proximal cities. After removing Boulder's effects and scoring the city, Denver goes from being Google's first ranked city to its fifth, and from Amazon's fourth ranked city to its seventh.

The case of Research Triangle combines the geographic network effect and the city collaboration effect. Raleigh and Durham are eliminated from the metropolitan list to examine this iteration of the data. Their data are instead averaged or added to form a new metro called "Research Triangle, NC." In terms of geographic network, Research Triangle is strong; it is ranked second with a score of 97.40, compared to Durham's 58.06 and Raleigh's 26.48. After statistical analysis, Research Triangle is ranked 23rd according to Google's parameters and 17th according to Amazon's.

[10] No data is available for Philadelphia's ranking according to Google's parameters prior to including the network variable.

Previously Raleigh was ranked 17th for Google and 18th for Amazon, while Durham was ranked 15th for Google and 23rd for Amazon. The results provide several important observations. First, the ranking for both cities, but more significantly for Durham, improves when considered as one Research Triangle area. More interesting but unsurprising, Google's rankings for the individual cities were higher before combining into Research Triangle. The results provide more evidence to distinguish Google's goals from Amazon's. Google is able to overlook a small population like that of Durham if the talent networks are strong enough to host a successful tech hub. Amazon's goals align instead with the size and strength of Research Triangle as a whole. Durham's addition to Research Triangle does in effect improve Raleigh's ranking on the short list.

Ultimately, the geographic network effects are significant to Amazon but are not significant to Google. In fact, the regression analysis indicates that Google "prefers"—whether purposefully or solely by nature of the tech communities—investing in cities which are not supported by geographic networks. Hence, Durham and Raleigh produce reduced scores according to Google when combined into one Research Triangle. In terms of Amazon's preferences, geographic network and city collaboration increase city attractiveness as the results of these three case studies show.

IX. Conclusion

In this paper, the index above measures the entrepreneurial attractiveness of 99 metropolitan areas across the United States according to Amazon and Google's locational preferences. The index presents the data according to 34 input variables spanning a range of city characteristics from demographics to business data. PCA reduces these variables to seven workable components, according to the correlation matrix of the data. Weights are determined through regression analysis on Google's tech hubs and Amazon's short list. These weights then produce an objective, research-based index and ranking as presented in Tables 6 and 7. Other versions of the set of weights come

from an arbitrary hypothesis, analysis of three case study cities, and a survey of graduating student locational preferences.

The weights given by regression analysis are parallel with Amazon's preferences for their new headquarters as stated in their announcement proposal request. Unfortunately, Amazon has yet to make a final decision concerning HQ2, so we cannot observe what characteristics ultimately determine the winner. These characteristics may be different from those produced by the regression results. The regression represents the first round of choices made by Amazon. Because having a sizeable population was Amazon's primary criterion, it is unsurprising that size is the most significant component according to the regression. The final round will not emphasize size, but likely talent and lifestyle conditions instead. Therefore, it is equally as important to apply non-empirical iterations of the index, observing how cities score in different situations.

Amazon serves as only one case of a company searching for the city most conducive to their innovative directives. Other companies could propose preferences different from Amazon's. Yet, even in such case the index presented in this study incorporates a corporate decision maker's priorities. Further, the use of two differing companies with differing objectives provides insight into forms of entrepreneurial attractiveness. Corporate success and entrepreneurial success each require city infrastructure, environments, and economies conducive to their innovation goals. A corporation does not draw success from how strong the tech start-up relationships in a city are, but does draw from university talent in the city.

The model captures city strengths according to a general framework for entrepreneurial attractiveness. The determinants do not imply that a city cannot be entrepreneurially attractive without a distinct set of qualities. For example, Durham is the least populous city in the sample yet scores in the top 25. The weights and the index it creates provide a framework for the combination

of characteristics most crucial to successful entrepreneurship. The strength of the model is allowing for outlying elements to be considered in the scoring process.

One of the realities of the data when producing the index is that many cities with attractive qualities are hidden in the data by more successful cities. Each city's values for each variable might be impressive, but not impressive *enough* to stand out over many of the higher-ranking cities. This is especially relevant considering the emphasis placed on size in the empirical ranking. Ranking the cities does not account for how close one city might be to reaching the top of the list, so the indexed scores are included in Appendix 7.

Several of the qualities that the data ignore are the synergies, histories, and trajectories that define each city. These qualitative aspects also contribute to the entrepreneurial ecosystem and serve a large role in how entrepreneurs view the city. This is especially important as we extend our discussion from entrepreneurial attractiveness to entrepreneurial success. For example, a city might have a youthful demographic and a large presence of tech degrees, but much of the success comes from how the city and its residents use and value these elements. Additional variables may include entrepreneurial social networks, potential for growth booms, industrial concentration, government involvement, and university involvement. Political leanings are another viable factor in a corporation or individual's decision making, especially as tensions rise when parties favor certain types of business over others.

This paper presents a linear, single equation regression model to determine entrepreneurial attractiveness. In the case of both Amazon and Google, this model may not be the single most accurate; there may be multiple models driving their decisions. For example, Amazon's decision process for HQ2 involves two rounds of narrowing down their list. One of these hasn't been made yet; therefore, it is impossible to empirically include this two-step process using this paper's methods. Such a model would require two sequential linear regressions, determining which

characteristics are most important for each round of decisions. On the other hand, the model might not even be linear in the first place.

To expand on the research in this paper, more extensive and more objective research is necessary. Even the use of Amazon alone as a case study brings a hint of subjectivity, as the company's own views are emphasized without considering others. It is also important, therefore, to consider the opinions of individual entrepreneurs and employees who will work for the companies in question. The survey of undergraduate students used in the research accomplishes this goal; including these data offers a new interpretation of the weights used in the creation of the index.

Further extensions of the research are also possible. First, future data sets should include longitudinal data to observe whether entrepreneurial attractiveness changes over time for certain cities, and which variables affect these changes. Such data could be used to influence the trajectories of certain cities as they change over time. Second, it is important to expand this research to be used across the world. Currently, analysis of United States metropolitan areas cannot be applied to other countries and their cities for several reasons. First, Amazon's search for a new headquarters is inapplicable to international cities simply due to Amazon's intention to remain in North America. However, Google's investments do extend overseas, and could offer viable further analysis. Second, the characteristics affecting the attractiveness of United States metropolitan areas could be significantly different from those across the world. The United States' size allows for the comparison across American cities. However, globally, many countries have only several contenders with populations over one million, so cities have to be compared across countries. In this case, national level variables are more significant to an international data set than they are to the United States' data set. Eventually, international analysis will provide important insights, especially in the comparison of United States data to international data. As the global phenomenon of

entrepreneurship and its widespread acclaim continues to inspire cities and governments across the world, objective research to guide policy and initiatives is critical to its success.

Appendices

Appendix 1. Overview of Variables

A1. Overview of Data

Total Entries: 104

Variable Name	Description of Variable	Type of Variable	Year	Missing Values	Source
city	name of the metropolitan area	text	2016	0	US Census
code	MSA code for the metropolitan area	text	2016	0	US Census
country	country of city's existence	text	2016	0	US Census
continent	continent of city's existence	text	2016	0	Google Maps
state	state of city's existence	text	2016	0	US Census
region	region of city's existence	text	2016	0	Google Maps
population	total city population	numeric	2016	0	US Census
size	total area of the metropolitan area in square meters	numeric	2016	0	US Census and Gazetteer
sizekm	total area of the metropolitan area in square kilometers	numeric	2016	0	US Census and Gazetteer
airtraffic	all commercial enplanements at the airport	numeric	2016	0	FAA Enplanements Data 2015-2016
hubsize	size of the airport according to FAA hub size: International Largest=4; Large Hub=3; Medium Hub=2; Small Hub=1; Non-Hub=0	indicator	2016	0	FAA Enplanements Data 2015-2016
intlairport	whether or not the city has an international airport	binary	2016	0	FAA Enplanements Data 2015-2016
geonetwork	arbitrary measure of city networking abilities, calculated as a sum of its own rank and all the cities' within 150 miles ranks divided by the distance to the city	numeric	2016	0	Manually Calculated Value
conferences	number of tech conferences hosted by the city or private venues in the city during the span of one year; includes women's and media tech-specific conferences as well	numeric	2016	0	The Bizzabo Blog
techhub	whether or not the city is a Google or Amazon designated tech hub	binary	2016	0	Google and Amazon
google_techhub	whether or not the city is a Google designated tech hub	binary	2016	0	Google for Entrepreneurs
amazon_techhub	whether or not the city is on Amazon's short list for HQ2	binary	2018	0	Amazon Inc.
googlefiber	whether or not the city has Google Fiber	binary	2016	0	Google Fiber
creativeclass	percentage of creative class employment in each city	numeric	2010	0	Richard L. Florida
degrees_science	percentage of bachelor's degrees in the science and engineering related fields	numeric	2016	0	US Census, American Community Survey, 1-Year
degrees_business	percentage of bachelor's degrees in the business field	numeric	2016	0	US Census, American Community Survey, 1-Year
edu_attain	percentage of inhabitants who have a bachelor's degree or higher	numeric	2016	0	US Census, American Community Survey, 1-Year
edu_millennials	percentage of inhabitants aged 25-34 who have a bachelor's degree or higher	numeric	2016	0	US Census, American Community Survey, 1-Year
patent	number of patents filed during one year within the metropolitan area	numeric	2013	0	US Cluster Mapping by Harvard University
patent_percapita	number of patents filed during one year per capita	numeric	2013	0	US Cluster Mapping by Harvard University
clustereff	strength of the industry cluster environment	numeric	2015	0	US Cluster Mapping by Harvard University
gdpgrowth	10 year growth rate in per capita gdp ppp; 2006-2016	numeric	2016	1	US Bureau of Economic Analysis
gdp_percapita	gross domestic product (GDP) per capita PPP, adjusted for inflation (USD $)	numeric	2016	1	US Bureau of Economic Analysis
realgdp_mil	gross domestic product (GDP) (Millions of USD $)	numeric	2016	1	US Bureau of Economic Analysis
rentalrate	median rental price per square foot for a single family residence (USD $)	numeric	2016	3	Zillow Research
homevalue	median home value per square foot for all homes, averaged over one year (USD $)	numeric	2016	5	Zillow Research
costofliving	estimation of consumer goods, services, transportation, and rent in the city compared to New York City, excluding rent	numeric	2016	1	Numbeo
costofliving_rent	estimation of consumer goods, services, transportation, and rent in the city compared to New York City, including rent	numeric	2016	1	Numbeo

pricetoincome	relative measure of apartment affordability; rental price multiplied by the average square footage of a one-bedroom apartment in the US (750), divided by per capita GDP	numeric	2016	3	Zillow Research and US Census
purchasingpower	relative purchasing power in buying goods and services in a given city for the city's average salary	numeric	2016	1	Numbeo
qualityoflife	indexed combination of measures of living conditions, including cost, traffic, safety, public goods available, pollution, and climate	numeric	2016	0	Numbeo
commute	mean commute time for all persons age 16 or older (minutes)	numeric	2016	0	US Census, American Community Survey
climate	an indexed measure of climate likeability that takes into account temperatures, humidity, and extreme weather, but ignores individual preference	numeric	2016	0	Numbeo
crime	an indexed measure of crime that takes into account crimes committed and the population's feelings about crime in the area	numeric	2016	0	Numbeo
medianage	median age of the population	numeric	2016	0	US Census, American Community Survey
hispanic	percentage of population that reports as having hispanic or latino origins	numeric	2016	0	US Census, American Community Survey, 5-Year
white	percentage of population that reports as white only, and not hispanic or latino	numeric	2016	0	US Census, American Community Survey, 5-Year
diversity	percentage of population that reports as having any ethnicity other than white only, including hispanic or latino	numeric	2016	0	US Census, American Community Survey, 5-Year
black	percentage of population that reports as black or african american only, and not hispanic or latino	numeric	2016	0	US Census, American Community Survey, 5-Year
asian	percentage of population that reports as asian only, and not hispanic or latino	numeric	2016	0	US Census, American Community Survey, 5-Year
poverty	poverty rate	numeric	2016	0	US Census, American Community Survey
unemployment	seasonally adjusted monthly unemployment rate, April	numeric	2016	0	US Bureau of Labor Statistics
annualpayroll_all	annual labor payroll, all sectors, $1000	numeric	2015	2	US Business Dynamics Statistics
annualpayroll_tech	annual labor payroll, NAICS 54, $1000	numeric	2015	3	US Business Dynamics Statistics
techratio_pay	percent of payroll that goes to the NAICS 54 sector	numeric	2015	3	US Business Dynamics Statistics
establishments_all	total establishments over all sectors	numeric	2015	2	US Business Dynamics Statistics
establishments_tech	total establishments, NAICS 54	numeric	2015	2	US Business Dynamics Statistics
techratio_est	percent of establishments that are NAICS 54 sector	numeric	2015	2	US Business Dynamics Statistics
employment_all	total employment over all sectors	numeric	2015	2	US Business Dynamics Statistics
employment_tech	total employment, NAICS 54	numeric	2015	3	US Business Dynamics Statistics
techratio_employ	percent of employment that is part of the NAICS 54 sector	numeric	2015	3	US Business Dynamics Statistics
employchange	10 year growth rate in employment for all sectors; 2005-2015	numeric	2015	2	US Business Dynamics Statistics
techemploychange	10 year growth rate in tech employment for NAICS 54 sector; 2005-2015	numeric	2015	2	US Business Dynamics Statistics
jobcreation	number of jobs created during the span of ten year; 2005 to 2015	numeric	2015	0	US Cluster Mapping by Harvard University
jobcreation_percapita	number of jobs created per capita during a 10 year period; 2005 to 2015	numeric	2015	0	US Cluster Mapping by Harvard University
estentries	total establishment entries during 2014	numeric	2014	1	US Business Dynamics Statistics
entries_percapita	number of establishment entries per capita	numeric	2014	1	US Business Dynamics Statistics
estexits	total establishment exits during 2014	numeric	2014	1	US Business Dynamics Statistics
entries_exits	ratio of all establishment entries to establishment exits	numeric	2014	1	US Business Dynamics Statistics
netentries_exits	net value of all establishment entries minus the establishment deaths	numeric	2014	1	US Business Dynamics Statistics

Appendix 2. Descriptive Statistics

A2a. Descriptive Statistics - All Observations

Variable	Obs	Mean	Std. Dev.	Min	Max
population	104	2,101,984	2,726,696	529,657	20,200,000
sizekm	104	11,224.15	9,120.17	1,555.56	70,612.33
airtraffic	104	7,271,745	12,100,000	0	63,900,000
hubsize	104	2.40	1.22	0	4
geonetwork	104	28.73	24.88	0	136.20
conferences	104	2.48	6.53	0	38
degrees_science	104	43.81	4.32	35.50	63.50
edu_attain	104	32.61	7.18	16.20	50.90
edu_millennial	104	35.43	9.38	13.80	59.80
patent	104	973.73	1,834.79	9	11,374
patent_percapita	104	0.00040	0.0006	0	0.0057
clustereff	104	0.4486	0.1358	0.1520	0.8655
gdpgrowth	103	0.0053	0.0929	-0.2392	0.4477
gdp_percapita	103	50,217.97	14,752.45	20,361	119,695
rentalrate	102	0.98	0.33	0.67	2.40
homevalue	104	140.35	89.67	59.25	614.75
pricetoincome	101	0.1814	0.0508	0.1075	0.3363
purchasingpower	103	143.6	24.51	83.16	203.07
commute	104	25.52	3.46	20	36.70
crime	103	49.82	13.54	16.09	81.82
climate	104	77.35	11.25	51.44	97.16
qualityoflife	104	180.9	18.25	137.94	221.82
medianage	104	37.83	3.80	24.7	52
diversity	104	36.93	17.65	7.40	93.10
hispanic	104	16.31	16.57	1.60	91.50
white	104	63.07	17.65	6.90	92.60
asian	104	4.72	5.84	0.70	41.90
black	104	12.64	10.20	0.40	48.90
poverty	104	0.1351	0.0362	0.0800	0.3140
unemployment	104	9.72	2.31	5.20	17.00
payroll_all	102	46,100,000	72,900,000	51,848,829	564,000,000

payroll_tech	101	5,909,419	10,900,000	201,800	72,800,000
techratio_pay	101	0.1017	0.04502	0.0318	0.3275
establishments_all	102	51,340.50	74,406.43	8,744	572,361
establishments_tech	102	6,779.57	10,361.82	665	71,639
techratio_est	102	0.1185	0.0265	0.0695	0.2173
employment_all	101	844,010.30	1,127,002	135,488	8,123,112
employment_tech	101	69,162.94	113,414	4,564	745,940
techratio_employ	101	0.0676	0.0254	0.0263	0.2034
employchange	102	0.0706	0.0975	-0.1014	0.3488
techemploychange	102	0.1192	0.1846	-0.2973	0.5935
jobcreation	104	24,995.75	85,193.42	-119,017	544,950
jobcreation_percapita	104	0.0030	0.0293	-0.0496	0.0880
estentries	103	4,683.07	7,251.04	756	54,894
entries_percapita	104	0.0019	0.0005	0	0.0034
estexits	103	3,920.31	6,163.64	650	48,537
entries_exits	103	1.17	0.0975	0.9913	1.47

A2b. Descriptive Statistics - Google's Tech Hub Network

Variable	Obs	Mean	Std. Dev.	Min	Max
population	7	3,574,028	2,894,379	559,535	9,512,999
sizekm	7	14,828.16	6,580.21	4,552.50	23,496.91
airtraffic	7	18,500,000	15,700,000	5,401,714	48,600,000
hubsize	7	3.57	0.53	3	4
geonetwork	7	16.34	20.21	0	58.06
conferences	7	4.71	4.15	0	12
degrees_science	7	45.65	5.04	38.80	53.80
edu_attain	7	40.43	7.17	30.40	50.90
edu_millennial	7	45.66	6.91	34	55.60
patent	7	1,936.86	1,160.68	198	3,339
patent_percapita	7	0.0007	0.0004	0.0001	0.0012
clustereff	7	0.4807	0.1605	0.1520	0.6287
gdpgrowth	7	0.0453	0.0749	-0.0707	0.1583
gdp_percapita	7	60,975.07	5,200.05	52,084	68,586
rentalrate	7	1.06	0.19	0.88	1.36

homevalue	7	144.68	47.15	100.83	237.71
pricetoincome	7	0.1561	0.0262	0.1157	0.1877
purchasingpower	7	157.90	29.26	127.65	197.24
commute	7	26.89	2.49	24.60	31.80
crime	7	49.57	15.72	27.26	71.09
climate	7	71.35	12.7	51.46	84.07
qualityoflife	7	184.58	27.5	144.87	221.76
medianage	7	36.90	1.69	34.40	40.10
diversity	7	35.61	10.25	22.90	46.70
hispanic	7	14.22	10.18	4.20	31.90
white	7	64.39	10.25	53.30	77.10
asian	7	4.68	1.33	2.50	6.20
black	7	14.03	8.58	3.10	26.50
poverty	7	0.1196	0.0230	0.0880	0.1490
unemployment	7	9.05	2.96	5.75	14.60
payroll_all	7	84,100,000	73,600,000	13,900,000	237,000,000
payroll_tech	7	11,200,000	9,596,672	3,035,306	31,200,000
techratio_pay	7	0.1447	0.0488	0.0787	0.2186
establishments_all	7	88,601.57	75,020.71	12,362	243,420
establishments_tech	7	12,425.43	10,427.07	1,961	33,708
techratio_est	7	0.1427	0.0282	0.0986	0.1727
employment_all	7	1,518,832	1,269,715	241,065	4,120,166
employment_tech	7	128,183.10	103,965.10	31,217	338,120
techratio_employ	7	0.0903	0.0255	0.0533	0.1295
employchange	7	0.1181	0.1242	-0.0534	0.3308
techemploychange	7	0.1527	0.2452	-0.1577	0.5935
jobcreation	7	49,676	96,692.52	-119,017	180,880
jobcreation_percapita	7	0.0246	0.0361	-0.0277	0.0880
estentries	7	8,120.71	6,537.78	1,070	21,461
entries_percapita	7	0.0023	0.0004	0.0019	0.0030
estexits	7	6,760.43	5,852.57	882	18,854
entries_exits	7	1.24	0.1218	1.10	1.47

A2c. Descriptive Statistics - Amazon's Short List

Variable	Obs	Mean	Std. Dev.	Min	Max
population	17	5,553,585	5,019,144	559,535	20,200,000
sizekm	17	14,560	5,910.34	4,552.50	24,035.45
airtraffic	17	22,900,000	19,500,000	3,567,864	63,900,000
hubsize	17	3.59	0.51	3	4
geonetwork	17	30.77	25.27	0	83.51
conferences	17	7.06	9.83	0	38
degrees_science	17	45.11	4.31	38.80	53.80
edu_attain	17	39.54	6.55	30.50	50.90
edu_millennial	17	44.92	7.54	31.50	58.00
patent	17	2,297.18	1,972.66	198	7,080
patent_percapita	17	0.0005	0.0003	0.0001	0.0012
clustereff	17	0.4855	0.1546	0.1520	0.7682
gdpgrowth	17	0.0496	0.0759	-0.0858	0.1812
gdp_percapita	17	61,971.09	7,683.97	47,438	76,968
rentalrate	17	1.16	0.44	0.71	2.09
homevalue	17	159.49	79.79	80.58	380.58
pricetoincome	17	0.1669	0.0523	0.1125	0.2656
purchasingpower	17	135.14	23.03	100	197.24
commute	17	28.86	3.71	24	36.7
crime	17	46.60	11.77	27.26	65.37
climate	17	79.47	8.60	60.00	95.90
qualityoflife	17	176.16	24.52	137.94	221.76
medianage	17	37.29	2.15	34.40	43.10
diversity	17	41.60	15.56	13.90	69.70
hispanic	17	17.45	13.13	1.60	45.00
white	17	58.40	15.56	30.30	86.10
asian	17	5.77	3.39	2.10	15.20
black	17	15.78	7.94	3.10	32.90
poverty	17	0.1229	0.0206	0.0840	0.1540
unemployment	17	8.94	1.59	5.75	11.80
payroll_all	17	134,000,000	136,000,000	13,900,000	564,000,000
payroll_tech	17	19,400,000	20,000,000	3,035,306	72,800,000
techratio_pay	17	0.1459	0.0586	0.0787	0.3275

establishments_all	17	143,539.20	141,945.90	12,362	572,361
establishments_tech	17	20,439.94	18,783.52	1,961	71,639
techratio_est	17	0.1436	0.0301	0.0986	0.2173
employment_all	17	2,285,697	2,043,219	241,065	8,123,112
employment_tech	17	215,043.30	206,140.20	31,217	745,940
techratio_employ	17	0.0928	0.0341	0.0533	0.2034
employchange	17	0.1071	0.0828	0.0176	0.3308
techemploychange	17	0.1725	0.1598	-0.0803	0.5935
jobcreation	17	97,671.53	106,754.80	-28,467	394,471
jobcreation_percapita	17	0.0233	0.0248	-0.0047	0.0880
estentries	17	13,704	13,855.01	1,070	54,894
entries_percapita	17	0.0023	0.0005	0.0016	0.0034
estexits	17	11,451.59	12,065.24	882	48,537
entries_exits	17	1.22	0.0952	1.08	1.47

Appendix 3. Correlation Coefficients

Appendix 3. Correlation Coefficients

	population	sizekm	airtraffic	hubsize	geonetwork	conferences	degrees_science	edu_attain	edu_millennial	patent	patent_percapita	clustereff	gdpgrowth	gdp_percapita	rentalrate	homevalue	pricetoincome	purchasingpower	commute	crime	climate	qualityoflife
population	1.00																					
sizekm	0.36	1.00																				
airtraffic	0.88**	0.32	1.00																			
hubsize	0.57*	0.36	0.68*	1.00																		
geonetwork	0.49	-0.03	0.47	0.46	1.00																	
conferences	0.66*	0.13	0.65*	0.43	-0.03	1.00																
degrees_science	0.11	-0.03	0.12	0.17	0.19	0.34	1.00															
edu_attain	0.25	-0.05	0.33	0.44	0.09	0.38	0.53*	1.00														
edu_millennial	0.28	-0.12	0.32	0.43	0.18	0.39	0.46	0.93**	1.00													
patent	0.60*	0.10	0.58*	0.46	-0.02	0.82**	0.56*	0.53*	0.54*	1.00												
patent_percapita	0.07	-0.05	0.09	0.17	0.01	0.45	0.66*	0.54*	0.51*	0.77*	1.00											
clustereff	0.12	0.01	0.21	0.09	0.05	0.23	0.46	0.18	0.05	0.29	0.32	1.00										
gdpgrowth	0.13	-0.07	0.05	0.10	-0.03	0.23	0.30	0.35	0.43	0.47	0.51*	-0.01	1.00									
gdp_percapita	0.37	-0.04	0.41	0.47	0.07	0.53*	0.47	0.77*	0.78*	0.71*	0.65*	0.15	0.54*	1.00								
rentalrate	0.50*	-0.06	0.45	0.33	0.08	0.71*	0.58*	0.52	0.50*	0.79*	0.59*	0.30	0.38	0.62*	1.00							
homevalue	0.29	-0.02	0.30	0.27	0.00	0.61*	0.64*	0.47	0.38	0.73*	0.65*	0.37	0.37	0.58*	0.90**	1.00						
pricetoincome	0.10	-0.01	0.00	-0.22	0.04	0.15	0.15	-0.27	-0.34	0.08	-0.03	0.28	-0.18	-0.41	0.42	0.37	1.00					
purchasingpower	-0.20	0.12	-0.17	-0.10	0.07	-0.26	0.02	-0.02	-0.05	-0.13	0.07	0.05	-0.07	-0.13	-0.24	-0.15	-0.05	1.00				
commute	0.66*	0.23	0.65*	0.45	0.13	0.54*	0.39	0.27	0.23	0.54*	0.17	0.34	0.07	0.38	0.54*	0.47	0.21	-0.18	1.00			
crime	0.05	-0.09	0.06	0.00	0.10	-0.06	-0.13	-0.34	-0.30	-0.09	-0.19	-0.02	-0.16	-0.12	-0.13	-0.17	-0.03	-0.13	0.14	1.00		
climate	0.17	-0.05	0.15	0.10	-0.03	0.17	0.17	-0.09	-0.13	0.25	0.11	0.14	0.01	0.11	0.22	0.31	0.13	-0.11	0.47	0.17	1.00	
qualityoflife	-0.40	-0.09	0.36	-0.20	-0.05	-0.31	0.03	0.21	0.17	-0.20	0.10	-0.11	0.05	0.01	-0.24	-0.16	-0.27	0.26	-0.44	-0.49	-0.09	1.00
medianage	-0.04	-0.32	-0.05	-0.04	0.12	-0.02	0.00	0.01	0.03	-0.03	-0.01	-0.10	-0.15	-0.04	0.05	-0.04	0.07	-0.21	0.07	0.03	0.07	-0.01
diversity	0.32	0.22	0.33	0.22	-0.17	0.29	0.14	-0.17	-0.27	0.26	0.07	0.43	-0.04	0.04	0.27	0.35	0.32	0.02	0.45	0.25	0.48	-0.32
hispanic	0.20	0.27	0.17	0.00	-0.16	0.16	0.05	-0.31	-0.40	0.12	0.01	0.37	-0.05	-0.19	0.14	0.19	0.48	0.13	0.21	0.03	0.14	-0.20
white	-0.32	-0.22	-0.33	-0.22	0.17	-0.29	-0.14	0.17	0.27	-0.26	-0.07	-0.43	0.04	-0.04	-0.27	-0.35	-0.32	-0.19	-0.45	-0.25	-0.48	0.32
asian	0.27	-0.02	0.30	0.34	-0.04	0.52*	0.56*	0.31	0.25	0.61*	0.53*	0.38	0.33	0.51*	0.73*	0.87**	0.25	-0.21	0.47	-0.02	0.37	-0.22
black	0.08	-0.05	0.12	0.13	0.05	-0.08	-0.21	0.01	0.05	-0.12	-0.21	-0.11	-0.22	0.04	-0.25	-0.32	-0.40	-0.01	0.14	0.43	0.36	-0.10
poverty	-0.07	0.10	-0.13	-0.23	-0.19	-0.17	-0.33	-0.64*	-0.62*	-0.26	-0.30	0.00	-0.25	-0.52*	-0.38	-0.37	0.19	0.08	-0.22	0.33	0.10	-0.21
unemployment	0.02	0.02	-0.02	-0.17	0.06	0.05	0.03	-0.49	-0.51*	0.00	-0.05	0.12	-0.21	-0.35	-0.03	0.04	0.40	0.05	0.19	0.31	0.16	-0.40
payroll_all	0.97**	0.26	0.85**	0.53*	0.05	0.75*	0.19	0.35	0.39	0.70*	0.19	0.15	0.21	0.48	0.59*	0.37	0.08	-0.23	0.66*	0.02	0.16	-0.38
payroll_tech	0.91**	0.20	0.80**	0.52*	0.10	0.74*	0.28	0.43	0.45	0.70*	0.22	0.25	0.21	0.51*	0.61*	0.42	0.07	-0.24	0.68*	0.02	0.19	-0.37
techratio_pay	0.32	-0.02	0.35	0.38	0.04	0.33	0.55*	0.66*	0.55*	0.41	0.34	0.36	0.15	0.49	0.41	0.37	-0.10	-0.05	0.39	-0.12	0.21	0.05
establishments_all	0.99**	0.29	0.87**	0.53*	0.02	0.68*	0.11	0.27	0.31	0.61*	0.08	0.11	0.13	0.38	0.53*	0.31	0.12	-0.23	0.64*	0.03	0.16	-0.40
establishments_tech	0.98**	0.27	0.90**	0.57*	0.01	0.69*	0.16	0.34	0.34	0.64*	0.12	0.18	0.13	0.42	0.56*	0.36	0.11	-0.23	0.67*	0.02	0.19	-0.39
techratio_est	0.35	0.15	0.51*	0.56*	-0.14	0.43	0.46	0.68*	0.51*	0.50*	0.43	0.48	0.10	0.54*	0.43	0.45	-0.12	0.00	0.39	-0.24	0.17	0.06
employment_all	0.99**	0.31	0.88**	0.58*	0.02	0.68*	0.12	0.30	0.34	0.64*	0.10	0.10	0.17	0.42	0.53*	0.31	0.06	-0.22	0.65*	0.05	0.15	-0.39
employment_tech	0.94**	0.23	0.83**	0.55*	0.07	0.70*	0.23	0.29	0.41	0.66*	0.16	0.20	0.18	0.47	0.57*	0.37	0.07	-0.23	0.67*	0.03	0.18	-0.38
techratio_employ	0.38	0.03	0.41	0.46	0.03	0.42	0.58*	0.70*	0.62*	0.53*	0.45	0.39	0.23	0.61*	0.47	0.44	-0.16	-0.04	0.42	-0.16	0.20	0.01
employchange	0.04	0.20	0.10	0.12	-0.20	0.05	-0.04	0.12	0.07	0.11	0.14	0.08	0.20	0.05	-0.04	0.00	-0.03	0.21	-0.03	-0.45	0.01	0.21
techemploychange	0.04	0.21	0.10	0.17	-0.13	0.00	0.11	0.11	0.07	0.03	0.06	0.11	0.13	0.08	0.00	0.02	-0.10	0.08	0.03	-0.35	0.02	0.09
jobcreation	0.55*	0.33	0.55*	0.44	-0.21	0.39	0.22	0.24	0.19	0.46	0.20	0.27	0.27	0.36	0.30	0.24	-0.05	-0.06	0.41	-0.18	0.19	-0.06
jobcreation_percapita	0.22	0.31	0.30	0.39	-0.33	0.24	0.21	0.28	0.21	0.33	0.29	0.29	0.34	0.28	0.19	0.23	-0.04	0.06	0.19	-0.37	0.05	0.05
estentries	0.99**	0.30	0.88**	0.54*	-0.01	0.68*	0.11	0.27	0.28	0.61*	0.08	0.13	0.12	0.37	0.53	0.32	0.13	-0.22	0.64*	0.02	0.18	-0.39
entries_percapita	0.42	0.17	0.55*	0.55*	-0.29	0.43	0.19	0.52*	0.40	0.43	0.27	0.23	0.08	0.42	0.42	0.37	-0.01	-0.13	0.37	-0.25	0.09	-0.03
estexits	0.98**	0.28	0.87**	0.52*	0.01	0.68*	0.10	0.26	0.28	0.60*	0.07	0.12	0.11	0.36	0.53*	0.31	0.14	-0.22	0.63*	0.03	0.17	-0.39
entries_exits	0.18	0.39	0.28	0.37	-0.34	0.22	0.10	0.22	0.11	0.20	0.14	0.23	0.04	0.12	0.06	0.14	-0.02	-0.02	0.24	-0.25	0.07	0.10

*Indicates moderate correlation where the coefficient is 0.50 or greater

**Indicates strong correlation where the coefficient is 0.80 or greater

Appendix 3. Correlation Coefficients, Cont.

	medianage	diversity	hispanic	white	asian	black	poverty	unemployment	payroll_all	payroll_tech	techratio_pay	establishments_all	establishments_tech	techratio_est	employment_all	employment_tech	techratio_employ	employchange	techemploychange	jobcreation	jobcreation_percapita	estentries	entries_percapita	estexits	entries_exits
population																									
sizekm																									
airtraffic																									
hubsize																									
geonetwork																									
conferences																									
degrees_science																									
edu_attain																									
edu_millennial																									
patent																									
patent_percapita																									
clustereff																									
gdpgrowth																									
gdp_percapita																									
rentalrate																									
homevalue																									
pricetoincome																									
purchasingpower																									
commute																									
crime																									
climate																									
qualityoflife																									
medianage	1.00																								
diversity	-0.34	1.00																							
hispanic	-0.35	0.77*	1.00																						
white	0.34	-1.00**	-0.77*	1.00																					
asian	-0.07	0.48	0.17	-0.48	1.00																				
black	0.04	0.14	-0.35	-0.14	-0.21	1.00																			
poverty	-0.14	0.48	0.54*	-0.48	-0.24	0.16	1.00																		
unemployment	0.19	0.26	0.38	-0.26	0.06	-0.14	0.41	1.00																	
payroll_all	-0.02	0.26	0.13	-0.26	0.33	0.07	-0.14	-0.03	1.00																
payroll_tech	-0.02	0.28	0.12	-0.28	0.36	0.08	-0.18	-0.07	0.95**	1.00															
techratio_pay	-0.01	0.10	-0.06	-0.10	0.23	0.13	-0.39	-0.24	0.36	0.54*	1.00														
establishments_all	-0.01	0.28	0.17	-0.28	0.26	0.07	-0.09	0.00	0.98**	0.91**	0.32	1.00													
establishments_tech	-0.02	0.31	0.18	-0.31	0.30	0.08	-0.12	-0.02	0.96**	0.94**	0.42	0.98**	1.00												
techratio_est	-0.13	0.20	0.03	-0.20	0.33	0.07	-0.37	-0.27	0.37	0.50*	0.80**	0.35	0.47	1.00											
employment_all	-0.03	0.27	0.15	-0.27	0.28	0.08	-0.11	-0.02	0.99**	0.92**	0.34	0.99**	0.98**	0.36	1.00										
employment_tech	-0.03	0.28	0.13	-0.28	0.32	0.10	-0.16	-0.06	0.96**	0.99**	0.51*	0.94**	0.97**	0.48	0.96**	1.00									
techratio_employ	-0.07	0.11	-0.07	-0.11	0.31	0.13	-0.40	-0.26	0.43	0.62*	0.95**	0.38	0.49	0.84**	0.41	0.59*	1.00								
employchange	-0.34	0.11	0.25	-0.11	-0.01	-0.20	-0.03	-0.24	0.04	0.04	0.06	0.02	0.04	0.26	0.03	0.03	0.11	1.00							
techemploychange	-0.27	0.04	0.11	-0.04	-0.05	-0.08	-0.17	-0.22	0.02	0.04	0.27	0.03	0.05	0.29	0.03	0.05	0.26	0.51*	1.00						
jobcreation	-0.24	0.34	0.34	-0.34	0.25	-0.12	-0.08	-0.17	0.53*	0.51*	0.27	0.50*	0.53	0.43	0.54*	0.51*	0.32	0.52*	0.37	1.00					
jobcreation_percapita	-0.44	0.28	0.38	-0.28	0.21	-0.27	-0.09	-0.22	0.22	0.23	0.26	0.19	0.23	0.50*	0.21	0.22	0.33	0.77*	0.56*	0.72*	1.00				
estentries	-0.02	0.31	0.20	-0.31	0.27	0.06	-0.08	0.01	0.97**	0.90**	0.32	1.00**	0.99**	0.38	0.98**	0.94**	0.39	0.04	0.04	0.54*	0.22	1.00			
entries_percapita	0.03	0.04	0.03	-0.04	0.17	-0.09	-0.40	-0.15	0.42	0.42	0.48	0.46	0.52*	0.72*	0.43	0.44	0.52*	0.10	0.20	0.39	0.44	0.50*	1.00		
estexits	-0.01	0.29	0.18	-0.29	0.26	0.07	-0.07	0.01	0.97**	0.90**	0.31	1.00**	0.98**	0.35	0.98**	0.93**	0.37	0.01	0.03	0.50*	0.19	1.00**	0.47	1.00	
entries_exits	-0.31	0.18	0.28	-0.18	0.10	-0.19	-0.20	-0.05	0.15	0.17	0.29	0.15	0.20	0.48	0.17	0.18	0.32	0.53	0.37	0.53*	0.71*	0.20	0.56*	0.15	1.00

*Indicates moderate correlation where the coefficient is 0.50 or greater

**Indicates strong correlation where the coefficient is 0.80 or greater

Appendix 4. Component Correlations and Computed Weights

A4a. Principal Components, Rotated

	Comp1	Comp2	Comp3	Comp4	Comp5	Comp6	Comp7	Unexplained
lnpop	0.3857	-0.0159	-0.0469	0.0442	0.0403	0.0164	-0.0030	0.07827
airtraffic	0.3617	-0.0351	-0.0538	0.0600	-0.0071	0.1063	0.0141	0.1688
hubsize	0.3108	-0.0057	-0.0029	0.0105	0.0290	-0.0914	0.1736	0.26
geonetwork	-0.0082	-0.0820	0.1087	0.0206	-0.1931	0.0935	-0.6683	0.2061
conferences	0.1806	0.2513	-0.0925	0.0050	-0.0459	0.1656	0.0449	0.3259
degrees_science	-0.1162	0.2566	0.3018	-0.0189	0.0030	0.0711	-0.1160	0.2485
edu_millennial	0.0937	0.1730	0.0834	-0.0227	-0.2658	-0.2124	-0.0735	0.1506
patent_percapita	-0.1181	0.4377	0.0819	-0.0238	0.0019	-0.0680	0.0959	0.2409
clustereff	-0.1198	0.0445	0.3682	0.0692	0.1673	0.2045	-0.0090	0.3811
gdpgrowth	-0.0169	0.3959	-0.1840	0.1366	-0.0239	-0.2272	-0.0982	0.3232
gdp_percapita	0.1123	0.3109	0.0420	-0.0294	-0.0401	-0.2577	-0.0336	0.1163
rentalrate	0.0635	0.3721	-0.0236	-0.0284	-0.1159	0.2293	-0.0252	0.111
pricetoincome	-0.0891	0.0718	-0.0452	0.0413	-0.0666	0.5770	-0.0115	0.1962
qualityoflife	-0.2192	0.0089	0.1404	0.0062	-0.1306	-0.2351	0.2441	0.4023
commute	0.2345	0.0268	0.0939	0.0213	0.0547	0.2388	-0.1671	0.2309
crime	0.1133	-0.0594	0.0282	-0.2125	0.3833	-0.0956	-0.1870	0.3225
climate	-0.0064	0.0833	0.1917	-0.0648	0.3359	0.0377	0.0530	0.6137
medianage	0.0251	-0.0091	-0.0130	-0.3471	-0.2467	0.1883	0.1502	0.4636
diversity	0.0357	0.0981	0.0093	0.1103	0.4961	0.1061	0.0062	0.1575
asian	-0.0237	0.4252	0.0073	-0.0294	0.1437	0.1150	0.0428	0.2378
poverty	-0.0349	-0.0756	-0.0680	0.0192	0.4332	0.0132	0.0321	0.2993
unemployment	-0.0130	0.0049	-0.0403	-0.1122	0.1610	0.3532	0.0022	0.4479
lnpayroll_tech	0.3168	0.0318	0.1016	-0.0092	-0.0310	-0.0980	0.0018	0.03533
techratio_pay	0.0456	-0.0642	0.4857	-0.0329	0.0097	-0.0786	-0.0010	0.1307
lnest_tech	0.3544	-0.0100	0.0386	0.0141	-0.0227	-0.0077	0.0819	0.04083
techratio_est	0.0723	-0.0154	0.3834	0.0367	0.0171	-0.0441	0.2141	0.1113
lnemploy_tech	0.3420	0.0096	0.0669	0.0080	-0.0173	-0.0924	0.0122	0.03794
techratio_employ	0.0647	-0.0012	0.4384	-0.0112	0.0173	-0.1124	-0.0017	0.1016
employchange	-0.0268	-0.0193	-0.0670	0.4841	-0.0582	-0.0124	0.0191	0.2322
techemploychange	-0.0050	-0.1412	0.1133	0.3950	-0.0819	-0.0127	-0.1035	0.4329
jobcreation	0.1620	0.0395	-0.0403	0.3436	0.0585	0.0102	0.0163	0.3058
jobcreation_percapita	0.0268	0.0579	0.0053	0.4442	0.0362	-0.0208	0.1249	0.1028
entries_percapita	0.1553	-0.0080	0.1172	-0.0248	-0.1516	0.1033	0.4163	0.2154
entries_exits	0.0482	-0.0748	0.0961	0.2683	-0.0053	0.0697	0.3098	0.3179

A4b. Computed Components and Variable Weights

	Comp1	Comp2	Comp3	Comp4	Comp5	Comp6	Comp7	Weight (%)
lnpop	3.21	0.00	0.03	0.02	0.01	0.00	0.00	3.28
airtraffic	2.82	0.02	0.04	0.04	0.00	0.09	0.00	3.01
hubsize	2.08	0.00	0.00	0.00	0.01	0.07	0.19	2.35
geonetwork	0.00	0.08	0.15	0.00	0.31	0.07	2.84	3.46
conferences	0.70	0.80	0.11	0.00	0.02	0.23	0.01	1.86
degrees_science	0.29	0.83	1.12	0.00	0.00	0.04	0.09	2.37
edu_millennial	0.19	0.38	0.09	0.01	0.59	0.37	0.03	1.65
patent_percapita	0.30	2.42	0.08	0.01	0.00	0.04	0.06	2.91
clustereff	0.31	0.03	1.67	0.05	0.23	0.35	0.00	2.63
gdpgrowth	0.01	1.98	0.42	0.21	0.00	0.43	0.06	3.10
gdp_percapita	0.27	1.22	0.02	0.01	0.01	0.55	0.01	2.09
rentalrate	0.09	1.75	0.01	0.01	0.11	0.43	0.00	2.40
pricetoincome	0.17	0.07	0.03	0.02	0.04	2.75	0.00	3.07
qualityoflife	1.04	0.00	0.24	0.00	0.14	0.46	0.38	2.26
commute	1.19	0.01	0.11	0.01	0.02	0.47	0.18	1.98
crime	0.28	0.04	0.01	0.51	1.22	0.08	0.22	2.35
climate	0.00	0.09	0.45	0.05	0.94	0.01	0.02	1.55
medianage	0.01	0.00	0.00	1.35	0.50	0.29	0.14	2.31
diversity	0.03	0.12	0.00	0.14	2.04	0.09	0.00	2.42
asian	0.01	2.28	0.00	0.01	0.17	0.11	0.01	2.60
poverty	0.03	0.07	0.06	0.00	1.56	0.00	0.01	1.72
unemployment	0.00	0.00	0.02	0.14	0.21	1.03	0.00	1.41
lnpayroll_tech	2.17	0.01	0.13	0.00	0.01	0.08	0.00	2.39
techratio_pay	0.04	0.05	2.90	0.01	0.00	0.05	0.00	3.06
lnest_tech	2.71	0.00	0.02	0.00	0.00	0.00	0.04	2.78
techratio_est	0.11	0.00	1.81	0.02	0.00	0.02	0.29	2.25
lnemploy_tech	2.52	0.00	0.06	0.00	0.00	0.07	0.00	2.66
techratio_employ	0.09	0.00	2.36	0.00	0.00	0.10	0.00	2.56
employchange	0.02	0.00	0.06	2.62	0.03	0.00	0.00	2.73
techemploychange	0.00	0.25	0.16	1.75	0.06	0.00	0.07	2.28
jobcreation	0.57	0.02	0.02	1.32	0.03	0.00	0.00	1.96
jobcreation_percapita	0.02	0.04	0.00	2.21	0.01	0.00	0.10	2.38
entries_percapita	0.52	0.00	0.17	0.01	0.19	0.09	1.10	2.08
entries_exits	0.05	0.07	0.11	0.81	0.00	0.04	0.61	1.69
Percent of Variance	21.58%	12.63%	12.30%	11.19%	8.29%	8.27%	6.36%	80.62%

*Due to the high correlations among the variables in the data set, these weights are not necessarily used to determine the principal component descriptions, but are generally reflective of the results.
*Highlighted cells indicate the component to which the variable contributes the most variance.

Appendix 5. Entrepreneurial Attractiveness Survey Questions

- Introduction: *You are invited to take part in a research study conducted by Anna Kropf, an undergraduate student in Economics at Duke University. This survey aims to determine which economics-related factors are most important to students as they choose where to find jobs after college. The survey asks a series of questions separated into several sections; first about you yourself, and second about your future career's location. The survey is completely anonymous, with no link to your name or email, and should take less than 15 minutes to complete. To take part in the survey, please continue to the next section.*
- Demographic Questions:
    1. *What major(s) are you studying?*
    2. *What career or industry would you pursue after graduation?*
    3. *Could you see yourself as an entrepreneur, either now or in the future? Yes/No/Unsure*
    4. *What is the year of your expected graduation? 2018/2019/2020/2021/2022/2023*
    5. *Please Specify: Hispanic or Latino/Black or African American/Asian/American Indian or Alaska Native/Hawaiian Native or Pacific Islander/White/Other*
    6. *Please Identify: Female/Male/Other*
- Location Determinant Questions:
    - *In this section, please weight each of the following components according to how much importance they hold in your decision of job location within the United States. The weights should sum to 100. For example, if you were indifferent to all options, you would weight each with a value of 14.2. Factors can be negative if they negatively affect your choice of a city, as long as all of the weights still sum to 100. For example, if you rank one with a value of -100, the combined weights of all other components must sum to 200 (i.e. -100 + 200 = 100). Detailed descriptions are included. Please note: each component may include both positives and negatives according to your personal opinion; that is where you can use weights to decide which components include the factors that are most important to you.*
    - Size of the Economy: *This component describes the city as being very large in many regards, including population, airport size, diversity, and sheer number of establishments, payroll and employment. There are long commute times and high rental prices for living in the urban area. Further, the tech sector is large, along with high rates of education and diversity.*
    - Tourism Economy: *This component describes the city as having an economy with high tourism. The population is diverse, but there are high unemployment and poverty rates. There is high congestion indicated by long commutes, a well as high crime. The quality of life is low, but climate is good. Further, there is a low number of tech establishments and low educational attainment among the residents.*
    - Talent Infrastructure and Clustering: *This component describes the city as having a clustered economy, where the different industries are highly interconnected. There is high diversity and a high presence of tech and professional establishments. Further, the city is clustered among other talent centers, which allows for an extensive network of talent and movement.*
    - Skilled Workforce: *This component describes the city according to its talented workforce and companies. The educational attainment is high and the average income per person is high and growing; affordability is high. Poverty and unemployment rates are low, but diversity is low. The ratio of employment and payroll in tech and professional industries compared to all industries is high, so there is a large tech presence.*
    - Inability to Drive Growth and Innovation: *This component describes the city as having a prominent technology sector but one that has been unable to grab the attention of tech conferences and*

large companies. The tech sector does not drive the growth of per capita income but is largely interconnected with other local industry. The age of the population is older.

- Economic Instability and Old Industry: *This component describes the city's industry and engineering as old rather than innovative. The city has small airports and long commute times. The age of the population is relatively high, and unemployment is high. There is a lack of tech industry presence. The cost of living in the city is high, with high rental prices even relative to income and low purchasing power.*
- Geographic Isolation: *This component describes the city as being located in a very isolated area, where other cities are not a day drive away.*

- City choice
  - *If you could choose three cities in which to hold a job after college…*
    - *…which three cities would you prefer?*
    - *…and which three cities are least desirable?*

Appendix 6. Regression Estimates

A6a. Estimates of the Binary Logistic Model: Google

Variable	Coefficient	Std. Error	Z-Statistic
constant	-4.41	0.63	-7.01
C1: Size	0.62	0.29	2.11
C2: Tourism	-0.03	0.13	-0.21
C3: Talent Networks	0.22	0.36	0.62
C4: Skilled Workforce	-0.21	0.33	-0.63
C5: Inability to Drive Growth	-0.30	0.38	-0.78
C6: Economic Downturn	-1.11	0.35	-3.18
C7: Isolated Tech Growth	0.70	0.34	2.08
Obs.	99		
Log-Likelihood	-17.70		
Wald chi^2(7)	21.91		
Prob > chi2	0.0026		
Pseudo R^2	0.3003		

A6b. Estimates of the Binary Logistic Model: Amazon

Variable	Coefficient	Std. Error	Z-Statistic
constant	-3.76	0.69	-5.43
C1: Size	1.07	0.19	5.61
C2: Tourism	-0.40	0.26	-1.52
C3: Talent Networks	-0.49	0.40	1.24
C4: Skilled Workforce	0.07	0.20	0.35
C5: Inability to Drive Growth	-0.65	0.35	-1.84
C6: Economic Downturn	-0.88	0.40	-2.22
C7: Isolated Tech Growth	-0.02	0.31	-0.09
Obs.	99		

Log-Likelihood	-22.34
Wald chi^2(7)	39.05
Prob > chi^2	0.0000
Pseudo R^2	0.5080

A6c. Estimates of the Binary Logistic Model: Combined

Variable	Coefficient	Std. Error	Z-Statistic
constant	-3.94	0.82	-4.81
C1: Size	1.28	0.24	5.33
C2: Tourism	0.26	0.27	0.97
C3: Talent Networks	0.43	0.35	1.22
C4: Skilled Workforce	-0.19	0.23	-0.81
C5: Inability to Drive Growth	-0.53	0.42	-1.26
C6: Economic Downturn	-1.26	0.50	-2.53
C7: Isolated Tech Growth	0.24	0.39	0.61
Obs.	99		
Log-Likelihood	-19.48		
Wald chi^2(7)	46.02		
Prob > chi^2	0.0000		
Pseudo R^2	0.6285		

Appendix 7. Full Indexed Ranking of Metropolitan Areas in the United States

A7a. Indexed Ranking of Metropolitan Areas for Amazon, Weighted By Regression

Rank	Metro Area	Score
1	**Washington-Arlington-Alexandria, DC-VA-MD-WV**	34.55
2	**New York-Newark-Jersey City, NY-NJ-PA**	19.65
3	**Chicago-Naperville-Elgin, IL-IN-WI**	18.35
4	**Denver-Boulder, CO**	18.25
5	**Atlanta-Sandy Springs-Roswell, GA**	16.83
6	**Dallas-Fort Worth-Arlington, TX**	16.65
7	**Boston-Cambridge-Newton, MA-NH**	16.57
8	**Philadelphia-Camden-Wilmington, PA-NJ-DE-MD**	16.27
9	**Baltimore-Columbia-Towson, MD**	15.39
10	**Austin-Round Rock, TX**	14.13
11	**Minneapolis-St. Paul-Bloomington, MN-WI**	13.99
12	**Houston-The Woodlands-Sugar Land, TX**	13.17
13	**Charlotte-Concord-Gastonia, NC-SC**	12.54
14	**Seattle-Tacoma-Bellevue, WA**	12.32
15	**Los Angeles-Long Beach-Anaheim, CA**	10.67
16	**Kansas City, MO-KS**	10.60
17	**Phoenix-Mesa-Scottsdale, AZ**	10.55
18	**Raleigh, NC**	10.29
19	**Salt Lake City, UT**	9.69
20	**Pittsburgh, PA**	9.43
21	**San Francisco-Oakland-Hayward, CA**	9.35
22	**Miami-Fort Lauderdale-West Palm Beach, FL**	9.09
23	**Durham-Chapel Hill, NC**	8.91
24	**St. Louis, MO-IL**	8.06
25	**Indianapolis-Carmel-Anderson, IN**	7.69
26	**Detroit-Warren-Dearborn, MI**	7.61
27	**Tampa-St. Petersburg-Clearwater, FL**	7.54
28	**Portland-Vancouver-Hillsboro, OR-WA**	7.42
29	**Bridgeport-Stamford-Norwalk, CT**	6.85
30	**Madison, WI**	6.66
31	**Orlando-Kissimmee-Sanford, FL**	6.42
32	**San Diego-Carlsbad, CA**	6.10
33	**Nashville-Davidson-Murfreesboro-Franklin, TN**	5.83
34	**Columbus, OH**	5.70
35	**Cincinnati, OH-KY-IN**	5.60
36	**Hartford-West Hartford-East Hartford, CT**	5.44
37	**Omaha-Council Bluffs, NE**	4.79

38	**Richmond, VA**	4.75
39	**Cleveland-Elyria, OH**	3.65
40	**Albany-Schenectady-Troy, NY**	3.24
41	**Milwaukee-Waukesha-West Allis, WI**	3.04
42	**Des Moines-West Des Moines, IA**	2.89
43	**Virginia Beach-Norfolk-Newport News, VA**	1.95
44	**Jacksonville, FL**	1.43
45	**San Antonio-New Braunfels, TX**	1.17
46	**Colorado Springs, CO**	1.04
47	**Harrisburg-Carlisle, PA**	0.95
48	**Oklahoma City, OK**	0.85
49	**Buffalo-Cheektowaga-Niagara Falls, NY**	0.74
50	**New Orleans-Metairie, LA**	0.70
51	**Tulsa, OK**	-0.76
52	**Sacramento-Roseville-Arden-Arcade, CA**	-1.20
53	**Las Vegas-Henderson-Paradise, NV**	-1.36
54	**Louisville/Jefferson County, KY-IN**	-1.41
55	**San Jose-Sunnyvale-Santa Clara, CA**	-1.53
56	**Dayton, OH**	-1.71
57	**Charleston-North Charleston, SC**	-1.85
58	**Portland-South Portland, ME**	-3.12
59	**Baton Rouge, LA**	-3.16
60	**Rochester, NY**	-3.30
61	**Birmingham-Hoover, AL**	-3.31
62	**Grand Rapids-Wyoming, MI**	-3.45
63	**Columbia, SC**	-3.52
64	**Greenville-Anderson-Mauldin, SC**	-3.59
65	**Boise City, ID**	-3.74
66	**Akron, OH**	-3.86
67	**Provo-Orem, UT**	-4.04
68	**Little Rock-North Little Rock-Conway, AR**	-4.46
69	**Knoxville, TN**	-4.67
70	**Worcester, MA-CT**	-4.83
71	**Syracuse, NY**	-5.53
72	**Ogden-Clearfield, UT**	-5.73
73	**New Haven-Milford, CT**	-6.44
74	**Lancaster, PA**	-7.06
75	**Wichita, KS**	-7.28
76	**Oxnard-Thousand Oaks-Ventura, CA**	-7.33
77	**Providence-Warwick, RI-MA**	-7.65
78	**Allentown-Bethlehem-Easton, PA-NJ**	-7.99
79	**Toledo, OH**	-8.55
80	**Augusta-Richmond County, GA-SC**	-9.08

81	**Spokane-Spokane Valley, WA**	-9.22
82	**Chattanooga, TN-GA**	-9.80
83	**Jackson, MS**	-9.82
84	**Palm Bay-Melbourne-Titusville, FL**	-9.87
85	**Cape Coral-Fort Myers, FL**	-9.87
86	**Memphis, TN-MS-AR**	-10.16
87	**Scranton-Wilkes-Barre-Hazleton, PA**	-11.01
88	**Tucson, AZ**	-11.39
89	**Winston-Salem, NC**	-12.99
90	**Riverside-San Bernardino-Ontario, CA**	-14.06
91	**Springfield, MA**	-16.19
92	**El Paso, TX**	-17.82
93	**Honolulu, HI**	-17.89
94	**Deltona-Daytona Beach-Ormond Beach, FL**	-18.02
95	**Bakersfield, CA**	-21.57
96	**Fresno, CA**	-25.04
97	**Modesto, CA**	-25.11
98	**Stockton-Lodi, CA**	-30.83
99	**McAllen-Edinburg-Mission, TX**	-33.20

A7b. Indexed Ranking of Metropolitan Areas for Google, Weighted By Regression

Rank	Metro Area	Score
1	**Denver-Boulder, CO**	46.96
2	**Washington-Arlington-Alexandria, DC-VA-MD-WV**	46.04
3	**Kansas City, MO-KS**	41.52
4	**Dallas-Fort Worth-Arlington, TX**	39.44
5	**Minneapolis-St. Paul-Bloomington, MN-WI**	37.07
6	**Seattle-Tacoma-Bellevue, WA**	37.01
7	**St. Louis, MO-IL**	34.39
8	**Omaha-Council Bluffs, NE**	32.22
9	**Austin-Round Rock, TX**	31.75
10	**Atlanta-Sandy Springs-Roswell, GA**	31.27
11	**Pittsburgh, PA**	28.55
12	**Portland-Vancouver-Hillsboro, OR-WA**	27.97
13	**Chicago-Naperville-Elgin, IL-IN-WI**	27.32
14	**Boston-Cambridge-Newton, MA-NH**	24.61
15	**Durham-Chapel Hill, NC**	23.95
16	**Salt Lake City, UT**	23.61
17	**Raleigh, NC**	23.14

18	**Bridgeport-Stamford-Norwalk, CT**	22.16
19	**Nashville-Davidson-Murfreesboro-Franklin, TN**	22.04
20	**Philadelphia-Camden-Wilmington, PA-NJ-DE-MD**	21.29
21	**Houston-The Woodlands-Sugar Land, TX**	21.21
22	**Cleveland-Elyria, OH**	20.89
23	**Phoenix-Mesa-Scottsdale, AZ**	20.67
24	**Detroit-Warren-Dearborn, MI**	20.21
25	**Des Moines-West Des Moines, IA**	19.75
26	**San Francisco-Oakland-Hayward, CA**	18.91
27	**Oklahoma City, OK**	18.40
28	**Madison, WI**	18.08
29	**Indianapolis-Carmel-Anderson, IN**	16.38
30	**San Diego-Carlsbad, CA**	15.98
31	**Columbus, OH**	15.43
32	**San Jose-Sunnyvale-Santa Clara, CA**	15.37
33	**Miami-Fort Lauderdale-West Palm Beach, FL**	15.20
34	**Charlotte-Concord-Gastonia, NC-SC**	15.10
35	**Virginia Beach-Norfolk-Newport News, VA**	15.01
36	**Tulsa, OK**	14.83
37	**Orlando-Kissimmee-Sanford, FL**	14.32
38	**Cincinnati, OH-KY-IN**	13.78
39	**Milwaukee-Waukesha-West Allis, WI**	12.97
40	**Buffalo-Cheektowaga-Niagara Falls, NY**	12.54
41	**Tampa-St. Petersburg-Clearwater, FL**	12.49
42	**Baltimore-Columbia-Towson, MD**	11.01
43	**Los Angeles-Long Beach-Anaheim, CA**	10.62
44	**Little Rock-North Little Rock-Conway, AR**	9.99
45	**Richmond, VA**	9.71
46	**Albany-Schenectady-Troy, NY**	8.99
47	**New York-Newark-Jersey City, NY-NJ-PA**	6.90
48	**Jacksonville, FL**	6.84
49	**Birmingham-Hoover, AL**	6.09
50	**San Antonio-New Braunfels, TX**	5.51
51	**New Orleans-Metairie, LA**	5.27
52	**Boise City, ID**	4.25
53	**Rochester, NY**	4.20
54	**Louisville/Jefferson County, KY-IN**	3.74
55	**Colorado Springs, CO**	1.82
56	**Las Vegas-Henderson-Paradise, NV**	1.38
57	**Dayton, OH**	0.87
58	**Charleston-North Charleston, SC**	0.01
59	**Wichita, KS**	-2.84
60	**Grand Rapids-Wyoming, MI**	-3.25

61	**Harrisburg-Carlisle, PA**	-5.17
62	**Knoxville, TN**	-5.34
63	**Akron, OH**	-5.93
64	**Jackson, MS**	-6.08
65	**Greenville-Anderson-Mauldin, SC**	-6.09
66	**Portland-South Portland, ME**	-7.75
67	**Memphis, TN-MS-AR**	-8.03
68	**Sacramento-Roseville-Arden-Arcade, CA**	-8.13
69	**Provo-Orem, UT**	-10.03
70	**Spokane-Spokane Valley, WA**	-10.70
71	**Syracuse, NY**	-11.03
72	**Columbia, SC**	-11.29
73	**Baton Rouge, LA**	-11.79
74	**Hartford-West Hartford-East Hartford, CT**	-12.08
75	**Palm Bay-Melbourne-Titusville, FL**	-13.04
76	**Toledo, OH**	-14.52
77	**Chattanooga, TN-GA**	-16.47
78	**Lancaster, PA**	-19.87
79	**Tucson, AZ**	-23.18
80	**Oxnard-Thousand Oaks-Ventura, CA**	-23.68
81	**Ogden-Clearfield, UT**	-24.33
82	**Providence-Warwick, RI-MA**	-26.53
83	**Allentown-Bethlehem-Easton, PA-NJ**	-26.86
84	**Augusta-Richmond County, GA-SC**	-27.41
85	**New Haven-Milford, CT**	-27.65
86	**Winston-Salem, NC**	-27.74
87	**Scranton-Wilkes-Barre-Hazleton, PA**	-28.04
88	**Cape Coral-Fort Myers, FL**	-29.19
89	**Honolulu, HI**	-33.99
90	**Worcester, MA-CT**	-36.56
91	**El Paso, TX**	-38.64
92	**Fresno, CA**	-48.26
93	**Springfield, MA**	-50.03
94	**Riverside-San Bernardino-Ontario, CA**	-53.94
95	**Bakersfield, CA**	-58.02
96	**Deltona-Daytona Beach-Ormond Beach, FL**	-60.23
97	**McAllen-Edinburg-Mission, TX**	-72.07
98	**Modesto, CA**	-76.23
99	**Stockton-Lodi, CA**	-85.03

References

Amazon.com, Inc. (2017). Amazon announces 238 proposals for HQ2 from across North America. Retrieved from https://www.amazon.com

Audretsch, David B., Erik E. Lehmann, and Susanne Warning. (2005). University Spillovers and New Firm Location. *Research Policy,* 34(7), 1113–122. Doi: 10.1016/j.respol.2005.05.009

Bomey, Nathan. (2017, October 21). See the Cities Hoping to Land 50,000 Amazon Jobs, Headquarters. *USA Today*. Retrieved from https://www.usatoday.com

Bunten, Devin et al. (2015). Entrepreneurship, Information, and Growth. *Journal of Regional Science*, 55(4), 560–84. Doi: 10.1111/jors.12157

The Capital Factory. (2017). Capital Factory - Austin's Center of Gravity for Entrepreneurship. Retrieved from https://www.capitalfactory.com

Cirillo, Christopher M. (2013, April 28). Birth of an Idea: The Creation of Research Triangle Park and Its Sustained Economic Impact on the Research Triangle Area. *Urban Economics*, Duke University.

The City of Austin. (2018). Innovation Office Programs. Retrieved from http://www.austintexas.gov

Collier, Sabrina (2017, June 7). Top Universities in the US 2018. Retrieved from https://www.topuniversities.com/university-rankings-articles/world-university-rankings/top-universities-us-2018

Delgado, Mercedes, Michael E. Porter, and Scott Stern. (2010). Clusters and Entrepreneurship. *Journal of Economic Geography,* 10(4), 496–518. Doi: 10.1093/jeg/lbq010

Etzkowitz, Henry. (2008). *Triple Helix: University-industry-government Innovation in Action*. Routledge.

Fairelie, Robert W., Arnobio Morelix, E.J. Reedy, and Joshua Russell. (2015, June). *The Kauffman Index: Startup Activity: National Trends*. The Kauffman Foundation.

FBI. (2015, August 17). Crime in the United States. *U.S. Department of Justice*. Retrieved from https://ucr.fbi.gov

Federal Aviation Administration. (2017, October 10). Passenger Boarding. *U.S. Department of Transportation*. Retrieved from https://www.faa.gov/airports/planning_capacity/passenger_allcargo_stats/passenger/

Fell, Jason. (2012, September 27). How Google Wants to Make Starting Up Easier for Entrepreneurs. *Entrepreneur Media, Inc.* Retrieved from https://www.entrepreneur.com/article/224522

Florida, Richard L. (2014, February 11). What Cities Really Need To Attract Entrepreneurs, According to Entrepreneurs. *CityLab*. Retrieved from https://www.citylab.com

Florida, Richard L. (2012). *The Rise of the Creative Class: Revisited*. New York, NY: Basic.

Galvanize. (2017). Galvanize. Retrieved from https://www.galvanize.com

Garfield, Leanna. (2018, January 11). Amazon Could Detonate a Gentrification 'Prosperity Bomb' in the City It Chooses for Its Second Headquarters. *Business Insider*. Retrieved from http://www.businessinsider.com

GEDI. (2017). Global Entrepreneurship Index. *Global Entrepreneurship Development Institute*. Retrieved from https://thegedi.org/global-entrepreneurship-and-development-index/

Giotopoulos, Ioannis, et al. (2016). Drivers of High-Quality Entrepreneurship: What Changes Did the Crisis Bring about? *Small Business Economics*, 48(4), 913–930. Doi: https://doi-org.proxy.lib.duke.edu/10.1007/s11187-016-9814-x

Google, Inc. (2011). Google For Entrepreneurs. Retrieved from https://www.googleforentrepreneurs.com

Google Inc. (2017). Google Maps. 2017. Retrieved from https://www.google.com/maps

Granovetter, Mark. (2005). The Impact of Social Structure on Economic Outcome. *The Journal of Economic Perspectives*, 19, 33–50. Retrieved from http://www.jstor.org/stable/4134991

Green, Dennis. (2017, October 18). These 10 Cities Have the Best Chance of Getting Amazon's $5 Billion Headquarters, According to the Data. *Business Insider*. Retrieved from http://www.businessinsider.com

The Kauffman Foundation. (2016). The Kauffman Index. Retrieved from https://www.kauffman.org/kauffman-index/

Klein, Adam. (2017, April 12). Email Interview with Adam Klein of American Underground.

McGraw, Marquise J. (2017). The Heterogeneous Impact of Airports on Population and Employment Growth in Cities. Self-published, 1–53.

Mooney, Attracta. (2016, February 19). Philadelphia's $5.7bn 'Quiet Crisis'. *The Financial Times*. Retrieved from https://www.ft.com

Mulvey, Jeanette. (2012, September 27). Top Exec Explains Google for Entrepreneurs. *Fox Business*. Retrieved from https://www.foxbusiness.com

National Science Foundation. (2015). Rankings by Total R&D Expenditures. Retrieved from https://ncsesdata.nsf.gov

Numbeo. (2017). Cost of Living. Retrieved from https://www.numbeo.com

O'Connell, Jonathan, and Darla Cameron. (2018, March 9). The Amazon Finalists Are Already Building Tech Economies. Which Are the Best? *The Washington Post*. Retrieved from http://www.newsobserver.com

Open Data Network. (2017). Gross Domestic Product. *Socrata*. Retrieved from https://www.opendatanetwork.com

Otet, Ama. (2016, June 23). Average Apartment Size in the US: Atlanta Has Largest Homes. *RENTCafe*. Retrieved from https://www.rentcafe.com

Pazzanese, Christina. (2018, February 26). The Quest to Win over Amazon, as Outlined by a Harvard Professor. *Harvard Gazette*. Retrieved from https://news.harvard.edu/gazette

Pollard, Matthew, and Ed Frauenheim. (2016, October 10). Why Austin Is the Rockstar of Small-Business Cities. *Fortune*. Retrieved from http://fortune.com

Porter, Michael E. (1990). *The Competitive Advantage of Nations*. Brighton, MA: Harvard Business Review.

Procopio, Joe. (2013, September 25). Why You Should Have Seen the Google Tech Hub at American Underground Coming. Retrieved from https://www.wraltechwire.com

The Research Triangle Park. (2017). RTP Company Directory. Retrieved from https://www.rtp.org/rtp-company-directory

Reyes, Jose Daniel and Roberts, Mark and Xu, Lixin Colin. (2017). The Heterogeneous Growth Effects of the Business Environment: Firm-Level Evidence for a Global Sample of Cities. World Bank Policy Research Working Paper No. 8114. The World Bank.

Sperling, Bert, and Ted Sperling. (2017, November 15). Atlanta Will Be Amazon's New Home. *Sperling's BestPlaces*. Retrieved from https://www.bestplaces.net

Steiber, Annika, and Sverker Alänge. (2013). The Formation and Growth of Google: A Firm-level Triple Helix Perspective. *Social Science Information*, 52(4), 575–604. Doi: 10.1177/0539018413497833

U.S. Bureau of Labor Statistics. (2016, April 15). "Unemployment Rates for Metropolitan Areas." *U.S. Department of Labor*. Retrieved from https://www.bls.gov

U.S. Census Bureau. (2012, September 1). 2010 Census Gazetteer Files. *US Department of Commerce*. Retrieved from https://www.census.gov/geo/maps-data/data/gazetteer.html

U.S. Census Bureau, Center for Economic Studies. (2014). Business Dynamics Statistics. *U.S. Department of Commerce*. Retrieved from https://www.census.gov/ces/dataproducts/bds/

U.S. Census Bureau. (2017, March 15). Metropolitan and Micropolitan Statistical Area Codes. *U.S. Department of Commerce*. Retrieved from https://www.census.gov

U.S. Census Bureau. (2017). National Population Totals Tables: 2010-2016. *U.S. Department Of Commerce*. Retrieved from https://www.census.gov

U.S. Census Bureau. (2015). UNITED STATES QuickFacts. *U.S. Department of Commerce*. Retrieved from https://www.census.gov

U.S. Cluster Mapping Project. (2014). Data by Region. *Harvard Business School*. Retrieved from https://www.clustermapping.us/cluster

U.S. News. (2018). The Best MBA Programs in America, Ranked. *U.S. News & World Report*. Retrieved from https://www.usnews.com

Vyas, S., and L. Kumaranayake. (2006). Constructing Socio-Economic Status Indices: How to Use Principal Components Analysis. *Health Policy and Planning*, 21(6), 459–68. Doi: 10.1093/heapol/czl029

Wessner, Charles W. (2009). *Understanding Research, Science, and Technology Parks: Global Best Practices: Report of a Symposium*. National Academies Press.

Ye, Victor Yifan, and Mia Mikic. (2016, June). "The Index of Asia-Pacific Regional Integration Effort. *East Asian Economic Review,* 20(2), 129–168. Retrieved from http://dx.doi.org/10.11644/KIEP.EAER.2016.20.2.307

Zillow Group. (2017). Zillow Research – Data. *Zillow Research*. Retrieved from https://www.zillow.com/research/