IHS
Making cities work

IHS is the international institute of urban management of Erasmus University Rotterdam

MSc Programme in Urban Management and Development

Rotterdam, The Netherlands

September 2018

Thesis

Title: Regional Innovation and Air Passenger Network

Name: Fitri N Solehuddin

Supervisor: dr. A.S. (Ajay) Bhaskarabhatla

Specialization: Urban Competitiveness and Resilience

UMD 14

MASTER’S PROGRAMME IN URBAN MANAGEMENT AND DEVELOPMENT

(October 2017 – September 2018)

Regional Innovation and Air Passenger Network

Fitri N Solehuddin

Indonesia

Supervisor: dr. A.S. (Ajay) Bhaskarabhatla

UMD 14 Report number: 1175

Rotterdam, September 2018

Summary

City competitiveness is an important subject to study since it determines the productivity and prosperity of (Cann, O., 2018). There is also a strand of thought that competitiveness reveals the quality of life (Rogerson, 1999). Later on, Ni, Kamiya, et al. (2017) suggest a comprehensive conception in the recent report of global urban competitiveness through six fundamental pillars: enterprise quality, local elements, local demand, software environment, hardware environment, global connection. Furthermore, they emphasize that innovation in technology will drive global urban competitiveness. Therefore, in this thesis research, I perceive urban competitiveness from the perspective of regional innovation capability.

Moreover, the air transportation network also plays an important role in urban competitiveness since the air transportation facilitates physical interaction of firms and market (Mukkala and Tervo, 2013). First, the availability of air infrastructure is one of the locational factors for firms to place subsidiary offices or production facilities for business expansion purposes. Secondly, air passenger flow allows knowledge transfer that leads to the development of innovation. Lastly, Kalayci and Yanginlar (2016) indicate a long-term positive relationship between air transportation, FDI, and economic growth. Hence, it is important to investigate the correlation between air transportation network and the cities competitiveness.

In this thesis research, I would like to discuss to what extent air transportation network affect regional innovation and how agglomeration influence the relation. I observe the relation between air transportation network and regional innovation of the United States (US) in Metropolitan Statistical Area (MSA) level in the year 2012. Moreover, I conduct three analysis to develop the discussion in this thesis research. First, I employ descriptive statistical analysis to explain the distribution of innovation. Secondly, I convey network analysis to define the status of each MSA in air transportation network by generating three network measurement, according to the social network approach by Freeman (1978): degree centrality, betweenness centrality, and closeness centrality. Lastly, I conduct correlation analysis to reveal the relationship between regional innovation and region's status in air transportation network (measured by degree centrality), and to find out how firms’ agglomeration influences the relation.

All in all, there are several findings resulted from this study. First, the distribution of regional innovation in the United States in 2012 is highly skewed, with two centers of concentration which are the north-east cluster and west cluster. Secondly, as the network analysis result, New York-Northern New Jersey-Long Island (NY-NJ-PA) appears as the MSA with the highest degree of centrality, while, Anchorage (AK) is the MSA with the highest betweenness centrality. Unexpectedly, South Bend-Mishawaka (IN-MI) appears as the MSA with the highest closeness centrality. Third, the relation between regional innovation and region status in air passenger network is statistically significant and shows the positive correlation. Lastly, agglomeration level is also a significant factor which influences the regional innovation. However, the interaction between air passenger and agglomeration levels weaken the influence of air passenger network on regional innovation.

Keywords

Urban innovation, regional innovation, air transportation, air passenger network, urban network, regional network.

Acknowledgments

First and foremost, I would like to praise Alloh SWT, Who gives me the strength and ability to undertake and persevere the whole UMD master program and finally conduct this thesis research. Without His Blessing, this would never happen.

Second, I would like to thank people from the Government of Netherland, who grant me the Netherland Fellowship Program (NFP) scholarship so that I could pursue my master degree at IHS-Erasmus University Rotterdam.

Third, I would also like to thank the following special people:

dr. A.S. (Ajay) Bhaskarabhatla, my supervisor, who has helped me to understand the theory behind this thesis topic and guided me during the process of this thesis development through frequent discussions.

The IHS team, especially UCR team members who have been working hard in facilitating the UMD program and have been willing to share their knowledge and to help us understand the course lesson.

My UMD friends from around the world, especially from Indonesia, who made this course full of joy. Thank you for the laugh, the stories and, of course, the food. It was a pleasure to taste many unique cuisines from each country.

Last but not least, I would like to express my deepest gratitude to my beloved family, my fathers, and mothers, my brothers, and sisters, especially my husband and my son, who always support me with their kindness and love. I dedicate this achievement to them.

Fitri N Solehuddin

Abbreviations

2SLS	Two-Stage Least Squares
AK	Alaska
ANC	Anchorage International Airport
BRCI	Brazil Russia China and India
CA	California
DC	District of Columbia
DE	Delaware
EU	European Union
EWK	Newark International Airport
FBI	Federal Bureau of Investigation
FDI	Foreign Direct Investment
FL	Florida
GA	Georgia
GDP	Gross Domestic Product
GIN	Global Innovation Network
GWR	Geographically Weighted Regression
HTS	High-Speed Train
IHS	Institute of Housing and Urban Development
IL	Illinois
IN	Indiana
IN	Indiana
JFK	John F Kennedy International Airport
MA	Massachusetts
MD	Maryland
MI	Michigan
MN	Minnesota
MNC	Multi-National Company
MSA	Metropolitan Statistical Area
NAICS	North American Industry Classification System
NFP	Netherland Fellowship Program
NH	New Hampshire
NJ	New Jersey
NOAA	National Oceanic and Atmospheric Administration
NY	New York

OD	Origin-Destination
OECD	Organization for Economic Cooperation and Development
OLS	Ordinary Least Square
PA	Pennsylvania
R&D	Research and Development
SME	Small-Medium Enterprise
TX	Texas
UCR	Urban Competitiveness and Resilience
UMD	Urban Management and Development
US	United States
USPTO	The United States Patent and Trademark Office
VA	Virginia
VC	Venture Capital
WA	Washington
WI	Wisconsin
WV	West Virginia

Table of Contents

List of Figures

List of Tables

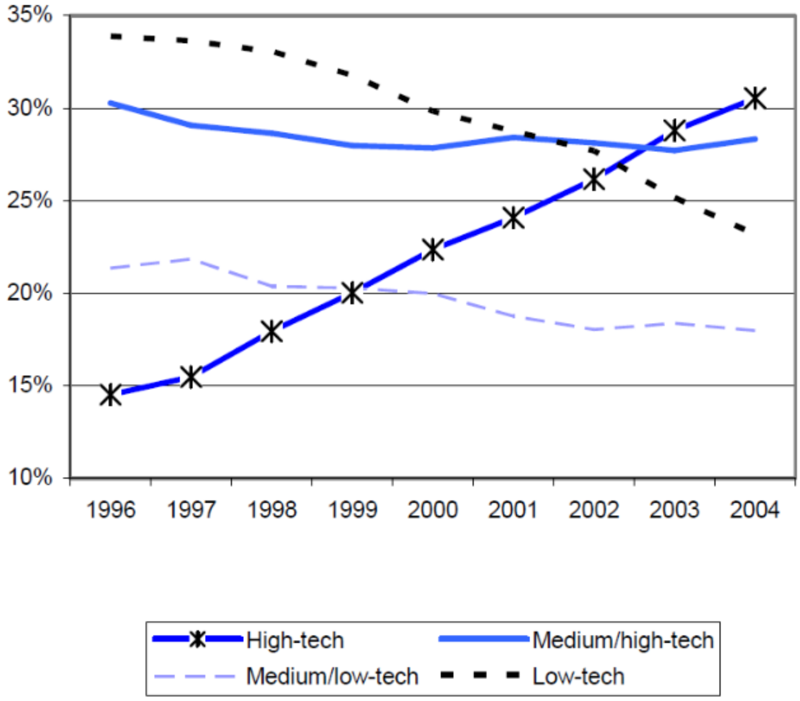
Chapter 1: Introduction

1.1 Background

Globalization reforms the way cities interact with each other through global competition and cooperation (Begg, 1999) that forced economic development (Rondinelli, Johnson Jr, et al., 1998). To win the global competition, cities must improve the quality of their highly skilled labor and the capacity for technological development. Meanwhile, to cooperate in the global network, cities should improve the standard of living, modernize the city services and extend the public infrastructures (Rondinelli, Johnson Jr, et al., 1998). Therefore, the global competition and cooperation among the cities are commonly associated with the global city competitiveness (Ni, Kamiya, et al., 2017, Taylor, 2014a).

Furthermore, the regional innovation and the global city network are the driving factors in determining global city competitiveness (Ni, Kamiya, et al., 2017). Firstly, the innovation level of a region determines how well a region could keep up in the globalization pace. Scholars argue that innovation can influence GDP (Soo, 2018, Sokolov-Mladenović, Cvetanović, et al., 2016) and market value (Hall, Jaffe, et al., 2005). For instance, Sokolov-Mladenović, Cvetanović, et al. (2016) find that 1% share on R&D expense in GDP increases 2.2 % GDP growth of EU countries during 2002-2012. While Hall, Jaffee, et al. (2005) find that one citation on patent improves 3% of its economic value in the market.

Figure 1 The changing structure of BRIC's1 manufacturing trade by technological intensity



Source: (OECD Innovation, 2007, p.8)

In fact, innovation also causes the structure of manufacturing trade in Brazil, Russia, China and India (BRCI) from 1996 to 2004 (OECD Innovation, 2007). According to Figure 1, initially in 1996, the sales of low-tech and medium/high tech product dominates the overall trade by having approximately 35% and 30% share consecutively. However during 2002 to 2003, a significant shift on BRCI's manufacturing trading occurs. Then, finally in 2004, the high-tech and medium/high-tech manufacturing product lead the trade share by taking account of around 30% and 25% of the overall BRCI’s trade share. In this case, there is the significant increase in high-tech manufacturing trade that significantly transforms BRCI’s economic structures.

Secondly, global city network plays an important role in urban competitiveness since it enables cities to develop across the boundaries. I will particularly see the global city network from the air transportation network since it facilitates the physical interaction between firms and market (Mukkala and Tervo, 2013); the physical interaction remains crucial to maintaining the communication quality of global interaction. Additionally, the air infrastructure availability is one of the locational factors for firms to expand their businesses. In this case, cities will get spillover effect from the firm’s expansion such as the job opportunity, knowledge spill-over and the taxes that the firms should pay to the cities.

In the United States (US) context, air transportation significantly contributes to the national economic growth. Oxford Economics (2011) reveal that 4.9% of the US GDP comes from the aviation sector and 9.3 million jobs are provided in the aviation sector in 2010 (6.8% of the workforce). Additionally, the direct taxes and social security payments from the aviation sector exhibit a substantial number which is $57.4 billion per annum. In fact that the air transportation sector footprints in US economy exhibit the significant influence, the Oxford Economics (2011) predict that the aviation sector will sustain the long-term of US economy. Thus, it is worthy to pay a big concern on the air transportation sector.

In this thesis research, I would like to profoundly discuss the correlation between regional innovation and air transportation network. I think it is essential to investigate the relation since both regional innovation and air transportation network are the crucial aspects of city competitiveness. In fact, there is a lack of discussion on exploring the relationship between these matters. I presume that both aspects will strongly relate to each other. Therefore, by exploring the relationship between regional innovation and air transportation network, the urban actors can define comprehensive measures to improve the regional competitiveness.

1.2 Problem statement

In this study, I would like to examine the effect of the region’s status in air passenger network on regional innovation. I argue that the air passenger network that commonly associated with the concentration of employment (Button, Lall, et al., 1999) and economic activities (Schaafsma, 2003) will create a favorable environment that attracts ‘creative class’ (Florida, 2002), in this case, talents in scientific research and development services. The region’s status alters to be an important factor in the locational decision in generating innovation. Therefore, the concentration of innovation will boost economic growth that contributes to improve regional competitiveness (Florida, 2002).

In the recent Global urban competitiveness report, Ni, Kamiya, et al., (2017) incorporate the patent index and global connection as the elements that determine the urban competitiveness index. They use the patent index to measure the index of local element pillar since they believe that patent can build the pleasant business environment for companies to grow. A supporting argument would be from Soo (2018) who argue that patent activities and GDP and population have a positive relationship. Besides patent, other scholars also use R&D instrument to measure the regional innovation. For example, Parisi, Schiantarelli, et al. (2006) find a positive correlation between R&D and labor productivity. Thus, innovation is significant in determining urban competitiveness.

Furthermore, the number of air routes is another underlying indicators of global connectivity pillar that represents worldwide interaction in the global network. For instance, air transportation activities trigger the presence of multinational firms in the cities that bring investment (Bannò and Redondi, 2014), economic activities, as well as job opportunities. Therefore, the amount of air transportation intensity could indirectly influence the economic development in the long-term period (Kalayci and Yanginlar, 2016). The air transport activities are another form of 'injection' for the local economy and to some extent even transform the economic structure (Button and Yuan, 2013).

Moreover, even though some studies suggest that there is a strong relationship between regional innovation and urban development as well as air transportation and urban development, there has not been any study yet which discuss the direct relations between regional innovation and air transportation network. I believe that the region's position in air transportation network may attract the creative class to gather in the city, share their knowledge with the city inhabitants and generates creative and innovative movement in the city. Therefore, the investigation of the relationship between the air transportation network and innovation in the city is essential for urban development.

1.3 Research objectives

The major goal of this research thesis is to reveal the information about how well connected a region is in the air transportation network system and to test the correlation between region's positions in air transportation network with its competitiveness in innovation. Therefore, the derivative goals are:

a. to identify the openness and connectedness of a region according to the air transportation network;
b. to investigate the correlation between air transportation-related factors with regional innovation.

1.4 Provisional research question(s)

The main provisional research question is to what extent does the air transportation network influence regional innovation in term of R&D and patent activities?
To solve the main question, these following sub-questions are necessary to address:

a. How is the distribution of regional innovation?
b. How connected a region is in the air transportation network?
c. What is the correlation between regional innovations, with its air transportation network? To what extent the connectedness of a region will influence R&D and patent activities?

1.5 Significance of the study

The research work contributes to enriching the current state of knowledge on how air transportation activities influence urban competitiveness, particularly in regional innovation aspect. I perceive the regions openness and connectedness from the physical accessibility of humans and goods transfer through the air transportation network. I expect this study may reveal the degree of the region's connectivity from their potential resources in the air transportation network. It is necessary to describe the region connectedness to understand how a region interacts with each other and get benefit from this interaction. Thus, Government can formulate relevant policies to help the weak cities become more attainable or to reveal the potential markets to expand the current investment.

1.6 Scope and limitation

The scope of the subject of this study is the relationship between regional innovation and air transportation network. For this study, I particularly focus on R&D and patent activities to measure the level of innovation in a region. Besides, I will particularly address air passenger network to define region status in air transportation network. I exclude the correlation between regional innovation and other transportation networks such as ground transportation and sea transportation network even though these relations might give significant impact to regional innovation. Hence, I suggest addressing the correlation between regional innovation and the whole transportation network in the future research.

The accessibility of the data becomes the major concern to delineate the scope of work of this study. Indeed, most of the previous studies on air transportation utilized the US air transportation database which is open to the public. I use the cross-sectional data of 2012, at the Metropolitan Statistical Area (MSA) level which is available at US government official website. I chose to do this study using the database in the year 2012 due to its complete information on variables that relevant for this research.

Lastly, I will mainly use OLS-regression to describe the relationship. Thus, I particularly focus on exercising the influence of air transportation network on regional innovation. In addition, I include the 2SLS regression result and Geographically Weighted Regression (GWR) as the robustness check to compare their results with the OLS result.

Chapter 2: Literature Review / Theory

2.1. Literature review

Many scholars have questioned the uneven economic power distribution over the world for decades, which lead to the conclusion that worldwide competition among the regions exists. Porter (1990) distinguishes the competition among the nations through the industry capability in satisfying factor and demand precondition. Using the prominent diamond diagram, he defines influential determinants to describe national comparative advantage that bolsters company capability to innovate and compete in the global market. Later on, Begg (1999) argues that competition exists alongside cooperation. The cooperation is essential to overcome the uncertainty in business circumstance through information exchange among the institutions (Begg, 1999). He suggests that economic development policy is needed in regional and urban level to help the cities adapt to changes. Thus, Taylor (2014b) associates competition with the hierarchical relationship of cities in term of economic or political power, while he associates cooperation with the horizontal links among the cities based on mutual relationship. Furthermore, he argues that competition does not befall among the cities, yet only occurs when cities share the same, geographical or sectoral, a niche in the market (Wall, 2009). Then, the latest study of The Global Competitiveness Report 2017 summarizes the existing determinant aspect of cities' competitiveness into a more comprehensive concept of urban competitiveness which incorporates both competition and coordination aspect (Ni, Kamiya, et al., 2017).

Table 1 Variables and Indicators of Global Urban Competitiveness 2017

Variables	Indicator
Enterprise quality	- number of multinational companies - Forbes 2000 index - industrial structure and industrial quality
Local demand	- population size - GDP - Per capita income
Local element	- patent index - unemployment rate - bank index - university index
Hardware environment	- PM2.5 emissions - benchmark hotel price - convenience roads and distance to the sea
Software environment	- crime rate - language diversity index - ease of doing business - the ratio of taxation by central versus local government
Global connection	- multinational corporation index - international reputation index - the number of air routes

Source: Ni, Kamiya, et al., (2017)

According to Tabel 1, the recent Global Competitiveness Report 2017, Ni, Kamiya, et al., (2017) include innovation level and air connectivity as two of the essential factors that determine urban competitiveness. Innovation (patent index) is one of the indicators that measures the quality of the local element of a region. The local element describes business environment quality. A region with a high level of innovation indicates the easiness for companies to do their businesses which result in the company high productivity. Moreover, the air connectivity (the number of air routes) is one of the indicators to measure urban connectivity with the global system. Air connectivity plays an important role in enabling economic, political and cultural connection to create a market for company's product. The more connected a city through the air transportation network, the bigger potential market that can be reached by its companies. All in all, the interplay between regional innovation and air transportation performance bolster the companies' productivity. Thus such interaction may significantly induce the regional competitiveness level.

2.1.1 Innovation and city competitiveness

There are several ways to conceptualize the notion of regional innovation. Griliches (1998) define innovation through R&D investment as the input and patents generation as the output of innovative activities. Meanwhile, Hall, Lotti, et al. (2009) distinguish innovation into two building blocks: the process of inventive activities and the product of innovation. They explore innovation from both aspects to investigate how and when the process of innovation occurs and explain how the products of innovation can contribute to company's productivity, in this case in small-medium enterprises (SME's) context. In a later study, Carlino and Kerr (2015) divide the concept of innovation into two parts: the idea generation part and the commercialization part of idea application. Through the process of commercialization, they perceive innovation as the process of knowledge transfer because of the circulated idea of new invention among the human capitals. In summary, it is obvious that scholars are conversant with the distinction between the product and the process of innovation.

Scholars have been utilizing a various approach to understand how innovation pertains to urban development. Carlino and Kerr (2015) summarize the approaches into three ways of innovation measurements:

a. *Investments in the innovation process.* The indicators for this measurements can be R&D employment and expenditures or firms fund start-up companies (VC investments).
b. *Patent and citations.* This approach measures the direct outcome of the innovation process through the patent database. The indicators for this approach can be the number of patents generated in a region, and the number of citation that a region make from the other region.
c. *Literature-based indicators.* This approach is according to the last output of the innovation stage namely product announcements in trade, engineering, and technical publication.

The first measurement of innovation is the amount of investment in the innovation process which commonly associated with R&D-related indicators. Parisi, Schiantarelli, et al. (2006) use R&D expenditures and new product introduction to determine innovation in the city. They find R&D expenditure strongly influence the probability of the new product introduction. Thus, they argue that R&D activities facilitate the new technology absorption so that it induces labor productivity growth. Bettencourt, Lobo, et al. (2007) measure innovation by the number of inventors and R&D employment. They consider innovation as the product and the process of inventive activities. Buzart and Carlino et al. (2016) utilized the combination of R&D lab location data and patent data to map the concentration of innovation in the US. They argue that the agglomeration of people and jobs facilitates the localized knowledge spillover. Hence, the agglomeration creates innovative movement.

Besides R&D-related indicators, scholars also use the ratio of VC investment per area metric to measure investment in innovation. The VC investment account for VC firms funds on start-up company in a region. Ferrary and Granovetter (2009) investigate the contribution of VC firms in Silicon Valley context. They argue that VC firms support the robustness of innovation network systems, and they point out financing as one out of five VC firms contribution to innovation. However, VC investment only applies to specific technological area and types of the firm (Carlino and Kerr, 2015).

Despite the clarity in representing the innovation process, investment in innovation also has several drawbacks. Indeed, the investment in innovation could not capture whether the innovation input activities effectively result in the successful inventions or not. The investment in innovation only records the process before product inventors. Some of the investment might result in the new products that have economic value. However, some might result in a product with no economic value or even might not give any result. Another drawback is regarding R&D data. The collection of R&D data at the local level is more difficult than the collection of patent data. R&D data collection requires a highly confidential survey to access the detail information about the innovation process. Thus, the data on R&D which openly are mostly available in a.

Secondly, Patent-related indicators are the prominent measurement of innovation. Counting patent rights is a decent way to indicates innovation in the cities. Griliches (1998) used the number of patents to differentiate the inventive activities of companies. However, this approach could not capture the significance of an invention. It assumed all inventions have the same value, whereas, the value of inventions are significantly divergent. Another study used citation patent to measure the significance of innovation to economic value. Hall, Jaffe, et al. (2005) examine the relationship between the quality of innovation (represent by patent-citation indicators) and market value. They found that one quality of citation on a patent can boost 3% of company market value.

Nevertheless, it takes times to gain the citation on the patent which causes this approach will irrelevant to analyze present condition. The other approach is by using the patent index to determine the level of the technological cutting edge of the cities. Ni, Kamiya, et al. (2017) used the patent index as one of the underlying indicators to create local element index of cities. They argue that patent index is direct indicators to measure city technological innovation comprehensively.

Regardless of its frequent uses in studying urban development, the patent as innovation indicators has disadvantages that scholars ought to acknowledge. Firstly, it is undeniable that not all innovative product was signed up as the patent properties (Hall, Jaffe, et al., 2005). Thus, there might be some innovative activities which are not available in the patent database. Another problematic issue is that not all patented products are commercialized. Thus, it's contribution to urban economic development become questionable.

Another disputed point is due to its skewed economic value (Griliches, 1998, Hall, Jaffe, et al., 2005, Carlino and Kerr, 2015) in the market force. An invention might have a merely high monetary value when it successfully implemented and accepted in the market, while other might be insignificant. Thus it might only reach the patent list without being forwarded to the market.

The last approach to measure innovation is due to the new product announcements on trade, engineering, and technical publication. This approach seems close to perfect in measuring innovation since it accounts for the final product generation and the product's economic value. This approach solves the issue of commercialization that arises on the investment of innovation approach and the patent-related approach. However, the data on new product announcement in local level is very limited. Another thing is that the new product announcement data may lead to bias results since the publication editors might announce the only influential new product. Thus, it may not represent the overall innovation.

Next discussion will investigate the correlation between innovation and urban competitiveness. Soo (2018) argued that the city is the focal point of innovation based-on his discovery of a positive correlation between GDP and density and patent application across OECD cities. Meanwhile, Andersson, Quingley, et al. (2009) conclude that worker productivity positively relates to patenting activities and investment in university research in Swedish regions. Another study from Bettencourt, Lobo, et al. (2007) shows the positive relation between population and R&D activities with creative industry employment. Parisi, Schiantarelli, et al. (2006) reveal that R&D give significant impact on labor productivity, while, Hall, Jaffe, et al. (2005) proved that 'quality' citation of the patent could boost firm monetary value in the market. All in all, innovation shows a persistent relationship with urban economic development.

Table 2 Literature analysis on Innovation

Author, year	Dependent variable(s)	Independent variables	Results
Soo, (2018)	GDP and Density	Patenting activities	the positive relationship between variables
Andersson, Quingley, et al. (2009)	Worker productivity and awarded-patent	decentralization of university-based research	the positive relationship between variables
Bettencourt, Lobo, et al. (2007)	R&D employment	creative industry employment and population	the positive relationship between variables
	Inventors	Population	Superlinear effect relationship
Parisi, Schiantarelli, et al. (2006)	labor productivity	R&D activities	R&D give significant impact on labor productivity
Hall, Jaffe, et al. (2005)	market value	patent citation	one value of patent citation on R&D activities boost 3% of market value

2.1.2 The geography of innovation

The geography of innovation seems to be the most popular research theme to study. It is a common belief that innovation is concentrated in particular geographical space and its distribution is highly skewed (Florida, 2002). Buzard, Carlino, et al. (2016) proofed that R&D labs location in indeed geographically clustered (see Figure 2.) by utilizing the R&D location data and patent data from USPTO. The maps provide evidence for the uneven distribution of inventive activities in two core clusters: Northeast corridor and California. Marshall (1898) argue that the short proximity between aggregate innovation class creates a dynamic interaction that leads to knowledge transfer. "*The geographic concentration of people and jobs in cities facilitates the spread of tacit knowledge"* (Carlino and Kerr, 2015, p.371). Thus, many scholars exercise the magnitude of innovation to provide the further evidence of innovation concentration in particular geographical scale (Murata, Nakajima, et al., 2014, Buzard, Carlino, et al., 2016, Kerr and Kominers, 2015).

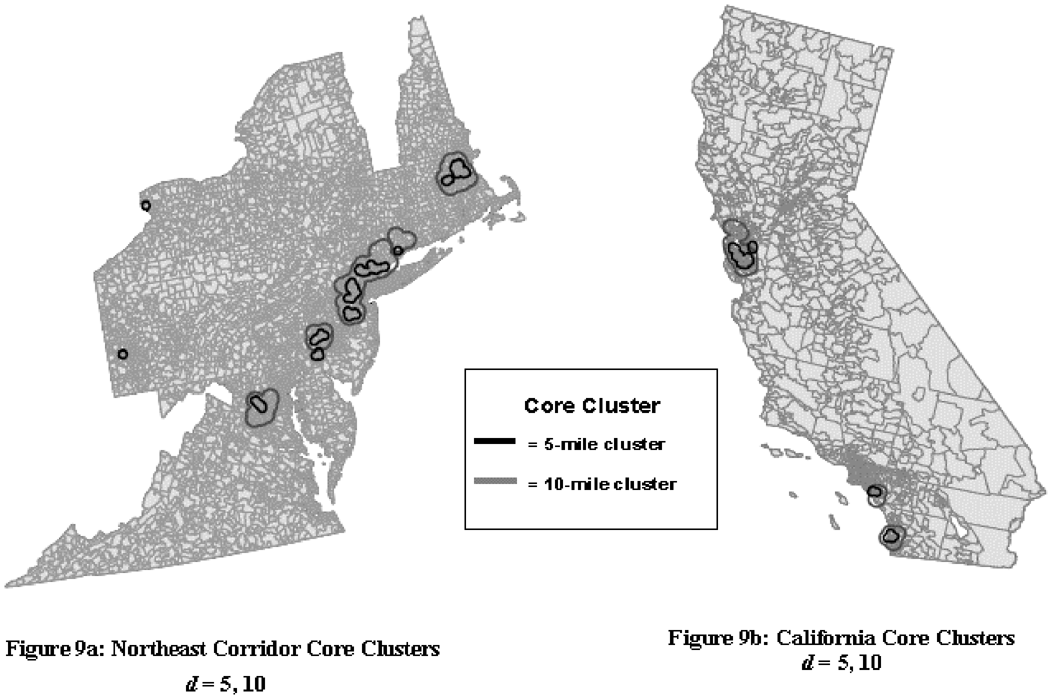


Figure 2 Innovation Clusters

Source: (Buzard, Carlino, et al., 2016)

Ernst (2009, p.35) classify four types of innovation concentration according to the new global innovation network (GIN):

1. *Global center of excellence* is characterized by the integration between basic-applied researches capabilities and advanced innovation infrastructure. For examples are regions such as the US, Japan, and the EU.
2. *Advance location* is benefited from the production activities of the MNCs. This type of regions has the excellent local skilled-labor, good quality of product and huge export markets, for example, Israel, Ireland Taiwan and Korea.
3. *Catching-up location* is the result of second offshore wave which includes China as the global factories and India is the global services. This type of region is highly integrated to the global production network which gives the opportunity for this region to accelerate their learning process and capability improvement, for examples Beijing in China and Delhi in India.
4. *"New Frontier" location* is the alternative locations when the caching up location no longer economically beneficial. The "new frontier" location is attractive for the firms due to the availability of cheaper yet highly motivated human resources. The regions that considered as the "New Frontier" location are Xian, Chengdu or Changqing in China and Ahmadabad or Pune in India.

Moreover, scholars try to explain for what reasons innovation clusters are developed. Firstly, Ellison, Glaeser, et al. (2010) argues that the agglomeration of industry occurs to reduce the transportation cost of people, product, and ideas. Also, Saxenian (1990) also finds that in Silicon Valley firms are intentionally agglomerated to learn effectively with each other to innovate. Therefore, the proximity between R&D facility to the market and the proximity among the creative workers are important to accelerate the information sharing within the creative groups. Secondly, Xibao (2007) suggest that the structural transformation of innovation activities cause innovation clustering. He argues that there is a shift from "technological intensive innovation" by universities to "marginal innovation" domination by firms in China (Xibao, 2007, p.355). Hence, the innovation activities emerge where the agglomeration of firms is located. Other reason, is the path dependency of the region. The quality of former innovation in a region may give benefit to the future innovation activities (Breschi, 2000, Feldman and Florida, 1994). Lastly, Bresnahan, Gambardella, et al. (2001) suggest that the clustering of innovation may occur due to the enforcement of the public policy, business strategy, and public institution. However, he does not suggest the typical top-down policy or strategy. He instead refers to a policy that rules the enabling condition for firms to innovate. In conclusion geographical proximity, firms' innovation activities and government intervention contribute to the emergence of a regional innovation cluster.

Furthermore, Florida (2002) devotes more his attention to the communities in the city and regional economic growth. He explores the relationship between the concentration of creative class, human capital, and the high-technology industry. He believes that the talent concentration and openness to diversity and creativity do not give direct impact on advanced technology establishments. Instead, he argues that talent concentration and openness of a city indicate that a city has a favorable environment for hosting innovative activities. He finds that New York and Los Angeles have the highest number of bohemians (excess 100,000) as well as both cities are the top 3 of bohemians' number/1000 people. Another finding is that the boho index is strongly correlated with talented human capital, while closely related to high-tech cluster formation. Florida (2002) introduced a study that explores the correlation between talent, diversity and regional income. The main unexpected finding of this relationship is that the diversity of an area turns out to be the most important factors for locational preferences of talents, while the companies are attracted to where the concentration of talent is. In this case, both diversity and talent give direct and indirect influence to regional economic growth.

Lastly, I will discuss what factors influence the improvement of innovation in the urban area. Audretsch and Feldman (2004) suggest the idea of geography as the platform where knowledge spill over transfer into innovation. They argue that a heterogeneous relationship exists between knowledge sharing and the formation of agglomeration. Concentrated human capital becomes the key factors in enabling information flows within an area. Another leading factor is the investment on R&D. Bettencourt, Lobo, et al., (2007) and Andersson, Quigley, et al. (2009) argue that investment in university R&D activities determines the magnitude of creative and innovative activities. Later on, Soo (2018) who perceive innovation in a more global context, believes that international competition drives innovation in the city to the extent that innovation creates welfare. In conclusion, cities are the main hub to facilitate both economic and inventive activities.

2.1.3 Air transportation network and urban development

In the present context, cities are integrated entities that spontaneously organized themselves into particular networks (Taylor, 2014a). The city network occurs through city competition and coordination relationship as an effort to boost their economic growth. In term of non-physical connectivity, cities form a network through resource transactional activities (Burger and Meijers, 2012) and economic power distribution (Alderson and Beckfield, 2004, Wall and Van der Knaap, 2011). For instance, Alderson and Beckfield (2004) investigated the city network analysis through the organizational relational linkage of multinational-company data. They observed the centrality of cities in global network system through the attractiveness in hosting multinational company offices. Indeed, the physical connectivity of a city remains to become an important aspect that companies take into account while placing their offices. Hard infrastructure is considered pivotal to sustain the dynamic flow of humans and goods which are the key aspects of maintaining the drift of the economic wheel. Transportation infrastructure facilitates face-to-face meeting which is the most effective way to transfer knowledge and innovation and ensure the movement of goods which is the core trading activity. Thus, effective transportation connectivity entails good transportation network management to bring desired economic value into a city. In this study, the main discussion will be focusing on the air transportation network, considering it is the most effective modes to connect with long-distance locations physically.

Button and Yuan (2013, p.331) describes the benefits of airport infrastructure for economic development into four categories:

- *Primary effects.* A region gains the short-term direct benefit from airport construction. For example, it generates income and employment in the airport construction site.
- *Secondary effects.* The local economy of the region gets advantages from the operation of the airport. For instance, it provides job opportunity in airport-related industries. Meanwhile, it also becomes tax resources for government.
- *Tertiary effects.* An airport stimulates the productivity growth of firms or individual by providing air services to support their mobility.
- *Perpetuity effects.* An airport could change the entire economic structure or function of a region.

However, this thesis research will only focus on the tertiary effects of air infrastructure to cities competitiveness.

Another study reveals that the status of an airport indicates its city economic development. Schaafsma (2003) compared airport position in air network system with a population size of a city to investigate whether air connection gives implication to economic development. In his discussion, he implied airport as a city itself, and he categorizes cities based on its airport status: cities with hub airports, cities as big Origin-Destination-world (OD-world cities), and cities with regional airports. Cities with hub airport benefited with more destination and frequencies from extensive airline network. It characterized through its central position in the market and its reliability to perform its schedule. Frankfurt, Singapore, and San Fransisco are good examples of hub-airport in Europe, south-east Asia and North America considering it is the main interchange connection of Star Alliance member airlines. The big OD-world cities, otherwise, do not provide such a huge interchange. Instead, big OD-world cities usually are the main point-to-point airport with huge numbers of international destination (sometimes with air transportation congestion), for example, London, Tokyo, and New York. In contrast, the last group of cities with the regional airports does not have so many international connections despite its big size, such as Detroit in North-America and Vienna in Europe. In conclusion, Schaafsma (2003) argues that cities with a global hub airport tend to become the center of economic activities.

Many studies have proved the strong relationship between air transportation and urban development. Goetz (1992) found that air passenger flows positively affect population size and employment rate. Button, Lall, et al. (1999) revealed that air transportation could be an attractive factor for the advanced technology industry. Another study was conducted to examine the implication of air freight infrastructure to economic growth in the periphery area (Button and Yuan, 2013). The result indicates a weak positive causal relationship between air freight and employment and income in suburb areas. By observing the placement of SMEs subsidiaries in cities, Bannò and Redondi (2014) prove that the operation of a new route in the peripheral region could boost the number of FDI on SMEs. Later on, Kalayci and Yanginar (2016) reveal a long-term relationship between economic growth, FDI and air transport; then the recent study found that airport delay could impact the employment performance in the tradeable service industry (Lakew and Bilotkach, 2018). In conclusion, the causal relationship between air transportation and cities development does exist.

Nevertheless, the debate on causality direction between these aspects keeps ongoing, as the endogeneity issues appear to become problematic in the correlation test. Debbage (1999) argued that there is a two-way relationship between those aspects; the economic growth initiates the construction of the airport so that the airport can expand. Thus it allows bigger capacity and networks. Second, the new development of the airport can lead to better economic growth. However, by running Granger non-causality analysis, Mukkala and Tervo (2013) argue that the different direction occurs depend on the position of an airport. Their study found different causal direction between cities in the core regions and peripheral regions. In core regions, the economic growth tends to induce the improvement of air infrastructure. In contrary, the improvement of the airport in periphery regions allows more human flows and goods transfer which foster economic activities. Another approach to seeking the endogeneity issue of these two aspects was conducted by Lakew and Bilotkach (2018) through two-stage least square (2SLS) analysis who suggest that airport delay directly/indirectly influence employment. Utilizing weather as instrument variables, they found that air transportation congestion impact to the employment performance is diverse. The air transportation delay has a negative effect on employment in tradable-services industry, while it cause no significant implication on the manufacturing industry.

Based on the literature review on correlation analysis between air transportation infrastructure and economic development, I try to identify the research gap from the previous literature. The findings show that none of the causal regression analysis incorporates air-network into their analytical model. I realize that it is not sufficient to indicate the airport position only through its size (passenger numbers). Thus, the investigation on how a region connects to another region through air transportation might be ample to justify the region's position in the network. The investigation needs to consider the number of the route served, to what kind of cities is the airport connected, and how often they provide the connection (frequencies). Though, the number of passenger and accessibility remain the important aspect to be included in the research as they reflect the potential value of the economy and the easiness to reach the targeted market. Hence, this thesis research will try to fill the gap by compounding the air network aspect of the proposed conceptual framework.

Table 3 Literature analysis on air transportation and urban development

Author, year	City variable(s)	Air transportation variable(s)	Methodology	Result
Lakew and Bilotkach (2018)	Employment	Airline delay (using Instrument Variable: weather)	two-stage least square (2SLS)	Airport delay negatively impact the employment performance in the tradable service industry
Kalayci and Yanginlar (2016)	Economic growth, FDI	Air passenger	Multilinear regression, Johannes cointegration,	The long-term positive relationship between economic growth, FDI and air transportation
Bannò and Redondi (2014)	FDI on SMEs	Operational of new routes	Comparison	New route in peripheral region boost the number of FDI on SME's
Mukkala and Tervo (2013)	Regional economic growth	Accessibility and air passenger	Granger non-causality analysis	In core regions: regional growth causes airport activity In the peripheral region: airport can boost the local economy
Schaafsma (2003)	City size	Airport network status	Comparison	Cities with a global hub airport tend to become the center of economic activities
Button and Lall, (1999)	Employment	Air passenger	Case study and OLS	Air-traffic is an attractive factor for advanced technology

2.1.4 Social network concept in the perspective of air transportation networks

For this study purposes, I use the social network approach to determine the region status in the air transportation system. First, I will explain the basic idea of centrality measurement according to the social network concept, which further detail I will explain in the methodology section. Then, I will discuss some previous studies in air transportation system that employ social network approach.

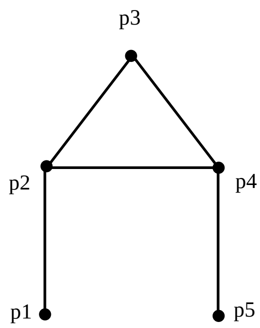


Figure 3 A graph of five points and five edges

Source: (Freeman 1978, p. 218)

Freeman (1978) suggest three centrality measures to describe the status of a node in the network system; in this study case, a node will represent a region, while an edge will represent a connection. First of all, degree centrality measures the magnitude of direct connections that a node has in a network system. For example in the above Figure 3 modes *p1* has only one direct connections with one member in the network which is *p2*. Thus, p1 has one-degree centrality. Whereas, *p2* has three degree centrality from its connection with *p1*, *p3* and *p4*. Furthermore, in this thesis study, a direct connection is translated as a direct and non-stop flight between a pair of OD regions. Such connection indicates the interactions between regions in the form of human flows that may happen for some reasons such as for work or businesses, educational purposes, family visit, or even for a recreational trip (the degree centrality equation will be provided in the chapter. 3.5.2).

Despite what kind of interaction behind the direct connections, degree centrality forms particular characteristics of a node which could be translated into several meanings. First, the degree centrality signifies the locational benefits of a node by being connected to other nodes (Guimera, Mossa, et al., 2005, Wang, Mo, et al., 2011). In term of regional innovation, the degree centrality could be associated with the abundant access to the source of information and technology. Second, Guimera, Mossa, et al., (2005) also suggest that degree centrality represent the attractiveness of a node. A region is attractive when there are particular airlines which serve direct flights to that region. The reason might because there is such concentration of market, human capital, jobs or even leisure attractions in a region that make people need or want to travel to the particular region. Moreover, the degree centrality also indicates the importance of a region for political motives. For instance, there are many flights accommodate New York route since New York host most of the firm's headquarters, or Washington DC since the White House located in this region.

The second measurement is betweenness centrality. It measures the number of occasions when a node acts as the connecting bridge in the shortest path of a pair of nodes. Based on Figure 3, there are three occasions that *p2* acts as the bridge in the shortest route of other nodes connection; *p2* connects *p1* and *p3*, *p1* and *p4* and *p1* and *p5*. On the other hand, *p3* does not have any betweenness centrality since p3 is not included in any nodes shortest path. Guimerá, Mossa, et al. (2003) argues that nodes with a high degree of betweenness play an important role as the network connectors. If a disruption occurs in *p2*, for example, then *p1* become isolated. Thus, nodes such as *p2* become important, especially for *p1*, to keep the network completely intact (the betweenness centrality equation will be provided in the chapter. 3.5.2).

The betweenness centrality of a node is commonly associated with hub function in the air transportation network system. For example, Guimerá, Mossa, et al. (2005) find that the city of Paris, London and Frankfurt are European cities with high betweenness centrality in the world air transportation system. Schaafsma (2003) categorize Paris and Frankfurt as the hub-cities since both of them host many connecting flights to other cities in Europe. However, Schaafsma (2003) argues that London does not ideally function as a hub since the London airports (Heathrow and Gatwick) do not meet the requirements of a hub; one of which is that London is hard to provide the reliable flight schedule.

Lastly, closeness centrality principally aims to measure the network distance of a node to reach the other nodes in the network. One network distance is equal to one direct connection or one edge in Figure 3, and the maximum closeness between two nodes is one. According to Figure 3, for example, the distance from *p2* to *p4* is one, while the distance from *p2* to *p5* is two. Moreover, the closeness centrality calculates the amount of possible closest distance that a node can have to reach other nodes. (the closeness centrality equation will be provided in the chapter. 3.5.2).

The closeness centrality of a region can be associated with the accessibility (Wang, Mo, et al., 2011) and the efficiency in spreading the information (Meligy, Ibrahem, et al., 2014). In term of accessibility, Wang, Mo, et al. (2011) argue that closeness centrality is the extent of the region being reachable by other regions. In term of the air transportation system, the network distance can be translated as the number of flights taken to get to a region. The less number of flights that it takes to reach a region, the closest this region to the other regions in the network. For example, it only takes one flight to get to Paris from Singapore, New York or London. However, it may take two or several flights to get to Lyon from those three cities since there is no airlines serve direct flight for these routes. In this case, Paris is more connected compared to Lyon. Secondly, in the context of communication, Meligy, Ibrahem, et al. (2014) argues that closeness centrality determines the productivity of a node in delivering information in the network system. However, I argue that closeness centrality may also represent the effectiveness of a region in getting knowledge from other regions which is useful to induce their activities in innovating. In fact that internet technology is dominating the flow of information, further research needs to be done on what kind of knowledge that can be brought only by air passengers and how air passenger brings such knowledge. Thus, this discussion becomes interesting to be developed in the future research.

Moreover, some scholars applying the social network approach in examining the air transportation system for particular purposes. First, Irwin and Kasarda (1991) observe the effect of airline structure revolution on employment. Their finding is Secondly, Guimerá, Mossa, et al. (2005) analyses the structure of the global air transportation network by calculating the degree, betweenness and closeness centrality of each city in the network. He finds that the cities with high degree centrality and closeness centrality does not necessarily have high betweenness centrality due to political and geographical reason. Next, Cheung and Gunes (2012) employ the degree centrality, betweenness centrality and other network features of airline connection to understand the characteristic of US air transportation system. They found that the US air transportation network performs the small-world characteristic, where the smaller airports more likely to cluster with the bigger airports. Thirdly, Wang, Mo, et al. (2011) exercise the implications of centrality measurement on socioeconomics. As a result, they found that degree, betweenness and closeness centrality are highly correlated with air passenger volume, population, and GDP. Thus, by analyzing the air-network centrality of a region, the researcher could get insights on understanding urban development.

2.1.5 Region's centrality status in air network and its' regional innovation

To draw the red thread from the previous discussion, I will discuss further the potential linkages between air transportation network and regional innovation. At first, I argue that the centrality of regions in air transportation network affecting its regional innovation level due to the locational reason of innovation facilities. Furthermore, the agglomeration of R&D facilities and creative people bolster the influence of air transportation on regional innovation. Lastly, I will discuss the interplay between air passenger flow and employment in the creative sector to address the endogeneity issue.

Air transportation infrastructure is the pivotal factor in determining R&D facility locations. Cornet and Rensman (2001) suggest that proximity and the quality of international airport is a crucial factor to locate the green-field R&D facilities. In term of proximity, traveling time and type of modes to the airports are the main concern. The reliable airport transportation services to the R&D locations such as high-speed train (HTS) is quite beneficial to support researchers' high mobility and to lessen the environmental cost of transportation (Givoni and Banister, 2006). In term of quality, the number of air connectivity and the reliability of the flight schedule are also essential to support the R&D firms' international activities. The region with the international airport is usually more attractive for the R&D firms than a region without an international airport. Besides, the performance of the airlines may also be important for R&D activities. Airport with high congestion causes airlines delay which negatively affects employment productivity (Lakew and Bilotkach, 2018). In result, the air traffic congestion may disrupt the R&D firm productivity.

Agglomeration of R&D firms (jobs) and creative class (researchers) plays an important role in promoting the influence of air connectivity on innovation growth. First, the R&D firms tend to clusters around the metropolitan area which is characterized by high air-connectivity (Malecki, 1980). The R&D firm concentration in a region indicates that a region has a conducive environment to support R&D activities. Thus, it attracts more R&D firms to locate in that region. Also, air connectivity becomes the main need for the creative class due to the needs of high mobility in their international career activities. The more connected a region is, the more attractive that region for creative class; thus the more concentrated the creative class in that region.

However, endogeneity issue remains the though discussion in revealing the causal mechanism between innovation growth and air connectivity growth of a region. To address the endogeneity issue, Neal (2012) observe the relation between air passenger flow and employment in the creative sector. He suggests that the hypothesis applies in particular economic condition (as shown in Figure 4). Firstly, the *flow generation hypothesis* occurs when the economy in a region is declining. In this condition, Neal (2012) argues that the high number of creative class triggers the high number of air passenger flow. The rise of air passenger flow may occur due to the mobility of these following three type of passengers who may relate to the creative sector. First is the tourist and the consumer who enjoy the creative products. Second is the other creative class who do their businesses because they think the people in this region are welcome to their ideas and works. Last is the other stakeholders who work to support the creative economy. Therefore, according to the first hypothesis, the improvement in innovation employment result in the air passenger growth.

On the other hand, the *structural advantage hypothesis* occurs when there is economic growth (Neal, 2012). In the positive economic situation, Neal (2012) explain that cities with high air passenger flow create two following environments which are favorable for creative economy activities. First, the high passenger flow brings larger potential creative consumer, greater creative employment, and latest information that useful for creative work. Second, air passenger flows help to spread of information that a region has a "*creative economy of scale*" (Neal, 2012, p. 2696). Thus, other creative classes from other regions are attracted to the region with high air passenger flow because it indicates the concentrations of resources for the creative economy. Therefore, the reciprocal between air connectivity and regional innovation may apply.

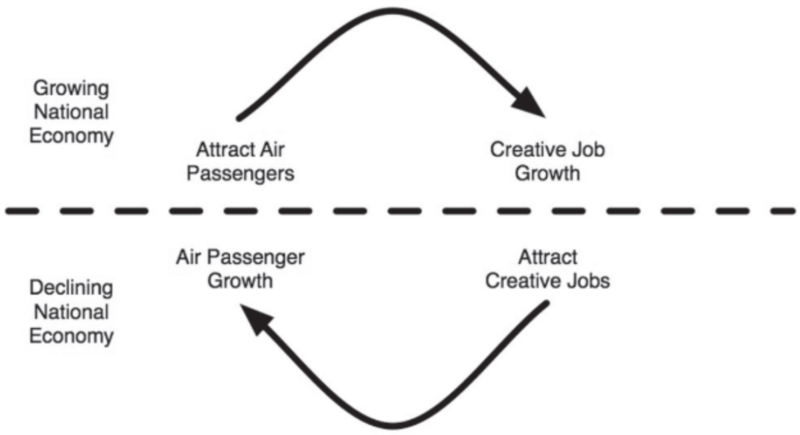


Figure 4 Schematic of an evolving urban development strategy

Source: (Neal 2012, p. 2706)

2.2 Conceptual Framework

For this study, I illustrate the link between regional innovation and air passenger network status of a region through the following conceptual framework:

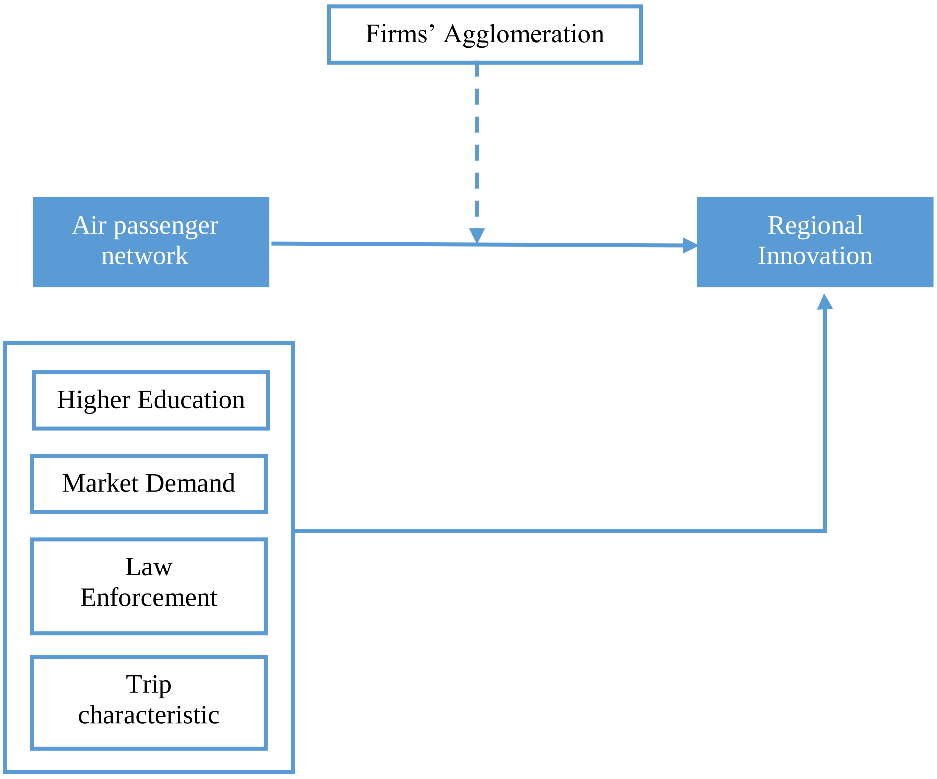


Figure 5 Conceptual Framework

In the conceptual framework, I will explain the regional competitiveness innovation by elaborating the air transportation network connectedness of a region into the context of economic geography. According to the literature review, innovation plays an important role as one of the underlying variables for the local element in competitiveness concept of The Global Urban Competitiveness Report 2017 (Ni, Kamiya, et al., 2017). Meanwhile, the air transportation network enables human capital interactions that lead to knowledge transfer which shape the competitiveness of a region.

In this conceptual framework, I perceive regional innovation from firms' activity on R&D and count of patent grants since they reflect the process of innovation and the output of innovation consecutively. According to Parisi, Schiantarelli, et al. (2006), investment in R&D induces regional productivity. Thus the investment in R&D may lead to social welfare. Meanwhile, Soo (2018) find that patenting activity has the positive effect on GDP and population. Hence, I believe it is important to find out in what circumstances inventive activities develop in an optimum way.

Meanwhile, I will explain the air passenger network status of a region through three centrality measurements according to the social network approach: degree centrality, closeness centrality, and betweenness centrality (Freeman, 1978). Each measurement indicates how well a region connected to the air transportation system in a different manner. The level of regional connectedness in air passenger network distinct region position in the air transportation system. That position is significant as it facilitates important flows of knowledge exchange through physical human capital interaction, and determines the benefit and strength of a region gained from air transportation network.

As shown in Figure 5, I draw a channel from the air passenger network to regional innovation, assuming that air passenger network may positively correlate with the regional innovation. I expect that regional connectedness status in air passenger network can induce human capital flows that lead knowledge spillover that in the end generates innovation. However, I also suspect there will be endogeneity issue in this relationship. The favorable environment which allows the generation of inventions in a city might attract the creative class to interact with each other then influenced the air transportation generation in the airport.

Furthermore, I found many kinds of literature discuss the correlation between agglomeration and knowledge spill over. Thus, I argue that the concentration of economic activities will influence the relationship between air transportation network system and regional innovation. I describe the degree of economic concentration per MSA by measuring region's agglomeration level. In this conceptual framework, I would like to particularly observe the moderating role of agglomeration in facilitating the knowledge transfer occurs in the relationship between air transportation network system and regional innovation.

Additionally, I use four control variables which I believe well explain the regional innovation. The first control variable is higher education level which usually associated with the universities related indicator. The other controls are market demand that measures the opportunity for firms to grow and the law enforcement that influence the business environment. Lastly, I also include the trip characteristic to identify the whether a region is a business destination or leisure destination.

All in all, by knowing the relationship between air transportation network system and regional innovation, the policymaker can have an insight on how should they will utilize air transportation infrastructure as the tools to shape the regional competitiveness.

Chapter 3: Research Design and Methods

3.1 Revised research question

To provide better focus, I reformulate the main research question in this study as follows: to what extent the regional status in air transportation network affect the regional innovation and how agglomeration influence its relation?

Hence, consecutive sub-questions are needed to answer the main question:

1. How is the distribution of the regional innovation in the United States? Which MSAs host the highest number of R&D activities? Which MSAs has the highest number of patent grants?
2. What is the region's status in the air transportation network in the United States? Which MSAs have highest measurements of centrality in air transportation network?
3. What is the correlation between regional innovation and region position in air transportation network? To what extent the connectedness of an MSA will influence its regional innovation?
4. How does the firms' agglomeration level in a region influence the relation between air transportation and regional innovation?

3.2 Operationalization variables and indicators

In this research, I employ a proxy of inventive activities in a region as the dependent variable. The unit measurement of a region will be in the Metropolitan Statistical Area (MSA) level. Furthermore, as the indicator to measure regional innovation level, I use the number of R&D establishments since it represents the process of inventive activities. The R&D establishment data are according to the number of establishment in industry scientific research and development services (NAICS code 5417). Meanwhile, another indicator to measure regional innovation is the number of patent grants by origin since it represents the output of regional innovation. To be acknowledged, the location in the origin patent data is according to the first inventor residence address which means that the inventive activities are not necessarily occurring in the same place. Nevertheless, I believe that patent origin data could still explain the characteristics of an environment which is favorable for the creative class to create the output of innovation.

The independent variable in this research is the region's status in the air transportation network system. I will measure the region's status in air transportation network through three network measurements according to the concept of centrality in social network introduced by Freeman (1978). The first measurement is degree centrality which is determined by the number of direct connection of a region to other regions in the network system. The degree centrality reflects the attractiveness of a region in the air transportation network system which further frequently associated with diffusion and congestion (Guimera, Mossa, et al., 2005). The second measurement is betweenness centrality which refers to the frequency of a region acts as the connecting point of two other regions in their shortest route. Betweenness centrality determines the importance of a region in the network system since it represents the hub function in the network system (Guimera, Mossa, et al., 2005). Lastly, closeness centrality refers to the region's distance to the network system. The closeness centrality measurement is obtained through the inversion of the average distance from a region to the other rest regions in the network. It is frequently associated with the effectiveness of a node in spreading the information to the rest of the network members. Guimera, Mossa, et al. (2005) associates the closeness centrality to determine the path length of connection in the global air transportation system.

All in all, among these network measurements, I believe that degree centrality can be one of good measurement to indicate whether an MSA is more attractive to creative class in advance technology and inventors. Thus, degree centrality is more sensible to explain regional innovation compare to betweenness centrality and closeness centrality. Additionally, I will indicate the regions which are holiday destination by using a dummy variable to differentiate whether the trip characteristic of two regions is more for leisure purposes or not.

Moreover, to have a closer investigation, I assume that the agglomeration level of the region influences the relationship between regional innovation and region's status in air transportation network. I consider agglomeration level to illustrate the concentration of human capital, since according to the literature such concentration positively correlates with innovation (Florida, 2002, Audretsch and Feldman, 2004, Bettencourt, Lobo, et al., 2007, Soo, 2018) as well as air transportation (Goetz, 1992, Schaafsma, 2003). In addition, I utilize human capital, market demand, and law enforcement variables to control the circumstances in the business environment. Lastly, I use the storm events as the instrumental variable to further investigate the endogeneity issue in the model.

Table 4 Research Operationalization

(y) Regional innovation (by Metropolitan Statistical Area)			
Variable	**Definition**	**Indicators**	**Unit of measurements**
Regional innovation (process)	The degree to which inventive activities take place in a region	R&D establishment	Number of industry establishment of scientific research and development services (NAICS code 5417)
Regional innovation (output)	The degree to which innovation output generated in a region	Patent grants	Number of patent grants
(x) Air transportation network (by Metropolitan Statistical Area)			
Variable	**Definition**	**Indicators**	**Unit of measurements**
Region's status in air transportation network	The degree to which the connectivity of a region within the air transportation network occurs	Degree centrality	Degree centrality value
		Betweenness centrality	Betweenness centrality value
		Closeness centrality	Closeness centrality value
Agglomeration (by Metropolitan Statistical Area)			
Variable	**Definition**	**Indicators**	**Unit of measurements**
Agglomeration level	The degree of business concentration in a region	Firms agglomeration	Number of firms/ sq miles
Instrumental Variable (by Metropolitan Statistical Area)			
Variable	**Definition**	**Indicators**	**Unit of measurements**
Weather	The degree of natural force influencing regional innovation and air transportation	Storm	The number of storm events

Controls (by Metropolitan Statistical Area)			
Variable	**Definition**	**Indicators**	**Unit of measurements**
Air trip characteristic	The trip characteristic to a region	Top 25 vacation destination	Dummy variable (1 = MSA that has top 25 vacation destination/s; 0= MSA that has no top 25 vacation destination)
Higher education	The human quality with higher education	Top 100 university	Dummy variable (1 = MSA that has top 100 university/s; 0= MSA that has no top 100 university)
Market Demand	The degree of marketability to absorb innovation output	Per capita income	Thousands of dollars
Law enforcement	The degree in law and regulation enforcement	Crime	Crime (event) per 1 million people

3.3 Research Strategy

I will conduct this study through desk research strategy on secondary data resources. I argue that desk research is the best approach for this study considering the utilization of statistical database that has a huge number of observations with various variables. In this case, I will use the existing database that has been collected by the United States Government (more detail about data resources is in the next section). Furthermore, I will particularly do quantitative analysis. Thus I will exclude the qualitative discussion.

In general, I elaborate three types of analysis to answer overall questions. First, I employ descriptive statistical analysis to explain the distribution of innovation. Through graphs and maps, I would like to describe how the concentration of inventive activities disperse in the overall United States area. Secondly, I convey network analysis to define the status of each MSA in air transportation network by generating three network measurement namely degree centrality, betweenness centrality, and closeness centrality. Lastly, I will conduct correlation analysis to reveal the relationship between regional innovation and region's status in air transportation network and how agglomeration influences the relationship. I will carry out the correlation analysis based on the operationalization scheme to covers the substance of the main question. For more details, the data analysis methods section will discuss the specific steps in each analysis.

3.4 Data Collection Methods

Overall, I will utilize secondary data from several websites belong to US Governments. For the descriptive part, I downoad the R&D establishment data from the US Census Bureau web page. The data on R&D establishments is available from the table titled "Professional, Scientific, and Technical Services: Geographic Area Series: Summary Statistics for the U.S., States, Metro Areas, Counties, and Places: 2012". The table refers to the establishment of industry scientific research and development services (NAICS code 5417) that represent the process of inventive activities in MSA level. Next, I obtain the patent origin data from the US Patent and Trademark Office website. The patent origin data is available from the table titled “Patenting In Technology Classes Breakout by Origin, US Metropolitan and Micropolitan Areas.” The patent origin data has the information about the number of patent grants per MSA.

For the network analysis, I will retrieve the United State air transportation data in 2012 from United State Bureau Transportation Statistic online portal. The air transportation database is available from the table namely "Air Carrier Statistics (Form 41 Traffic)- All Carriers". The database contains information regarding carrier details such as origin and destination, passenger number per route and freight per route. The database includes the international market when a flight serves at least one point located in the United States. It excludes the flight information which both points of service are outside United States territory.

Meanwhile, for correlation analysis, I will employ the R&D establishment data and the patent origin data which I utilize in the descriptive analysis as the dependent variable. Whereas, as the explanatory variables, I will employ the value of three network measurements (degree centrality, betweenness centrality, and closeness centrality) of an MSA that previously resulted in network analysis phase. Additionally, to consider the influence of agglomeration on the correlation between dependent and independent variables, I will measure the region's agglomeration level by dividing the number of the firms with the land area of MSA. For this purposes, I will refer to two databases from US Census Bureau; The data of firms number from the table titled "All sector: Geographic Area Series: Economy-Wide Key Statistic: 2012" and the MSA land area data from "Metropolitan Area Census Data: Population & Housing Density 2000". As the controls for the correlation analysis, I create a dummy variable that explains trip characteristic based on the latest rating of traveler choice by Tripadvisor, and another dummy variable for higher education based on top 100 universities of Forbes.

Additionally, I also retrieve the US weather data from the website of National Centres for Environmental Information by the US National Oceanic and Atmospheric Administration (NOAA). To be acknowledged, the data on storm events are provided in the state level. However, I assume that each MSA within a state gets the same number of the storm events. As for the MSA located in two or three state boundary, I use the average number of storm events. Lastly as the control variable as well, I also obtain the income per capita data from the US Census Bureau and the number of crime from Uniform Crime Reporting of FBI.

Table 5 Data Resources

Variable	Indicators	Table Name	Sources
Regional innovation (process)	R&D establishment	Professional, Scientific, and Technical Services: Geographic Area Series: Summary Statistics for the U.S., States, Metro Areas, Counties, and Places: 2012	US Census Bureau
Regional innovation (output)	Patent grants	“Patenting In Technology Classes Breakout by Origin, US Metropolitan and Micropolitan Areas”	US Patent and Trademark Office
Variable	**Indicators**	**Table Name**	**Sources**
Region’s status in air transportation network	Degree centrality	Air Carrier Statistics (Form 41 Traffic)- All Carriers Network status from network analysis phase	United State Bureau Transportation Statistic
	Betweenness centrality		
	Closeness centrality		

Variable	Indicators	Table Name	Sources
Agglomeration level	Firms agglomeration	- All sector: Geographic Area Series: Economy-Wide Key Statistic: 2012 - Metropolitan Area Census Data: Population & Housing Density 2000	US Census Bureau
Variable	**Indicators**	**Table Name**	**Sources**
Weather	Storm	Storm Event Database of 2012	National Oceanic and Atmospheric Administration
Variable	**Indicators**	**Table Name**	**Sources**
Air trip characteristic	The trip characteristic to a region	Top 25 destinations-United States	Tripadvisor
Higher education	Top 100 university	Forbes Top 100 university	Forbes
Market Demand	Per capita income	CA1 Personal Income Summary: Personal Income, Population, Per Capita Personal Income	Bureau of Economic Analysis
Law enforcement	Crime	Crime in the United States by Metropolitan Statistical Area, 2012	Uniform Crime Reporting-FBI

3.5 Data Analysis Methods

This research involves three analysis methods that are descriptive analysis, network analysis, and quantitative analysis. In the descriptive analysis, I explain the distribution of regional innovation in the United States to see the locational concentration of RnD establishments and patents by using ArcGIS pro. Then, in the network analysis, I define the MSA's status using the social network approach by measuring the centrality of MSAs in the air transportation network. I will use the result of network analysis as the independent variable in the regression model. Lastly, in the quantitative correlation analysis, I regress the correlation between regional innovation and MSA's status in the air transportation through data analysis and statistical software namely STATA.

3.5.1. Descriptive Statistic Method

In the descriptive analysis, I will project the regional innovation data in the United States maps to show the distribution of R&D and patent concentration. Besides, I also generate a map that shows the agglomeration level in each MSA to provide a general reference for the business environment in the United States. Afterward, I compare the R&D and patent distribution with agglomeration level maps to get the initial idea about the association of agglomeration to regional innovation. The result will indicate MSAs with the highest and lowest concentration of innovation and what is the agglomeration level of that MSAs.

3.5.2. Network Analysis Method

In the network analysis, I aim to define the region's status in air transportation network through three centrality measurements introduced by Freeman (1978): degree centrality, betweenness centrality, and closeness centrality. I will conduct the network analysis in STATA software to calculate the centrality magnitude of each MSA. I use the number of the passenger to weight the air transportation connection since the air passenger is commonly associated with the concentration of population and employment (Goetz, 1992) that eventually attracts creative class and inventors (Florida, 2002).

The explanation of each network measurement according to Freeman (1978) in a nutshell is as follows:

1. The degree of centrality is the proportion of the total number of direct connection a node has to other nodes compare to the total number of nodes that available in the network. The equation below explains the calculation of the degree centrality index (Freeman, 1978, p.220) :

$$C'_D(p_k) = \frac{\sum_{i=1}^{n} a(pi, pk)}{n-1}$$

$a(pi, pk)$	=	One if and only if node (p_i) and (p_k) are directly connected
		0 if otherwise
n	=	number of members (nodes) in the networks

2. The betweenness centrality is the ration of the frequency of a node stands between other pairs of nodes in their ‘shortest path’ divide by the maximum frequency of a node stands in ‘all path’ of other pairs of nodes. The betweenness equation as follows (Freeman, 1978, p.224):

$$C'_B(p_k) = \frac{2C_B(p_k)}{n^2 - 3n + 2}$$

$C_B(p_k)$	=	the sum of overall frequency of node (p_k) lays between two nodes in the shortest route
$\frac{n^2 - 3n + 2}{2}$	=	The maximum betweenness value can be achieved by $C_B(p_k)$ in the network

3. The closeness centrality is the inverse of the average distance of a node to reach other nodes in the network. Freeman (1978) re-introduce the closeness calculation by Beauchamp (1965)

$$C'_C(p_k) = \frac{n-1}{\sum_{i=1}^{n} d(pi, pk)}$$

$d(pi, pk)$	=	The sum of the distance from node (p_k) to other members (nodes) in the network
n	=	number of members (nodes) in the networks

3.5.3. Inferential Analysis Method

In the inferential analysis, I carry on the correlation analysis also in STATA software to reveal the relationship between regional innovation and MSA's status in the air transportation network. The correlation analysis aims to test the hypothesis whether there is a strong relation between inventive activities and the region's status in the air-transportation network and whether the agglomeration level influences their relations.

The correlation analysis involves two phases: the preparation phase and the analysis phase. In the preparation phase, I conduct the data cleaning process by following several steps. First, I merge all of the relevant data into a cross-sectional database. Then, I carry on the descriptive statistic to describe distribution tendency and spread of the row data. Afterward, the multicollinearity test is necessary to prevent collinearity among the independent variables. If STATA detects this multicollinearity issue, then one of the similar indicators should be dropped or substituted by other variables. I also apply the robust command in the regression to anticipate the heteroskedastic condition of the residuals variance. Lastly, I run the Kernel density estimation to test the normal distribution of residuals variance. The results of multicollinearity test and Kernel density test will be provided in the appendix.

For initial information, the following Table 6 contains a general description of the data used in this study, except for air transportation data. The descriptive statistic for air transportation variable will be provided from the result of network analysis.

VARIABLES	min	max	mean	median	variance
(log) R&D	0	6.935	2.549	2.398	2.368
(log) patent	0	9.352	4.029	3.932	3.45
Agglomeration (number of firms per sqmiles)	0.166	58.738	5.55	3.579	49.814
MSA with top 25 vacation destination	0	1	0.049	0.047	-
MSA with top 100 universities	0	1	0.166	0.139	-
Per capita income (thousand dollars)	22.77	107.912	40.311	39.091	77.623
Crime (per 1 million people)	161	5885.7	3256.4	3306.9	1465486

Table 6 Descriptive statistic of the data used in the regression

In the analysis part, I carry on OLS regression to find the correlation between regional innovation as the dependent variable and the MSA's status in air transportation network as the independent variable by taking into count the agglomeration level of an MSA as the moderating variables. The hypothesis of the study will be:

H_0 = there is no relation between regional innovation and air transportation network

H_1 = there is relation between regional innovation and air transportation network

H_2 = there is relation between regional innovation and air transportation network and the moderating effect of agglomeration

In the regression, I include other four control variables which are a vacation destination, human capital, market size, and law enforcement. STATA will generate statistical calculations to explain the significance of each variable and whether it has a positive or negative relation. Hence, the model for each regression would be as shown in the following equations.

Model 1.

$$Innovation_{MSA}^{r\&d} = \alpha_0 + \alpha_1 * Degree\ Centrality_{MSA}^{Passengers} + \alpha_2 * higher\ education + \alpha_3 * market\ demand + \alpha_7 * law\ enforcement + \alpha_8 * vacation\ destinations + \epsilon \qquad (1)$$

Model 2.

$$Innovation_{MSA}^{r\&d} = \alpha_0 + \alpha_1 * Degree\ Centrality_{MSA}^{Passengers} + \alpha_2 * higher\ education + \alpha_3 * market\ demand + \alpha_7 * law\ enforcement + \alpha_8 * vacation\ destinations + \alpha_8 * agglomeration + \epsilon \qquad (2)$$

Model 3.

$$Innovation_{MSA}^{r\&d} = \alpha_0 + \alpha_1 * Degree\ Centrality_{MSA}^{Passengers} + \alpha_2 * higher\ education + \alpha_3 * market\ demand + \alpha_7 * law\ enforcement + \alpha_8 * vacation\ destinations + \alpha_8 * agglomeration + +\alpha_8 * Degree\ Centrality_{MSA}^{Passengers} * agglomeration + \epsilon \qquad (3)$$

Model 4.

$$Innovation_{MSA}^{patent} = \alpha_0 + \alpha_1 * Degree\ Centrality_{MSA}^{Passengers} + \alpha_2 * higher\ education + \alpha_3 * market\ demand + \alpha_7 * law\ enforcement + \alpha_8 * vacation\ destinations + \epsilon \qquad (4)$$

Model 5.

$$Innovation_{MSA}^{patent} = \alpha_0 + \alpha_1 * Degree\ Centrality_{MSA}^{Passengers} + \alpha_2 * higher\ education + \alpha_3 * market\ demand + \ \alpha_7 * law\ enforcement + \alpha_8 * vacation\ destinations + \alpha_8 * agglomeration + \ \epsilon \qquad (5)$$

Model 6.

$$Innovation_{MSA}^{patent} = \alpha_0 + \alpha_1 * Degree\ Centrality_{MSA}^{Passengers} + \alpha_2 * higher\ education + \alpha_3 * market\ demand + \ \alpha_7 * law\ enforcement + \alpha_8 * vacation\ destinations + \alpha_8 * agglomeration + \ + \alpha_8 * Degree\ Centrality_{MSA}^{Passengers} * agglomeration + \ \epsilon \qquad (6)$$

Additionally, I conduct the robustness check to address endogeneity issue and spatial variety regression. Firstly, I employ the 2SLS regression to address the endogeneity issue considering that I use cross-section data for this thesis purposes. To run the 2SLS regression, I utilize weather as the instrument variable as similarly done by Lakew and Bilotkach (2018) in the previous study on air transportation and employment. I use the number of storms events as the indicator to measure the weather condition; I consider that the storm is the exogenous force that influences both the innovation activities and the air transportation. Moreover, I also provide the Durbin and Wu-Hausman test of the 2SLS regression result in the appendix. Secondly, I conduct Geographically Weighted Regression (GWR) to check the spatial variety of regression result. I map the spatial distribution of R-square and provide the comparison between the OLS regression and the GWR result. The detailed result of GWR will be provided in the appendix

3.6 Data Validity

Considering that I conduct desk research strategy for this thesis, I only use the secondary data from the reliable and valid resources. I obtained most of the data from the US government website; indeed, the data are mostly openly available online. Whereas, there are two dummy variables that I make due to the unavailability of data from the US government resources, which are: higher education and trip characteristic. For the higher education dummy variable, I obtain the data of the top 100 US University from the Forbes website. Meanwhile, for the

Chapter 4: Research Findings

4.1 Descriptive Analysis

4.1.1. The spatial distribution of R&D establishments

The R&D database shows the number of the establishment of industry scientific research and development services (NAICS code 5417) by MSA in 2012. The R&D data was recapped according to the firm's report on each of the company's business establishments. It covers 361 MSAs, and the average number of R&D establishment is approximately 45 establishments.

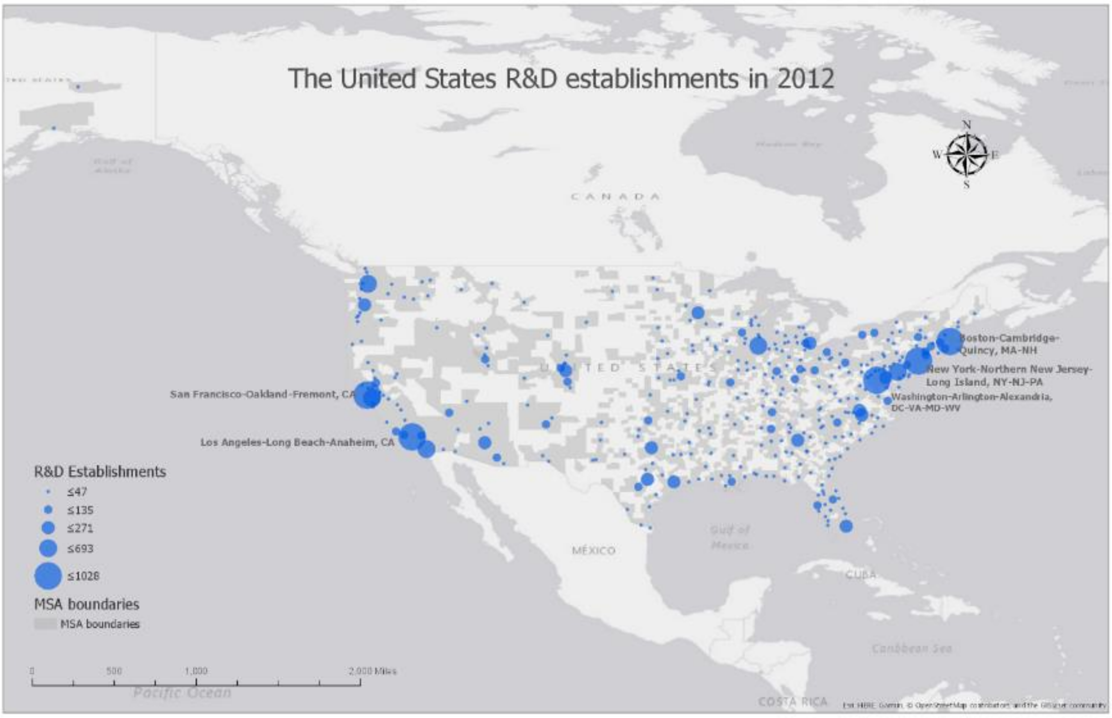


Figure 6 R&D Establishment in the United States in 2012

Source: The US Census Bureau

According to the map in Figure 6, the R&D establishments across the United States is unevenly distributed. The east part of the United States area seems to host more R&D activities than the west part. Additionally, the R&D activities occur more in the cost line of United States. Especially in the west part of the United States where the R&D activities mostly occur in the west coastal areas, and rarely take place in the areas that more to the middle land.

Furthermore, the R&D distribution in the United States is skewed into two areas of concentration: the west coast and the east coast. It is obvious that in both areas, the high number of R&D establishments occur in notable MSAs such as Boston-Cambridge-Quincy (MA-NH), New York-Northern New Jersey-Long Island, (NY-NJ-PA) in east coast; and San Francisco-Oakland-Fremont (CA), Los Angeles-long-beach-Anaheim (CA) in the west coast. Unpredictably, Washington-Arlington-Alexandria (DC-VA-MD-WV) also seems to be the favorable MSA for firms to host R&D activities.

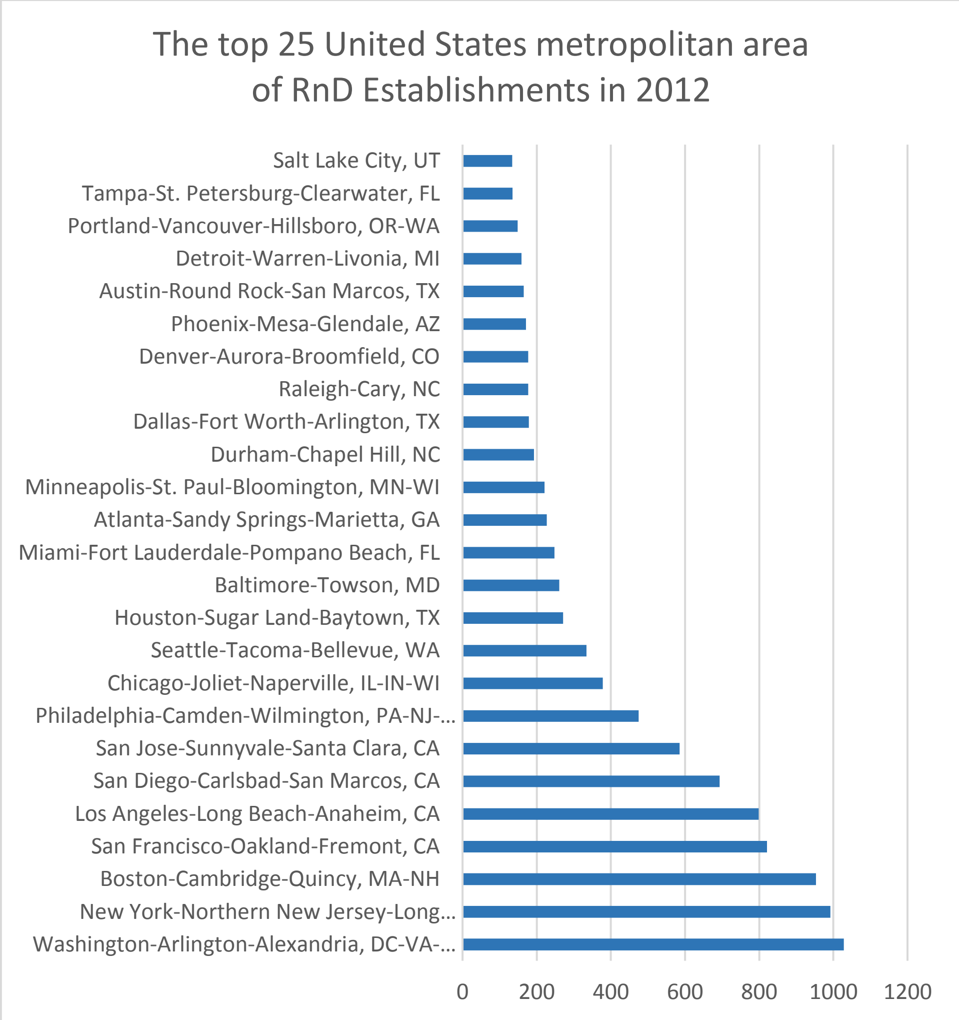


Figure 7 Top 25 the United States metropolitan areas of R&D Establishment in 2012

Source: The US Census Bureau

According to the chart in Figure 7, the coastal MSAs dominate the top 10 positions regarding R&D establishments. Surprisingly Washington-Arlington-Alexandria (DC-VA-MD-WV) host the highest number of R&D by exceeding one thousand establishments in a year. It is followed by New York-Newark-Jersey City (NY-NJ-PA) in the second place and Boston-Cambridge-Quincy (MA-NH) in the third place. These three top MSA comes from the US east coastal area where R&D establishment occurs more according to the previous discussion. Just then, the next four MSA that come from the west area of the United States, specifically from the California States, subsequently appears in the fourth to seventh positions: San Francisco-Oakland-Fremont (CA), Los Angeles-Long Beach-Anaheim (CA), San Diego-Carlsbad-San Marcos (CA) and San Jose-Sunnyvale-Santa Clara (CA). Meanwhile, the top ten rank in the chart only includes one MSA from the middle area of the US which is Chicago-Joliet-Naperville (IL-IN-WI).

Furthermore, the discussion focuses on the agglomeration level in the two areas where R&D establishments are mostly concentrated, namely the east coast area and west coast area. To get a better overview of the relations between agglomerations and regional innovation, I overlap the agglomeration map with R&D maps as seen in the following Figure 8 and 9.

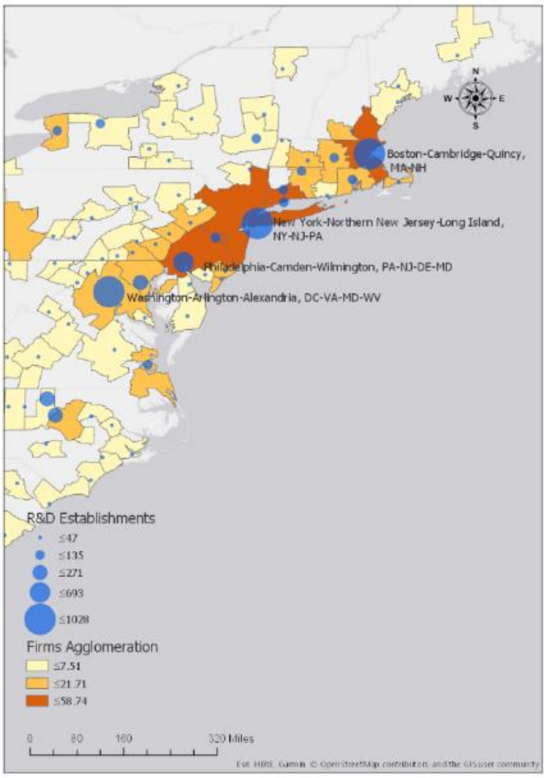


Figure 8 R&D establishment and MSA agglomeration in the US East Coast area in 2012

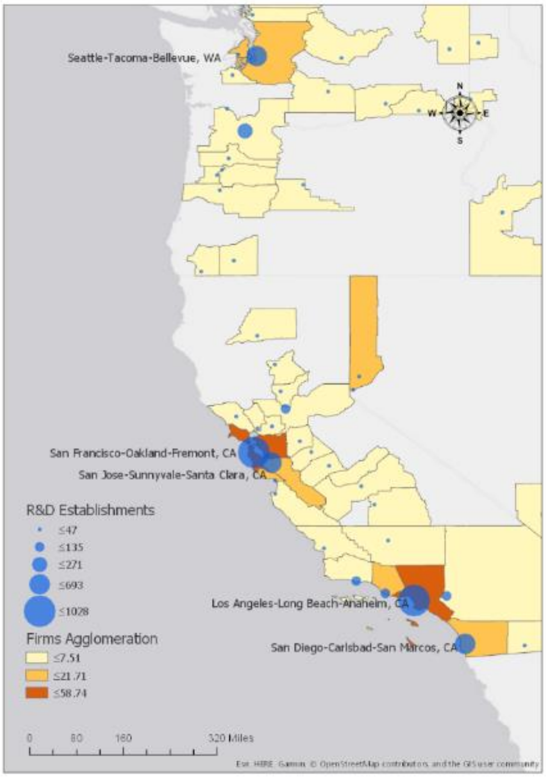


Figure 9 R&D establishment and MSA agglomeration in the US West Coast area in 2012

Based on the map in Figure 8, MSA with a high number of R&D establishments are likely to have the high agglomeration level as well, for example, New York-Northern New Jersey-Long Island (NY-NJ-PA) and Boston-Cambridge-Quincy (MA-NH). However, for MSA such as Washington-Arlington-Alexandria (DC-VA-MD-WV) which does not necessarily has the high agglomeration level, seems to remain favorable for firms to establish R&D activities. On the other hand, MSA such as Philadelphia-Camden-Wilmington (PA-NJ-DE-MD) where agglomeration level is high does not always attract companies to set up R&D activities there.

Meanwhile as shown in Figure 9, the number of R&D activities in the west coast area is most likely in line with the level of agglomeration. For instance, San Francisco-Oakland-Fremont (CA) and Los Angeles- Long Beach- Anaheim (CA) are MSAs with the most R&D establishments and also the highest level of agglomeration. Meanwhile, MSAs with the moderate number of R&D establishments such as Seattle-Tacoma-Bellevue (WA), San Jose-Sunnyvale-Santa Clara (CA) and San Diego-Carlsbad-San Marcos (CA) has a fair level of agglomeration.

4.1.2. The spatial distribution of patent grants

The United States patent data was derived from the U.S. Patent and Trademark Office (USPTO)'s Technology Assessment and Forecast database. It shows a count of the patent in technology classes defined by the U.S. Patent Classification System. It covers the number of patents granted in 2012 originated from approximately 300 Metropolitan Statistical Area (MSA).

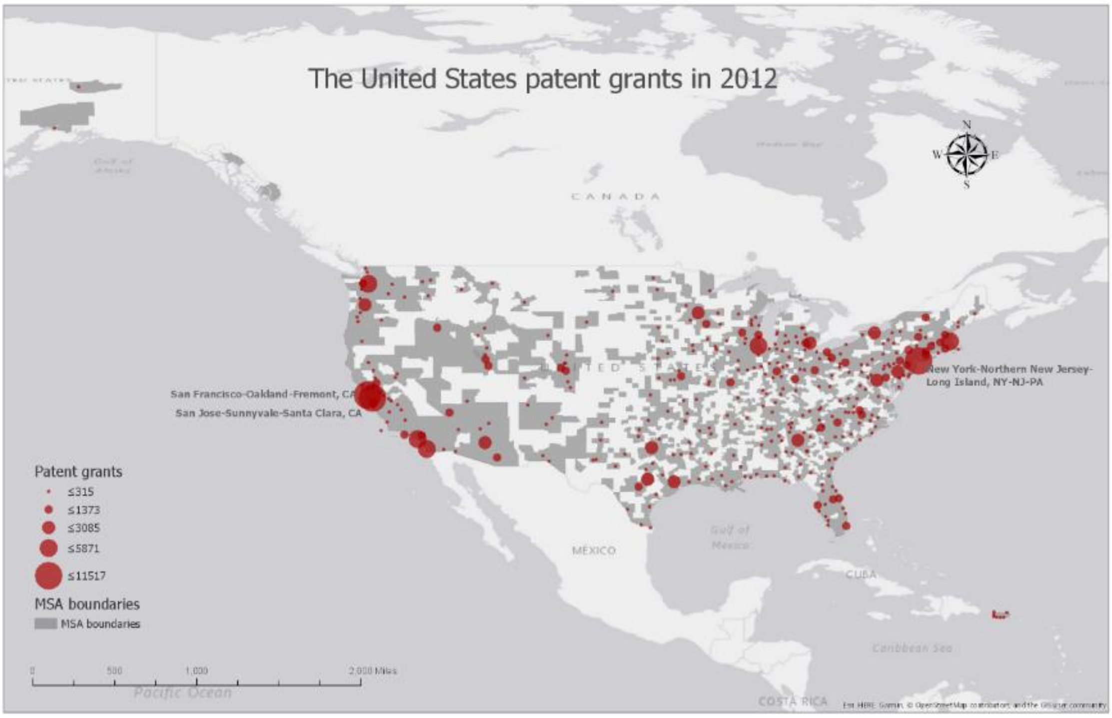


Figure 10 the United States patents distribution by Metropolitan Statistical Area in 2012

Source: The United States Patent and Trademark Office

According to Figure 10, the patent distributions in the United States are spatially uneven. The MSAs in the east part of the United States appears to be more favorable for researchers than MSA in the west part. It is also obvious that the MSA with a high number of patents is mostly located in the coastal area. In general, this condition is similar to what occurs in R&D distributions.

Furthermore, the patent distribution also has two areas of concentration that similar to R&D areas of concentration. The first concentration lays along the north-east area with New York-Northern New Jersey-Long Island (NY-NJ-PA) as the center of concentration since it has the most patent counts in the area. Meanwhile, in the west coast area of the US, San Francisco-Oakland-Freemont (CA) and San-Jose-Sunnyvale-Santa Clara (CA), becomes another bigger center since these two MSA are spatially adjacent to each other. All in all, these three leading MSAs obviously appear as the big hub of innovation due to their preeminent locational advantage as the center of urban growth in the US.

Figure 11 the top 25 United States metropolitan area of patents

Source: The United States Patent and Trademark Office

Based on the chart in Figure 11, the patents distribution in the United States are obviously concentrated in particular MSAs. For instance, San Jose-Sunnyvale-Santa Clara (CA) is distinctly in the first place in term of patent counts by becoming the only MSA that almost reach 12000 grants in a year. It proves that San Jose-Sunnyvale-Santa Clara (CA) is strongly attractive for the inventors. This result becomes more valid as the biggest innovation hub Silicon Valley is located in this MSA. The other MSAs in the top ten (10) of patent counts are mostly the similar MSAs as in the top 10 of R&D Establishments, except for Minneapolis-St. Paul-Bloomington (MN-WI) and Detroit-Warren-Livonia (MI). However, Washington-Arlington-Alexandria (DC-VA-MD-WV) which is the MSA that accommodate most of the R&D activities, is not even in the top 10 regarding patent grants. All in all, patent grants occurs more in the MSAs where have more R&D activities.

The next discussion will explain how the agglomeration relates to the patent counts distribution in the two area of concentration: the northeast hub and the west hub. In Figure 12 and 13, I will again overlap the agglomeration data with patent data to seek the general overview about their association.

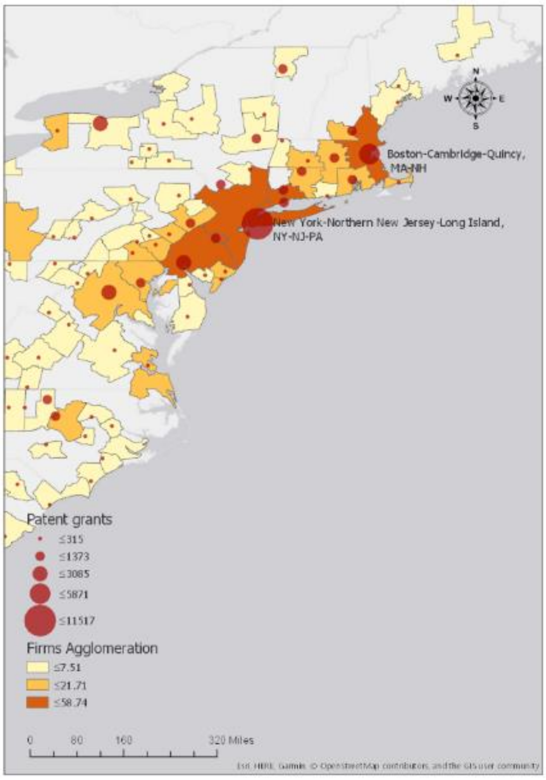


Figure 12 patent grants and MSA agglomeration in the US East Coast area in 2012

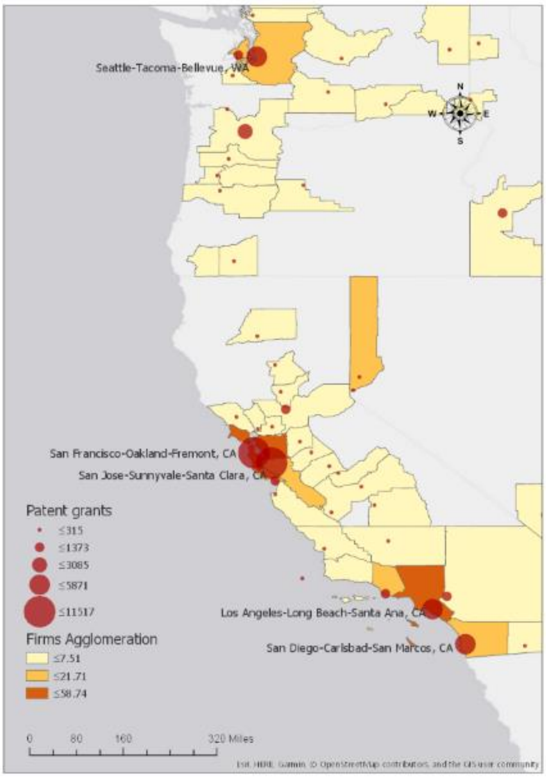


Figure 13 patent grants and MSA agglomeration in the US West Coast area in 2012

In general, the patent counts apparently in line with the agglomeration level. For instance, in the northeast area, both of New York-Northern New Jersey-Long Island (NY-NJ-PA) and Boston-Cambridge-Quincy (MA-NH) are located in where agglomeration level is range from 22 firms up to 60 firms per square miles. As well as in the west area, patents emerge most in the San Francisco-Oakland-Hayward, CA which is highly agglomerated. However, the San Jose-Sunnyvale-Santa Clara (CA), which is the MSA with the highest rank in term of patents counts, are not classified as the high agglomerated MSA. In this case, San Jose-Sunnyvale-Santa Clara (CA) might get the benefit from its neighbor MSA's agglomeration spill over.

All in all, the agglomeration level sounds to have a reciprocal relation with both R&D establishments and patent grants. This result is in line with what Audretsch & Feldman (2004) and Carlino and Kerr (2015) who argue that innovation emerges at where the concentration of economic are there. This understanding underlies the idea of the role that agglomeration play as the moderating variable between R&D establishment and air transportation network. Thus, I further discuss the agglomeration role in the inferential analysis section.

4.2 Network Analysis

For the network analysis in STATA, I utilize the air transportation data of "Air Carrier Statistics (Form 21 Traffic)-All Carriers" provided by United State Bureau Transportation Statistic. The data shows the airport to airport connection for both domestic and international flight. I identify the MSA code for each airport to get the connection information in MSA level. Thus, the information about Origin-Destination, in this case, the number of passengers flown per each OD connection can be provided in MSA level. Moreover, I include the connectivity to other OD cities outside di MSA and the United States since I believe that the connection to other regions especially foreign cities which significantly determines the magnitude of air passenger flow towards each MSA.

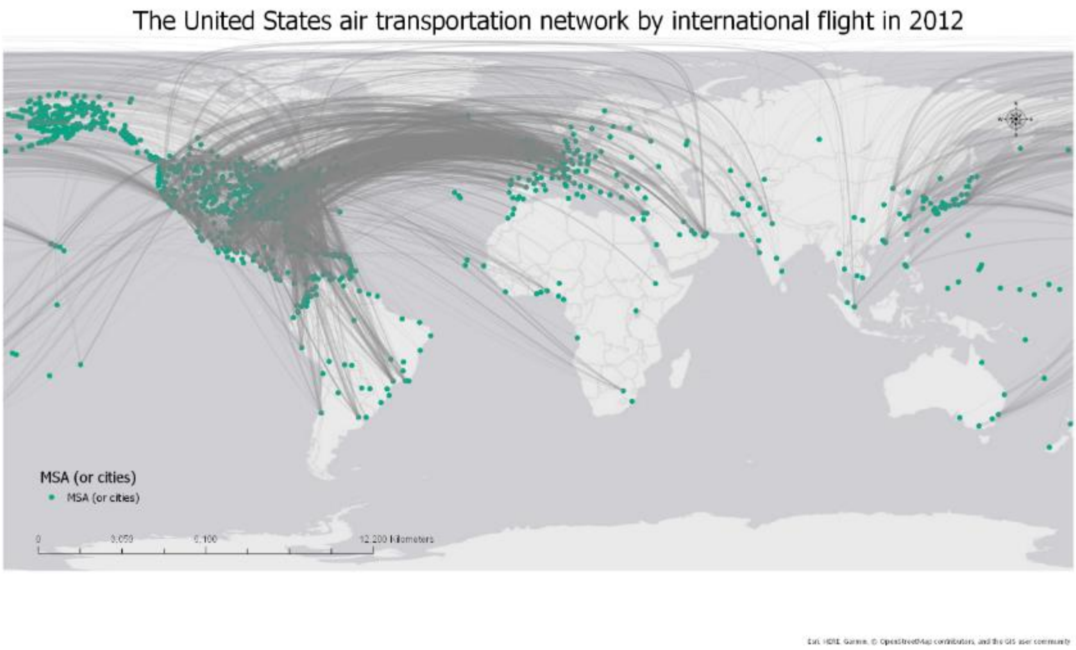


Figure 14 The United States air passenger network according to the international flight by MSA in 2012

Source: The United States Bureau Transportation Statistic

I illustrate the air passenger flow in Figure 14 to see the pattern of air connectivity according to the United States air passenger data in 2012. According to the international connection, it is obvious that the biggest group of passenger flows occurs between the US ODs and the European ODs. According to the data, the highest connection in this group occurs between New York-Northern Jersey City-Long Island (NY-NJ-PA) and London, United Kingdom (approximately 3.9 million passengers). Another significant group of flows also occurs with the South-American and Asian ODs, however not so many connections with African and Australian ODs.

Meanwhile, according to the following Figure 15, the US domestic flight evenly occurs in the entire area of the US. For instance, I found the high level of air passenger connectivity within the US mainland area (indicated by the dense grey lines across ODs). Additionally, the area of islands such as the state of Hawaii, the Commonwealth of Puerto Rico and the Commonwealth of Northern Mariana Island, mainly have direct access to the US mainland area. Meanwhile, in the Alaska area, the Anchorage (AK) and Fairbanks (AK) are the main hubs that connect the Alaskan ODs with the mainland ODs. Thus, I argue that the regions in the US are very well connected by air transportation.

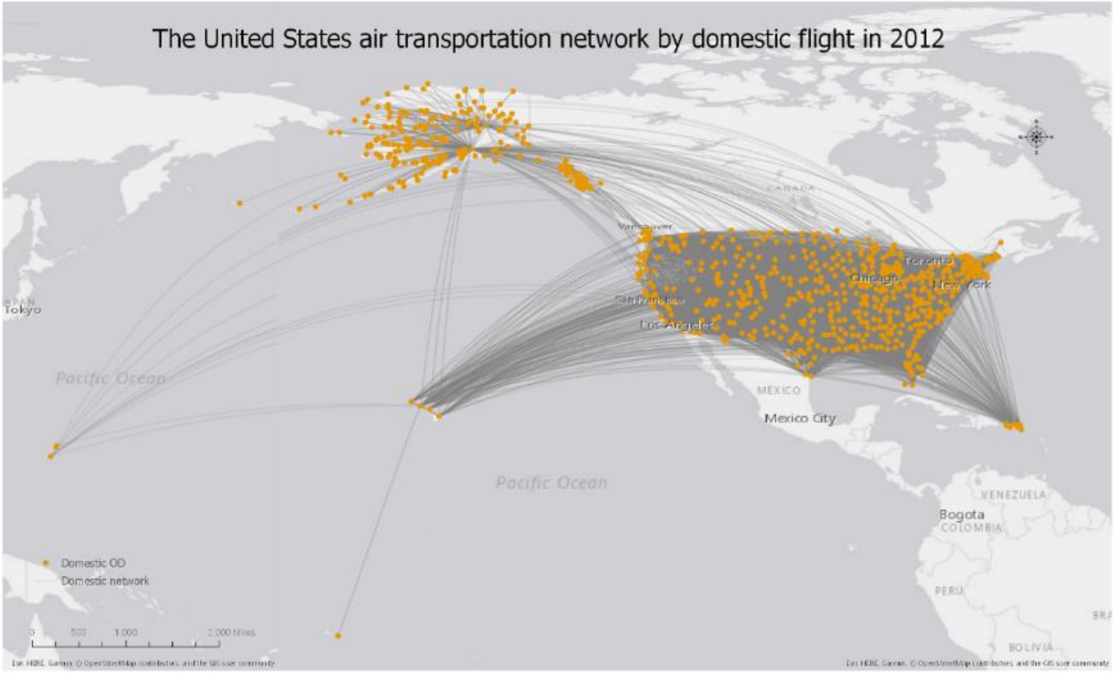


Figure 15 The United States air passenger network according to the domestic flight by MSA in 2012

Source: The United States Bureau Transportation Statistic

As a result from network analysis, approximately 1200 Origin-Destination in the entire network has their value of the degree centrality, betweenness centrality and closeness centrality (the summary of the overall result from network analysis is provided into Appendix 1). In part of the network, there is 474 OD of the United States MSA. The composition of OD indicates that the connection of air transportation in the United States mostly come from other OD cities outside the MSA. The following Table 6. Display the descriptive statistic of network centrality of the MSA in the United States.

Table 6 The descriptive statistics of the United States MSA's status in 2012

Network measurements	Obs	Mean	Std. Dev.	Min	Max
Degree centrality	474	1285.769	4595.024	.0015291	43307.85
Betweeness centrality	474	.0076288	.0214728	0	.3104869
Closeness centrality	474	.0008143	.0001395	.000023	.0008515

Source: The United States Bureau Transportation Statistic

According to Figure 16. I found that the MSAs with the high degree centrality are relatively disperse. For instance, New York-Northern Jersey City-Long Island (NY-NJ-PA), Atlanta-Sandy Springs-Roswell (GA), Miami-Fort Lauderdale-West Palm Beach (FL) appear with as the MSAs with high degree centrality from the east part of the country. Meanwhile, more to the center of the US, the high degree centrality fall in Chicago-Naperville-Elgin (IL-IN-WI) and Dallas-Fort Worth-Arlington (TX). These five MSAs account for the huge amount of domestic connections within the United States and many international connections to the European and African destinations. On the other hand, Los Angeles-Long Beach-Anaheim (CA) and San Francisco-Oakland-Hayward (CA) are the only two MSA from the west part of the United States. Besides their enormous international connections, Los Angeles-Long Beach-Anaheim (CA) and San Francisco-Oakland-Hayward (CA) are mainly connected to the South American and Asian destinations.

Additionally, it is noticeable that the largest circle on the map in Figure 17 lays in Anchorage (AK). This result is in line with what Cheung and Gunes (2012) find that the Anchorage airport was the airport with highest betweenness centrality in 2011 and 1991. Additionally, (Guimera, Mossa, et al., 2005) also find that Anchorage airport is the second highest betweenness degree in the world air network system. Guimera, Mossa, et al. (2005) argue that the geographical factor is the main reason to explain the high number of betwenness degree, which is sensible since Alaska State is separated from the US mainland by Canada. They also argue that people in Alaska need access to the political centers in the US. Thus, the Anchorage (AK) functioned as the main transfer hub to connect the other Alaskan regions to the US continent.

Meanwhile, the distribution of closeness centrality is more equal among MSAs in the US. Referring to the map in Figure 18, most of the MSAs have the same value of closeness centrality (there is no particular MSAs with a prominent point in the map). In fact, it is contradicting to what previously discussed in the case of China, where the degree centrality and the cities with high degree centrality are most likely appear as the cities with high closeness centrality (Wang, Mo, et al., 2011). I argue that in the case of US, most of MSAs have a direct connection to the other MSAs through the air transportation. Thus the network distance from each MSAs to reach the other MSAs are relatively similar.

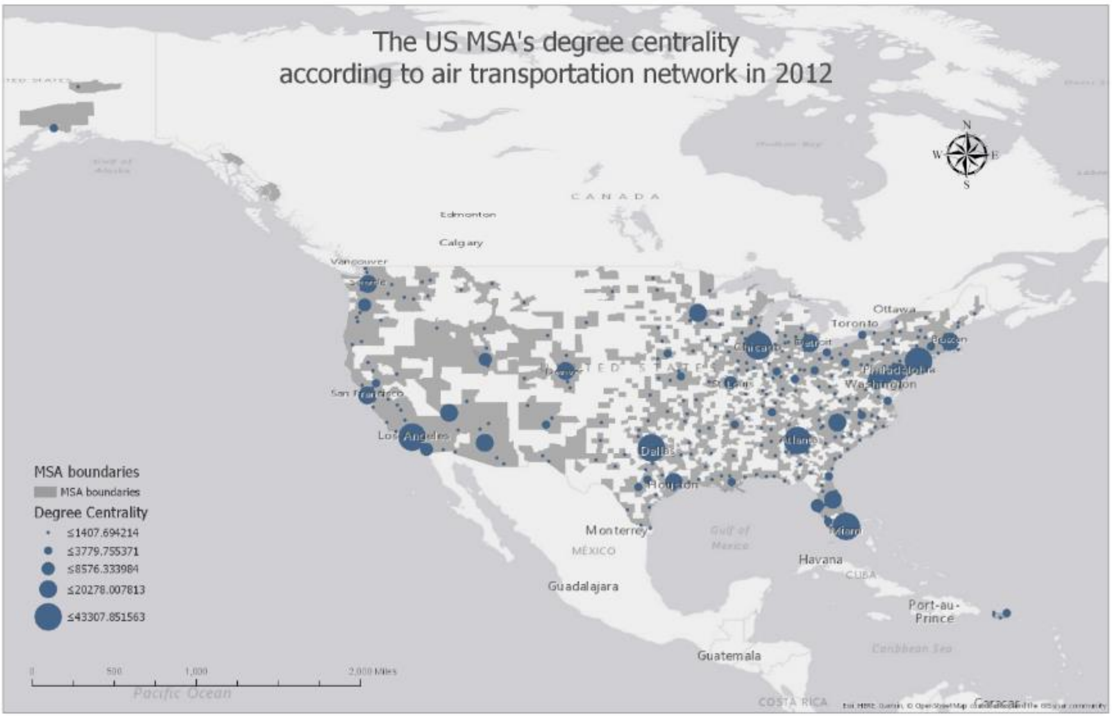


Figure 16 the United States MSA's degree centrality spatial distribution in 2012

Source: The United States Bureau Transportation Statistic

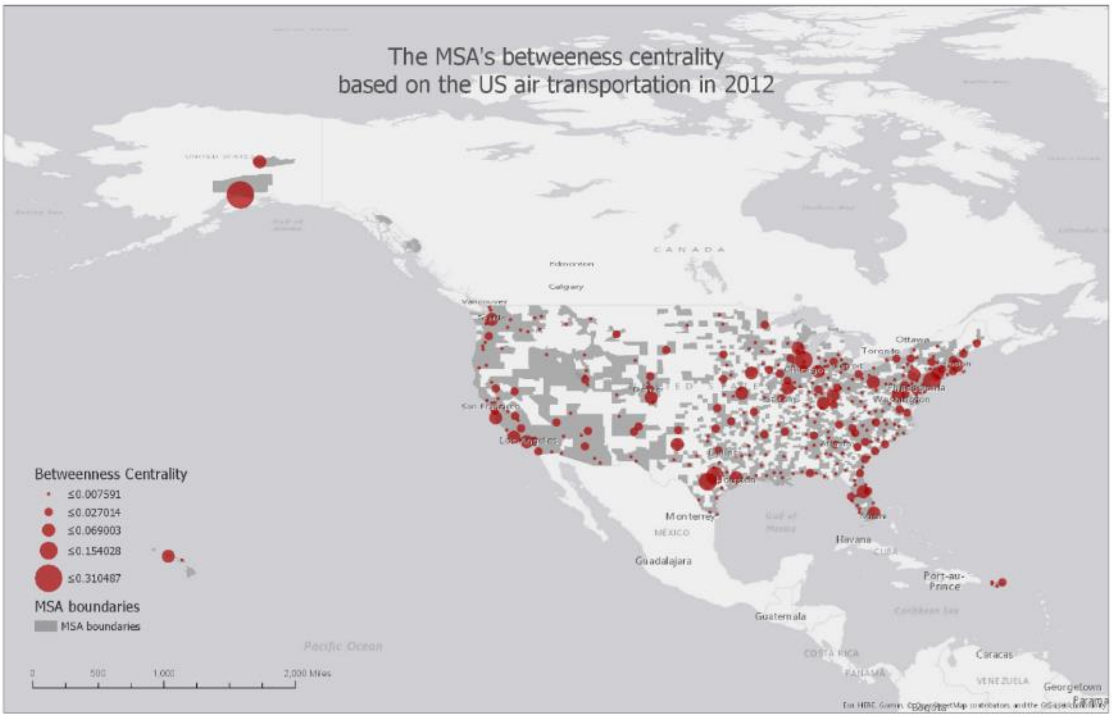


Figure 17 The United States MSA's betweenness centrality spatial distribution in 2012

Source: The United States Bureau Transportation Statistic

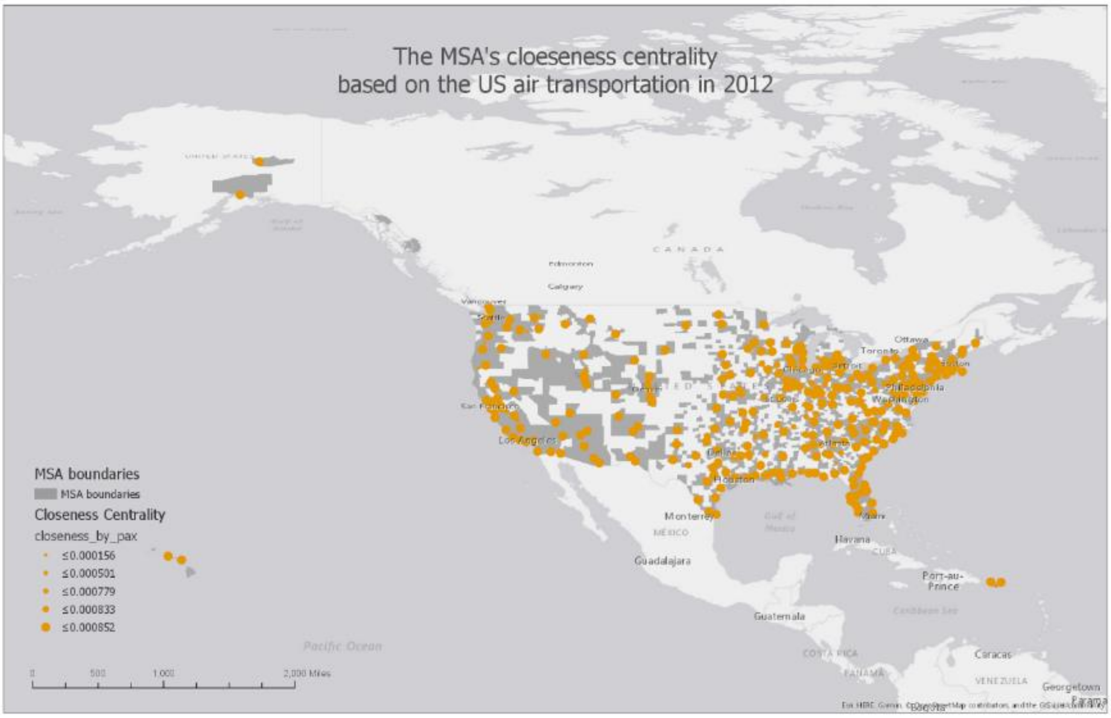


Figure 18 The United States MSA's closeness centrality spatial distribution in 2012

Source: The United States Bureau Transportation Statistic

To identify the MSA that has a high status of network centrality measurement, I display the top ten MSA in table 7. In the case of degree centrality, New York-Northern Jersey City-Long Island (NY-NJ-PA) is the MSA with the highest degree of centrality. This likely to happen since the two major airports in the United States namely Newark Int. Airport (EWK) and John F Kennedy Int. Airport (JFK) and are located in this MSA. In term of the betweenness centrality, it is obvious that Anchorage (AK) is the most important MSA with the highest betweenness centrality. In fact, Ted Stevens Anchorage International Airport (ANC) is one of the main airports that connects the domestic flights from any other Alaskan ODs to the rest of ODs in the United States. Lastly, the closeness centrality has the unexpected results since the cities appeared in the top ten list are the cities that barely discovered in this thesis discussion. For instance, South Bend-Mishawaka, IN-MI has an effective distance to reach the other OD di the network system by having the connection to many MSAs with high degree of centrality such as New York-Northern Jersey City-Long Island (NY-NJ-PA), Atlanta-Sandy Springs-Roswell (GA), Chicago-Naperville-Elgin (IL-IN-WI), and San Francisco-Oakland-Hayward (CA).

Rank	Degree Centrality	Betweenness centrality	Closeness centrality
1	New York-Northern Jersey City-Long Island, NY-NJ-PA	Anchorage, AK	South Bend-Mishawaka, IN-MI
2	Atlanta-Sandy Springs-Roswell, GA	New York-Newark-Jersey City, NY-NJ-PA	Scranton--Wilkes-Barre--Hazleton, PA
3	Chicago-Naperville-Elgin, IL-IN-WI	Milwaukee-Waukesha-West Allis, WI	Appleton, WI
4	Los Angeles-Long Beach-Anaheim, CA	Austin-Round Rock, TX	Milwaukee-Waukesha-West Allis, WI
5	Miami-Fort Lauderdale-West Palm Beach, FL	San Antonio-New Braunfels, TX	Austin-Round Rock, TX
6	Dallas-Fort Worth-Arlington, TX	Los Angeles-Long Beach-Anaheim, CA	Colorado Springs, CO
7	San Francisco-Oakland-Hayward, CA	Santa Maria-Santa Barbara, CA	Lexington-Fayette, KY
8	Denver-Aurora-Lakewood, CO	Miami-Fort Lauderdale-West Palm Beach, FL	Cincinnati, OH-KY-IN
9	Houston-The Woodlands-Sugar Land, TX	Boston-Cambridge-Newton, MA-NH	Greensboro-High Point, NC
10	Washington-Arlington-Alexandria, DC-VA-MD-WV	Fairbanks, AK	Salinas, CA

Table 7 the United States top 10 MSA ranked by degree, betweenness and closeness centrality in 2012

Source: Author, The United States Bureau Transportation Statistic

For further discussion in this thesis research, I will narrow down the definition of the MSA status in air transportation network as the degree centrality of an MSA. I believe that the degree centrality is the most sensible measurement in explaining regional innovation since it simply represents the magnitude of passengers' flows among MSAs (Guimera, Mossa, et al., 2005) which is commonly associated with socioeconomic concentration (Schaafsma, 2003) and accessibility to the sources of information. Referring to what Friedman (2002) argues, the concentration of people and the open-minded society are the main locational factors that attract the creative class, in this case, in advance technology classification. Therefore, in the next section, I will further discuss the relation of MSAs degree centrality to the regional innovations, by proceeding with the inferential analysis to statistically test their relation.

4.3 Inferential Analysis

In the following Table 8 and 9, I present the results of OLS regression to exercise the linear relation between the dependent variables (R&D Establishment and patent counts) and the independent variable of degree centrality. I also exercise the relation using negative binomial regression and provide the results in the appendix.

4.3.1 OLS regression

I exercise the relation in three model for each dependent variable. According to Table 8, OLS-Model 1 to OLS-Model 3 shows the regression between R&D establishments and air network degree centrality. Meanwhile, according to table 9, OLS-Model 4 to OLS-Model 6 show the regression between patent grants and air network degree centrality. In the OLS-Model 1 and OLS-Model 4, I test the influence of air network degree centrality on the dependent variables without incorporating the agglomeration variable into the model. In OLS-Model 2 and OLS-Model 5, I add the agglomeration variable to test the influence of both air network degree centrality and agglomeration on each dependent variables. Lastly, in OLS-Model 3 and OLS-Model 6, I interact the air network centrality and agglomeration to test the significance of their interaction on the dependent variables.

Table 8 the OLS regression result for the dependent variable: R&D establishment

	Dependent variable : (log) R&D establishment		
VARIABLES	OLS-Model 1	OLS-Model 2	OLS-Model 3
Air passenger network (degree centrality)	0.115***	0.096***	0.164***
	(0.021)	(0.019)	(0.029)
i.MSA with top 100 universities	1.057***	0.950***	0.840***
	(0.198)	(0.204)	(0.201)
Per capita income (thousand dollars)	0.036***	0.028**	0.022**
	(0.012)	(0.012)	(0.011)
Crime (rate per 1 million people)	-0.039	-0.044	-0.055
	(0.061)	(0.061)	(0.059)
i.MSA with top 25 vacation destination	0.247	0.150	0.433
	(0.294)	(0.291)	(0.263)
Agglomeration (number of firms per sq miles)		0.035**	0.068***
		(0.018)	(0.018)
Air passenger network##agglomeration			-0.004***
			(0.001)
Constant	0.866	1.081**	1.204**
	(0.557)	(0.543)	(0.495)
Observations	252	252	252
R-squared	0.489	0.500	0.542

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

According to Table 8, the passenger network variable exhibits a positive relation with R&D establishment and the relation is statistically significant in every model. In OLS-Model 1, 1 unit change of air passenger network affects 11.5% change of R&D establishments to the reciprocal direction, holding other factors constant. In OLS-Model 2, the influence of air passenger network remains significant, except the decrease of its coefficient to 9.6%, holding other factors including agglomeration constant. In the same model, the variable agglomeration, as expected, has a significant and positive impact on R&D establishments (Carlino and Kerr et al., 2015) by affecting 3.5% change on R&D, holding other factors constant.

At last, in OLS-Model 3, the influence of air passenger network and agglomeration on R&D establishments remain significant to the reciprocal direction. However, the interaction term between air passenger network and agglomeration is unexpectedly shown a negative sign. Thus, the interpretation of the negative interaction term in OLS-Model 3 is the higher value on agglomeration lessen the impact of air passenger network on R&D establishments.

According to Table 9, a similar situation occurs in the regression between air passenger network and patent grants where the influence of both passenger and agglomeration is significant in each model. In OLS-Model 4, one unit of the air passenger network accounts for 13.4% change on patent grants to the reciprocal direction, holding other factors constant. In OLS-Model 5, both air passenger network and agglomerations are positively related to the patent grants by having a coefficient of 9% and 8% consecutively. In OLS-Model 6, the interaction between air passenger network and agglomeration also indicates the negative sign which leads to the conclusion that the high agglomeration lower the air transportation effect on patent grants

Responding to the unexpected negative interaction term between air passenger network and agglomeration, I suspect this result appears due to the limited observation in the regression; in fact, the number of observation is considerably decreasing from approximately 400 to approximately 250 observations. However, if I refer to the coefficient of the interaction term in OLS-Model 3 and OLS-Model 6, the value is insignificant (nearly 0%) in both models.

Table 9 the OLS regression result for the dependent variable: patent grants

	Dependent variable : (log) Patent grants		
VARIABLES	OLS-Model 4	OLS-Model 5	OLS-Model 6
Air passenger network (degree centrality)	0.134***	0.090***	0.191***
	(0.024)	(0.024)	(0.030)
i.MSA with top 100 universities	1.324***	1.076***	0.913***
	(0.235)	(0.239)	(0.228)
Per capita income (thousand dollars)	0.041**	0.022	0.012
	(0.017)	(0.015)	(0.013)
Crime (rate per 1 million people)	-0.221***	-0.233***	-0.249***
	(0.076)	(0.075)	(0.071)
i.MSA with top 25 vacation destination	-0.100	-0.323	0.100
	(0.362)	(0.393)	(0.326)
Agglomeration (number of firms per sq miles)		0.082***	0.131***
		(0.024)	(0.027)
Air passenger network##agglomeration			-0.005***
			(0.001)
Constant	2.704***	3.197***	3.380***
	(0.756)	(0.666)	(0.594)
Observations	255	255	255
R-squared	0.482	0.522	0.583

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

4.3.2 Robustness check

4.3.2.1 Endogeneity check

As further investigation on endogeneity issue, I conduct 2SLS regression based on the equation of the OLS-Model 3 and OLS-Model 6. According to Table 10, the explanatory variable is hardly significant in explaining both R&D (as shown at 2SLS-Model 1 column) and patent (as shown at 2SLS-Model 2 column). Only firms' agglomeration is statistically significant in the 2SLS regression. On the other hand, the instrumented variable (air passenger network) and the interaction term (air passenger network##agglomeration) are not statistically significant in the 2SLS regression. The insignificant sign on the air passenger network variable is inconsistent with what Lakew and Bilotkach (2018) previously studied. In this case, the choice of instrument may cause the insignificant result. I argue that there may be a relevant instrument to measure weather such as the damage value caused by storm, or the number of injuries and death caused by the storm which better explain the weather as the instrumental variables. Thus, further exploration on instrumental variable may worth. Otherwise, I suggest another approach such as panel regression to investigate deeper the endogeneity issue.

Table 10 2SLS regression result

	Dep variable : (log) RnD IV: Storm	Dep variable : (log) Patent grants IV: Storm
VARIABLES	2SLS-Model 1	2SLS-Model 2
Air passenger network (degree centrality)	-0.741	-0.800
	(0.931)	(1.023)
1.MSA with top 100 universities	1.910	2.088
	(1.228)	(1.354)
Per capita income (thousand dollars)	0.040	0.033
	(0.030)	(0.033)
Crime (rate per 1 million people)	-0.101	-0.299
	(0.167)	(0.182)
1.MSA with top 25 vacation destination	2.446	2.304
	(2.286)	(2.515)
Agglomeration (number of firms per sq miles)	0.090*	0.155***
	(0.050)	(0.055)
Air passenger network##agglomeration	0.017	0.017
	(0.021)	(0.023)
Constant	1.160	3.310***
	(1.127)	(1.233)
Observations	252	255

Instrumented: Air passenger network (degree centrality)

Instruments: 1.MSA with top 100 universities, Per capita income, Crime,
1.MSA with top 25 vacation destination, Agglomeration,
Air passenger network*agglomeration, Storm

Standard errors in parentheses

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

4.3.2.2 Spatial influence

As the further investigation on spatial aspect, I conduct the GWR to check pattern on the geographical influence of the explanatory variables on the dependent variable. First, I refer to 'OLS-Model 3' to define the equation of 'GWR-Model' to test the influence of the degree centrality of a region on R&D establishments in geographically weighted regression. I exclude the explanatory variable crime considering it is not significant in explaining R&D establishments in the OLS regression. Meanwhile, to test the influence of degree centrality on patent grants in geographically weighted regression, I define GWR-Model 2 equation. I refer 'to OLS-Model 6' to define the GWR-Model 2 equation, but excluding the explanatory variable income because it is not significant in explaining patent grants in OLS regression.

GWR-Model 1

$$Innovation_{MSA}^{r\&d}(g) = \alpha_0(g) + \alpha_1(g) * Degree\ Centrality_{MSA}^{Passengers} + \alpha_2(g) * university + \alpha_3(g) * income + \alpha_4(g) * agglomeration + \alpha_5(g) * Degree\ Centrality_{MSA}^{Passengers} * agglomeration + \epsilon \qquad (9)$$

GWR Model 2

$$Innovation_{MSA}^{patent}(g) = \alpha_0(g) + \alpha_1(g) * Degree\ Centrality_{MSA}^{Passengers} + \alpha_2(g) * university + \alpha_3(g) * crime + \alpha_4(g) * agglomeration + +\alpha_5(g) * Degree\ Centrality_{MSA}^{Passengers} * agglomeration + \epsilon \qquad (10)$$

(g)=refers to the location (x,y) of each observation

Based on Table 11, both OLS regression results and GWR results are mostly consistent between each other in explaining R&D establishments. In term of R&D establishments, the coefficient of explanatory variables in GWR result exhibits the similar sign with the coefficient of explanatory variables in OLS regression result. However, the constant signs in both model show the opposite direction. Additionally, the average R^2 of GWR-Model 1 is higher than the OLS result. The GWR result means that, in general, the likelihood of GWR-Model 1 explain the real situation is higher than the likelihood of OLS regression result. Furthermore, according to the map in Figure 19, there are obvious clusters of MSA with high R^2 of GWR-Model 1 in the north-east area of the US (I provide the Moran's I Spatial Autocorrelation Report in the Appendix). Indeed, this pattern is congruent with the distribution of R&D establishments (Figure 5).

Moreover, according to Table 12, both OLS regression results and GWR results are consistent with each other in explaining patent grants. All the coefficient signs of explanatory variables in GWR result show the same signs as the coefficient signs in OLS regression. Also, the average R^2 of GWR-Model 2 is also higher than the OLS result. Thus the GWR-Model 2 explain more likely as the real situation. Additionally, the Figure 20 demonstrate the clusters of MSA with high R^2 of GWR-Model 2 are located in west-coast area (I provide the Moran's I Spatial Autocorrelation Report in the Appendix). The R^2 distribution of GWR-Model 2 shows the opposite situation described in Figure 18. In fact, the pattern in Figure 19 is congruent with the distribution of patent (Figure 9).

Table 11 the comparison between GWR and OLS regression result for the dependent variable: R&D Establishments

	Dependent variable : (log) R&D			
	GWR Model 1 result			OLS result
VARIABLES	β mean	β min	β max	β
Air passenger network (degree centrality)	0.231	0.086	0.934	0.167***
1.MSA with top 100 universities	0.716	0.214	1.322	0.817***
Per capita income (thousand dollars)	0.044	0.014	0.140	0.026**
Agglomeration (number of firms per sq miles)	0.156	0.037	0.412	0.076***
Air passenger network##agglomeration	-0.011	-0.076	-0.002	-0.003***
Constant	-0.109	-3.453	2.163	0.780**
R-squared	0.680			0.545
Adjusted R-squared	0.630			
AICc	845.85			
Residual Square	223.2			
Neighbors	121			

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

Figure 19 The local R-squared distribution of GWR-Model 1 result

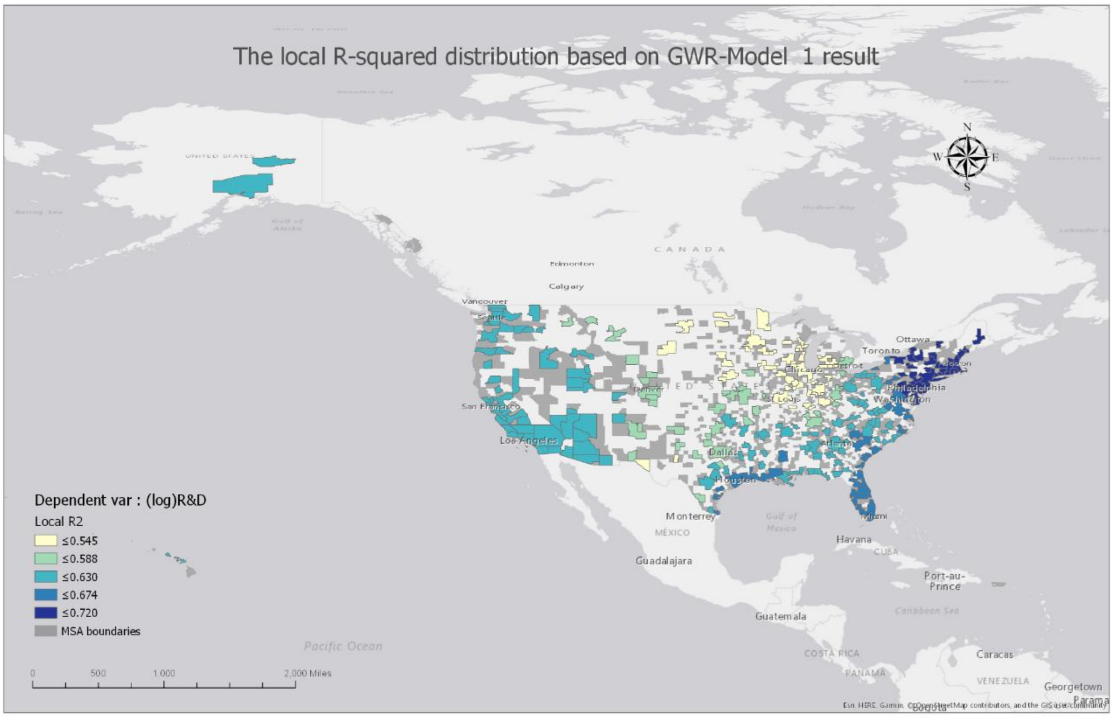
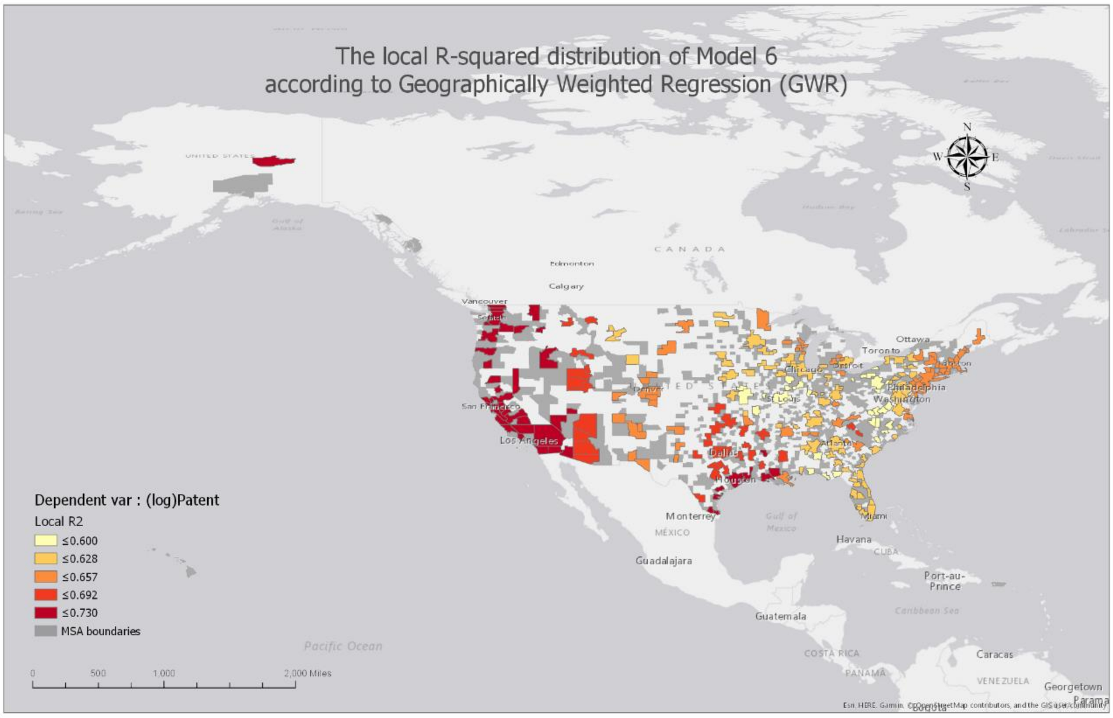


Table 12 the comparison between GWR and OLS regression result for the dependent variable: Patent grants

	Dependent variable : (log) patent			
	GWR Model 2 result			OLS result
VARIABLES	β mean	β min	β max	β
Air passenger network (degree centrality)	0.279	0.115	1.211	0.194***
1.MSA with top 100 universities	0.938	0.318	1.276	0.933***
Crime (rate per 1 million people)	-0.236	-0.402	-0.059	-0.265***
Agglomeration (number of firms per sq miles)	0.280	0.086	0.574	0.141***
Air passenger network##agglomeration	-0.017	-0.100	-0.003	-0.005***
Constant	3.405	1.617	4.268	3.862***
R-squared	0.716			0.581
Adjusted R-squared	0.667			
AICc	793.4			
Residual Square	250.7			
Neighbors	103			

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

Figure 20 The local R-squared distribution of GWR-Model 2 result



Chapter 5. Conclusion

5.1 Conclusion

To answer the main research question of this thesis research, I would like to conclude this thesis discussion by highlighting the several findings. Firstly, I find a highly unequal distribution of regional innovation in the United States in 2012. There is two main area of regional innovation concentration in the United States which I identify as the north-east cluster and the west cluster. This result is consistent with the finding from Buzard, Carlino et al. (2016). The north-east cluster host the top three MSA in R&D establishment which are Washington-Arlington-Alexandria (DC-VA-MD-WV), New York-Newark-Jersey City (NY-NJ-PA) and Boston-Cambridge-Quincy (MA-NH) consecutively from highest rank. On the other hand, the west cluster host the top two MSA in patent grants which are San Jose-Sunnyvale-Santa Clara (CA) and San Francisco-Oakland-Hayward (CA) respectively from the highest. Meanwhile, the third rank in patent grants falls to New York-Northern New Jersey-Long Island, (NY-NJ-PA) which is in the north-east cluster. Hence, New York-Newark-Jersey City (NY-NJ-PA) appear to be the MSA with excellent R&D establishment and patent grants.

Secondly, I generate the region status in air passenger network using social network centrality measurement: degree centrality, betweenness centrality, and closeness centrality (Freeman, 1978). I find that the top three MSA in term of degree centrality is New York-Northern New Jersey-Long Island, (NY-NJ-PA), Atlanta-Sandy Springs-Roswell (GA), and Chicago-Naperville-Elgin (IL-IN-WI) consecutively from the top rank. Meanwhile, the top three MSA with high betweenness centrality is Anchorage (AK), New York-Northern New Jersey-Long Island, (NY-NJ-PA) and Milwaukee-Waukesha-West Allis (WI) consecutively from the highest rank. Lastly, it is surprising that the top three of MSA in term of high closeness centrality are South Bend-Mishawaka (IN-MI), Scranton-Wilkes-Barre-Hazleton (PA) and Appleton (WI). Indeed, New York-Northern New Jersey-Long Island, (NY-NJ-PA) appear as the MSA with the excellent air passenger network centrality in term of degree and betweenness which also means the excellent air passenger transport connectivity.

Thirdly, according to OLS regression analysis, the air passenger network exhibit a positive correlation with both R&D establishment and patent grants. Excluding firms' agglomeration aspect in the regression, the air passenger network (measured by degree centrality) affects 11.5% change of R&D establishment, while it affects 13.4% increase in patent grants. This result is consistent with other previous studies on air transportation and urban development such as Kalayci and Yanginlar (2016) as well as Bannò and Redondi (2014).

Lastly, I find that the firm's agglomeration level significantly affect the regional innovation, yet weaken the influence of the air transportation on both R&D establishment and patent grants. However, the coefficient values of the interaction term between air passenger network and firms' agglomeration are negligible. The interaction term between air passenger network (measured by degree centrality) and firm's agglomeration (measured by firms number per square miles) weaken 0.4% of the air passenger network influence on R&D establishment, while lessening 0.5% of the air passenger network influence on patent grants. Additionally, The OLS regression result is consistent with the result of the geographically weighted regression, yet differ from the 2SLS regression which shows the air passenger network as the insignificant variable in explaining R&D establishment and patent grants. All in all, the firms' agglomeration has a moderating role in lowering the impact of air passenger network on regional innovation.

5.3 Policy recommendation

In accordance with this thesis results, I suggest that the policy on improving air passenger network centrality also improves the regional innovation level. At first, the US government may want to create the favorable environment for innovation activities by increasing the air connectivity of a region. Several attempts such as opening the new route, improving the route frequency or expanding the airport runaway capacity may contribute to improving network centrality of a region. Also, the government may also want to consider the existing MSA centrality status to generate the effective planning and to define further regulation in air transportation network. Therefore, the excellent air transportation system will significantly contribute to creating the conducive environment for innovation activities.

Additionally, the US government may also want to promote the firms' agglomeration to boost the regional innovation. The enabling factors such as the availability of clear and easy procedures for new business establishments and the effective and efficient public infrastructures for business activities also improve the innovation activities in a region. Therefore, by incorporating those enable factors in comprehensive policy documentation, Government may have an effective measurements to induce the regional innovation level.

5.2 Limitation and recommendation for future research

In spite of this thesis contribution, I acknowledge that there are some limitations due to variable selection and methodology. First, this thesis only based on cross-sectional data. Indeed, the thesis result could not be generalized since it does not have any time variance. Thus, I recommend the utilization of panel data on the future study of the relation between regional innovation and air passenger network.

Secondly, I only use the number of R&D and patents origin to measure regional innovation. Whereas, these variables do not inform the urban socio-economic values of regional innovation. Hence, the deeper exploration on how to convert the regional innovation into the urban socio-economic value (such as urban happiness and wealth) may worth for urban development study.

Next, in this study, I particularly consider the air passenger network to define MSAs centrality status. In fact, the MSA centrality status of the ground transportation and water transportation in the urban network system may also be the crucial determinants for regional innovation. Therefore, I suggest to further study the influence of integrated transportation network on regional innovation.

Lastly, I mainly test the one direction relationship of air passenger network on regional innovation through OLS regression. Even though I provide the overview of endogeneity check on the model, further investigation of causal direction on the relation between both variables worth to conduct. Thus, the urban actors or any stakeholders will get a better view in studying regional innovation and air transportation network.

Bibliography

Alderson, A. S. and Beckfield, J. 2004. Power and position in the world city system. *American Journal of Sociology,* 109 (4), pp. 811-851.

Andersson, R., Quigley, J. M. and Wilhelmsson, M. 2009. Urbanization, productivity, and innovation: Evidence from investment in higher education. *Journal of Urban Economics,* 66 (1), pp. 2-15.

Audretsch, D. B. and Feldman, M. P. 2004. Knowledge spillovers and the geography of innovation. Knowledge spillovers and the geography of innovation. 2004. Handbook of regional and urban economics. Elsevier. pp. 2713-2739.

Bannò, M. and Redondi, R. 2014. Air connectivity and foreign direct investments: economic effects of the introduction of new routes. *European Transport Research Review,* 6 (4), pp. 355-363.

Begg, I., 1999. Cities and competitiveness. *Urban Studies,* 36 (5-6), pp. 795-809.

Bettencourt, L. M., Lobo, J. and Strumsky, D. 2007. Invention in the city: Increasing returns to patenting as a scaling function of metropolitan size. *Research Policy,* 36 (1), pp. 107-120.

Breschi, S., 2000. The geography of innovation: A cross-sector analysis. *Regional Studies,* 34 (3), pp. 213-229.

Bresnahan, T., Gambardella, A. and Saxenian, A. 2001. 'Old economy'inputs for 'new economy'outcomes: Cluster formation in the new Silicon Valleys. *Industrial and Corporate Change,* 10 (4), pp. 835-860.

Burger, M. and Meijers, E. 2012. Form follows function? Linking morphological and functional polycentricity. *Urban Studies,* 49 (5), pp. 1127-1149. Available at: http://usj.sagepub.com/content/49/5/1127.full.pdf [Accessed 05/10/2016].

Button, K., Lall, S., Stough, R. and Trice, M. 1999. High-technology employment and hub airports. *Journal of Air Transport Management,* 5 (1), pp. 53-59.

Button, K. and Yuan, J. 2013. Airfreight transport and economic development: an examination of causality. *Urban Studies,* 50 (2), pp. 329-340.

Buzard, K., Carlino, G. A., Hunt, R. M., Carr, J. K., et al., 2016. Localized knowledge spillovers: Evidence from the agglomeration of American R&D labs and patent data.

Cann, O., 2018. What is competitiveness? Available at: https://www.weforum.org/agenda/2016/09/what-is-competitiveness/ [Accessed 9 June 2018].

Carlino, G. and Kerr, W. R. 2015. Agglomeration and innovation. Agglomeration and innovation. 2015. Handbook of regional and urban economics. Elsevier. pp. 349-404.

Cheung, D. P. and Gunes, M. H., eds., 2012. A complex network analysis of the United States air transportation, [Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2012)]. IEEE Computer Society. pp. 699-701.

Cornet, M. and Rensman, M., 2001. The location of R&D in the Netherlands: trends, determinants and policy (No. 14). CPB Netherlands Bureau for Economic Policy Analysis. Available at: http://www.cpb.nl/sites/default/files/publicaties/download/location-rd-netherlands-trends-determinants-and-policy.pdf .

Debbage, K. G., 1999. Air transportation and urban-economic restructuring: competitive advantage in the US Carolinas. *Journal of Air Transport Management,* 5 (4), pp. 211-221.

Ellison, G., Glaeser, E. L. and Kerr, W. R. 2010. What causes industry agglomeration? Evidence from coagglomeration patterns. *American Economic Review,* 100 (3), pp. 1195-1213.

Ernst, D., 2009. A new geography of knowledge in the electronics industry? Asia's role in global innovation networks.

Feldman, M. P. and Florida, R. 1994. The geographic sources of innovation: technological infrastructure and product innovation in the United States. *Annals of the Association of American Geographers,* 84 (2), pp. 210-229.

Ferrary, M. and Granovetter, M. 2009. The role of venture capital firms in Silicon Valley's complex innovation network. *Economy and Society,* 38 (2), pp. 326-359.

Florida, R., 2002. Bohemia and economic geography. *Journal of Economic Geography,* 2 (1), pp. 55-71.

Freeman, L. C., 1978. Centrality in social networks conceptual clarification. *Social Networks,* 1 (3), pp. 215-239.

Givoni, M. and Banister, D. 2006. Airline and railway integration. *Transport Policy,* 13 (5), pp. 386-397.

Goetz, A. R., 1992. Air passenger transportation and growth in the US urban system, 1950–1987. *Growth and Change,* 23 (2), pp. 217-238.

Griliches, Z., 1998. Patent statistics as economic indicators: a survey. Patent statistics as economic indicators: a survey. 1998. R&D and productivity: the econometric evidence. University of Chicago Press. pp. 287-343.

Guimerá, R., Mossa, S., Núñez-Amaral, L. and Turtschi, A. 2003. Structure and efficiency of the world-wide airport network.

Guimera, R., Mossa, S., Turtschi, A. and Amaral, L. A. 2005. The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles. *Proceedings of the National Academy of Sciences of the United States of America,* 102 (22), pp. 7794-7799.

Hall, B. H., Jaffe, A. and Trajtenberg, M. 2005. Market value and patent citations. *RAND Journal of Economics,* pp. 16-38.

Hall, B. H., Lotti, F. and Mairesse, J. 2009. Innovation and productivity in SMEs: empirical evidence for Italy. *Small Business Economics,* 33 (1), pp. 13-33.

Irwin, M. D. and Kasarda, J. D. 1991. Air passenger linkages and employment growth in US metropolitan areas. *American Sociological Review,* pp. 524-537.

Kalayci, S. and Yanginlar, G. 2016. The Effects of Economic Growth and Foreign Direct Investment on Air Transportation: Evidence from Turkey. *International Business Research,* 9 (3), pp. 154.

Kerr, W. R. and Kominers, S. D. 2015. Agglomerative forces and cluster shapes. *Review of Economics and Statistics,* 97 (4), pp. 877-899.

Lakew, P. A. and Bilotkach, V. 2018. Airport delays and metropolitan employment. *Journal of Regional Science,* 58 (2), pp. 424-450.

Malecki, E. J., 1980. Dimensions of R&D location in the United States. Available at: http://www.sciencedirect.com/science/article/pii/0048733380900244 .

Marshall, A., 1898. Principles of economics. Vol. 1. Macmillan And Co., Limited; London.

Meligy, A. M., Ibrahem, H. M. and Othman, E. A. 2014. Communication Centrality in Dynamic Networks Using Time-Ordered Weighted Graph. *International Journal of Computer Network and Information Security,* 6 (12), pp. 21.

Mukkala, K. and Tervo, H. 2013. Air transportation and regional growth: which way does the causality run? *Environment and Planning A,* 45 (6), pp. 1508-1520.

Murata, Y., Nakajima, R., Okamoto, R. and Tamura, R. 2014. Localized knowledge spillovers and patent citations: A distance-based approach. *Review of Economics and Statistics,* 96 (5), pp. 967-985.

Neal, Z., 2012. Creative employment and jet set cities: disentangling causal effects. *Urban Studies,* 49 (12), pp. 2693-2709.

Ni, P., Kamiya, M. and Ding, R., 2017. Global Investment Competitiveness Report 2017/2018: Foreign Investor Perspectives and Policy Implications. The World Bank.

OECD Innovation, 2007. Growth: Rationale for an innovation strategy. *Luxembourg: Statistical Office of the European Communities,* .

Oxford Economics, 2011. Economic Benefits from Air Transport in the US. Available at: https://www.iata.org/policy/Documents/Benefits-of-Aviation-US-2011.pdf .

Parisi, M. L., Schiantarelli, F. and Sembenelli, A. 2006. Productivity, innovation and R&D: Micro evidence for Italy. *European Economic Review*, 50 (8), pp. 2037-2061.

Porter, M., 1990. The competitive advantage of nations. *Harvard Business Review*, (March-April), Available at: http://www.economie.ens.fr/IMG/pdf/porter_1990_-_the_competitive_advantage_of_nations.pdf [Accessed 20 April 2018].

Rogerson, R. J., 1999. Quality of life and city competitiveness. *Urban Studies*, 36 (5-6), pp. 969-985.

Rondinelli, D. A., Johnson Jr, J. H. and Kasarda, J. D. 1998. The changing forces of urban economic development: Globalization and city competitiveness in the 21st century. *Cityscape*, pp. 71-105.

Saxenian, A., 1990. Regional networks and the resurgence of Silicon Valley. *California Management Review*, 33 (1), pp. 89-112.

Schaafsma, M., 2003. Airports and cities in networks. *disP-the Planning Review*, 39 (154), pp. 28-36.

Sokolov-Mladenović, S., Cvetanović, S. and Mladenović, I. 2016. R&D expenditure and economic growth: EU28 evidence for the period 2002–2012. *Economic Research-Ekonomska Istraživanja*, 29 (1), pp. 1005-1020.

Soo, K. T., 2018. Innovation across cities. *Journal of Regional Science*, 58 (2), pp. 295-314.

Taylor, P. J., 2014. Competition and cooperation between cities in globalization. *Urban Competitiveness and Innovation*, pp. 139.

Wall, R., 2009. Netscape: cities and global corporate networks. Erasmus University Rotterdam , Erasmus Research Institute of Management. Available at: http://hdl.handle.net/1765/16013 .

Wall, R. S. and Van der Knaap, G. 2011. Sectoral differentiation and network structure within contemporary worldwide corporate networks. *Economic Geography*, 87 (3), pp. 267-308.

Wang, J., Mo, H., Wang, F. and Jin, F. 2011. Exploring the network structure and nodal centrality of China's air transport network: A complex network approach. *Journal of Transport Geography*, 19 (4), pp. 712-721.

Xibao, L., 2007. A case study on the changes in the innovation capability of China's regions: A concept based on the innovation system [J]. *Management World*, 12 pp. 18-30.

Appendix

Appendix 1 the summary of overall the United States air transportation network analysis result in 2012

Network measurements	Obs	Mean	Std. Dev.	Min	Max
Degree centrality	1,225	516.8314	2816.988	.0007645	43307.85
Betweeness centrality	1,225	.0038819	.0139886	0	.3104869
Closeness centrality	1,225	.0007797	.0001907	.0000118	.0008515

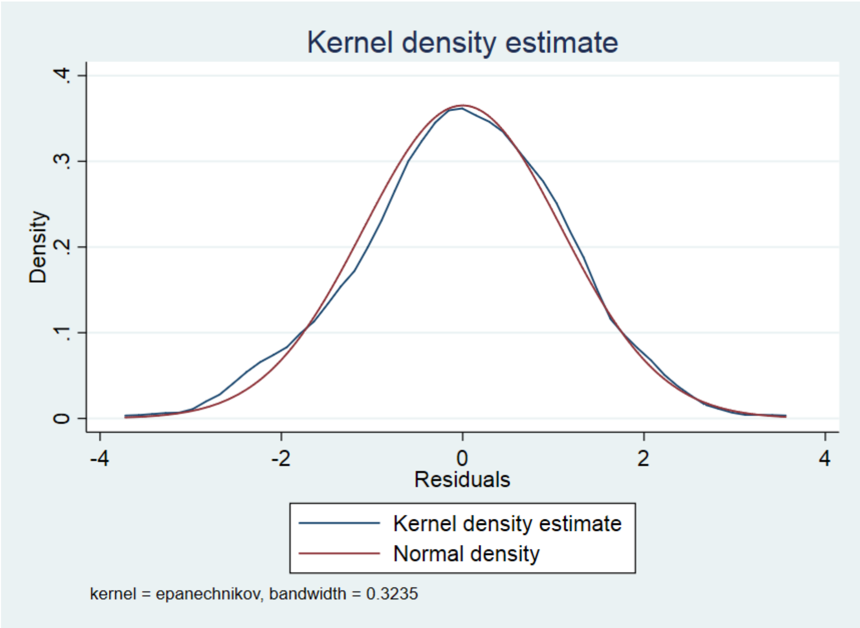
Appendix 2 the multicollinearity test results on OLS regression between R&D establishment and air passenger network

		VIF	
VARIABLES	Model 1	Model 2	Model 3
Air passenger network (degree centrality)	1.46	1.88	3.32
i.MSA with top 100 universities	1.22	1.28	1.30
Per capita income (thousand dollars)	1.20	1.44	1.48
Crime (rate per 1 million people)	1.11	1.11	1.11
i.MSA with top 25 vacation destination	1.34	1.36	1.40
Agglomeration (number of firms per sqmiles)		2.20	4.01
Air passenger network##agglomeration			2.71
Mean VIF	1.27	1.54	2.19

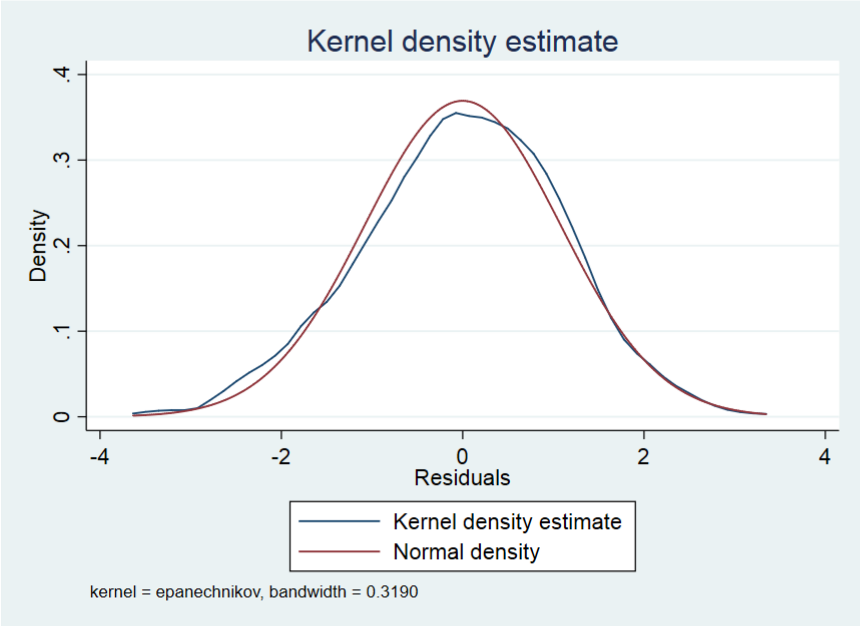
Appendix 3 the multicollinearity test results on OLS regression between patent grants and air passenger network

		VIF	
VARIABLES	Model 4	Model 5	Model 6
Air passenger network (degree centrality)	1.46	1.88	3.32
i.MSA with top 100 universities	1.22	1.28	1.30
Per capita income (thousand dollars)	1.20	1.44	1.48
Crime (rate per 1 million people)	1.11	1.11	1.11
i.MSA with top 25 vacation destination	1.34	1.36	1.40
Agglomeration (number of firms per sqmiles)		2.20	4.01
Air passenger network##agglomeration			2.71
Mean VIF	1.27	1.54	2.19

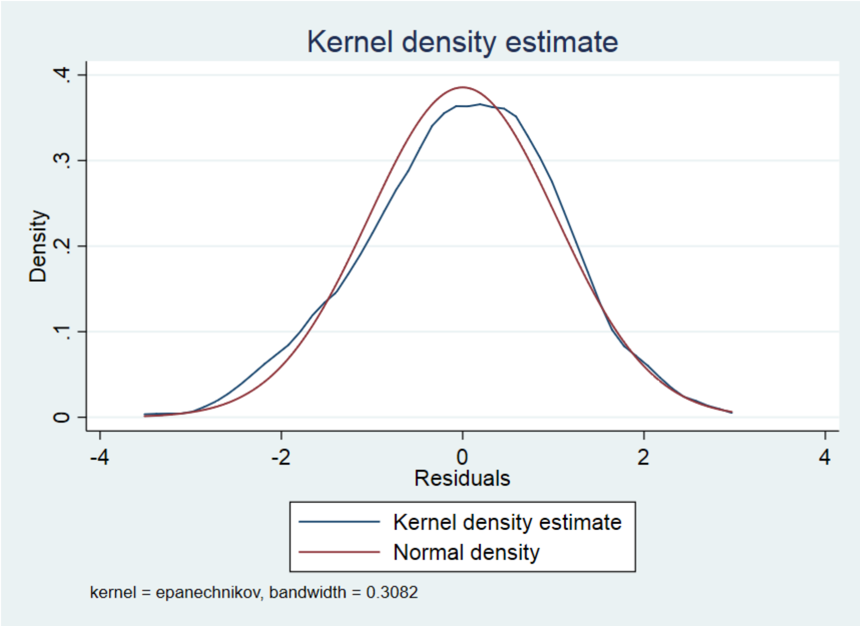
Appendix 4 Kernel density test for OLS regression on model 1



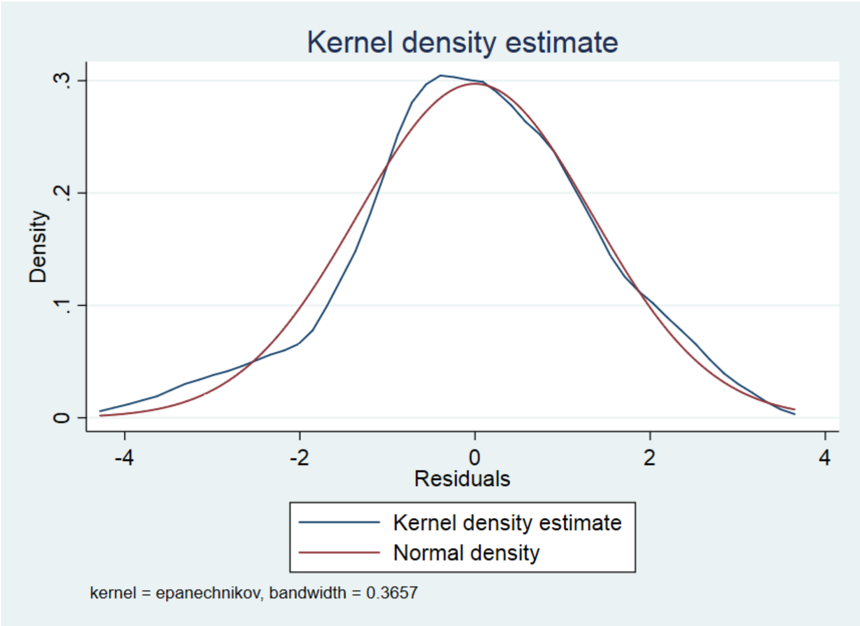
Appendix 5 Kernel density test for OLS regression on model 2



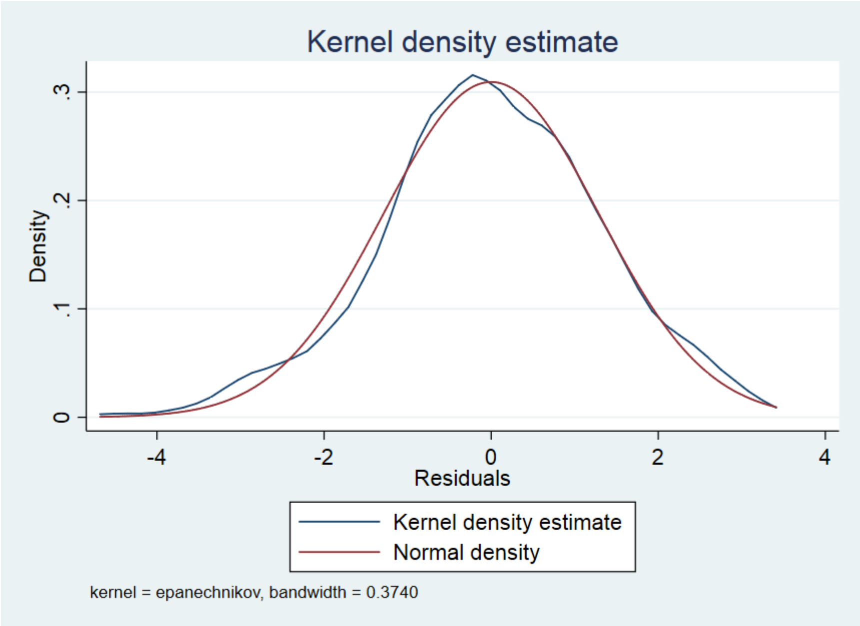
Appendix 6 Kernel density test for OLS regression on model 3



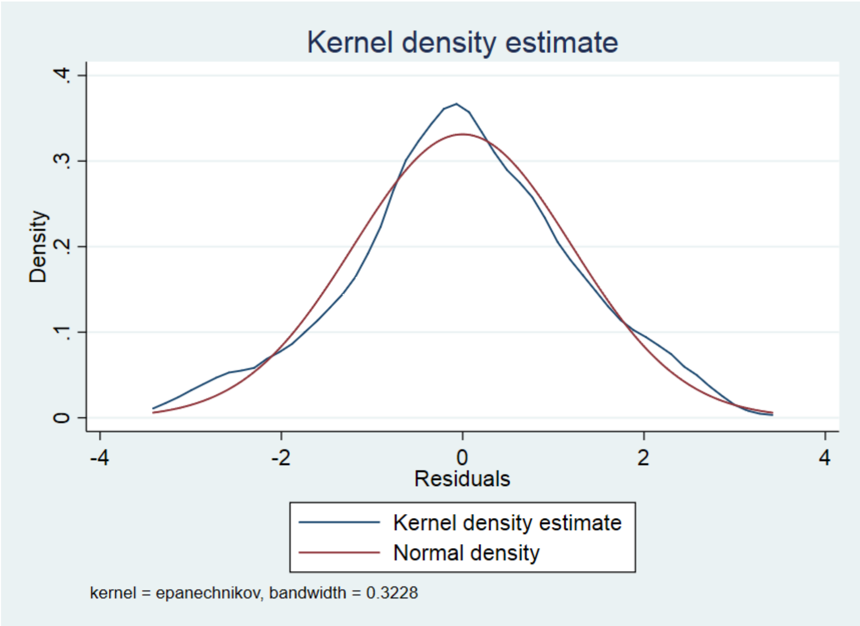
Appendix 7 Kernel density test for OLS regression on model 4



Appendix 8 Kernel density test for OLS regression on model 5



Appendix 9 Kernel density test for OLS regression on model 6



Appendix 10 the result of negative binomial regression on dependent variable R&D Establishment

	Number of R&D Establishment		
VARIABLES	Model 1	Model 2	Model 3
Air passenger network (degree centrality)	0.152***	0.121***	0.194***
	(0.038)	(0.032)	(0.049)
i.MSA with top 100 universities	0.689***	0.606***	0.586***
	(0.192)	(0.186)	(0.175)
Per capita income (thousand dollars)	0.048***	0.032**	0.024*
	(0.015)	(0.015)	(0.014)
Crime (rate per 1 million people)	-0.082	-0.093	-0.099
	(0.065)	(0.067)	(0.065)
i.MSA with top 25 vacation destination	-0.108	-0.058	0.071
	(0.265)	(0.248)	(0.241)
Agglomeration (number of firms per sqmiles)		0.050**	0.078***
		(0.021)	(0.023)
AiNetwork##agglomeration			-0.004***
			(0.001)
Constant	1.099*	1.565**	1.713***
	(0.624)	(0.621)	(0.578)
Observations	252	252	252

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

Appendix 11 the result of negative binomial regression on dependent variable patent grants

	Number of Patent grants		
VARIABLES	Model 4	Model 5	Model 6
Air passenger network (degree centrality)	0.163***	0.124***	0.216***
	(0.039)	(0.032)	(0.051)
i.MSA with top 100 universities	0.905***	0.854***	0.821***
	(0.235)	(0.238)	(0.233)
Per capita income (thousand dollars)	0.074***	0.048**	0.037**
	(0.018)	(0.020)	(0.019)
Crime (rate per 1 million people)	-0.185**	-0.217***	-0.242***
	(0.074)	(0.077)	(0.076)
i.MSA with top 25 vacation destination	-0.321	-0.227	-0.126
	(0.490)	(0.493)	(0.483)
Agglomeration (number of firms per sqmiles)		0.076***	0.116***
		(0.028)	(0.030)
Network##agglomeration			-0.006***
			(0.001)
Constant	2.045***	2.814***	3.076***
	(0.774)	(0.819)	(0.799)
Observations	255	255	255

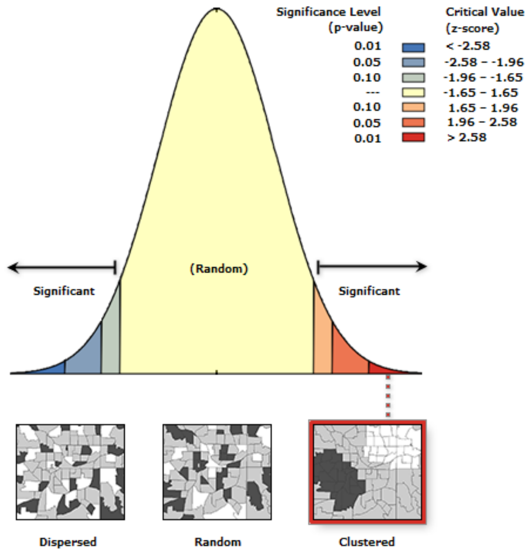
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix 12 Durbin and Wu-Hausman test of 2SLS regresion

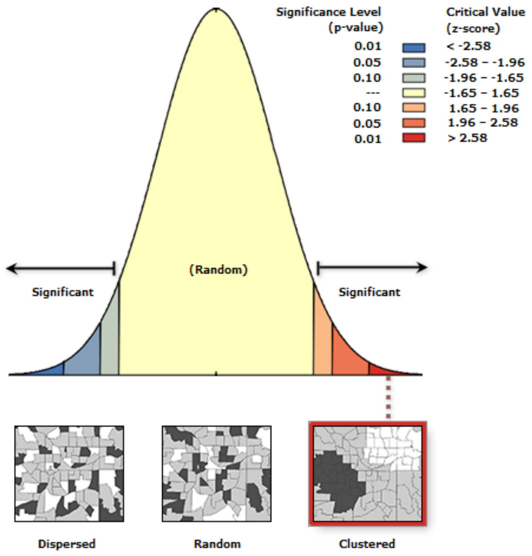
	Dependent Variable: (log)R&D	Dependent Variable: (log)Patent
Durbin (score) chi2 (1)	7.6013 (p=0.0058)	6.69396 (p=0.0097)
Wu-Hausman F(1,246)	7.5578 (p=0.0064)	6.63179 (p=0.0106)

H0: variables are exogenous

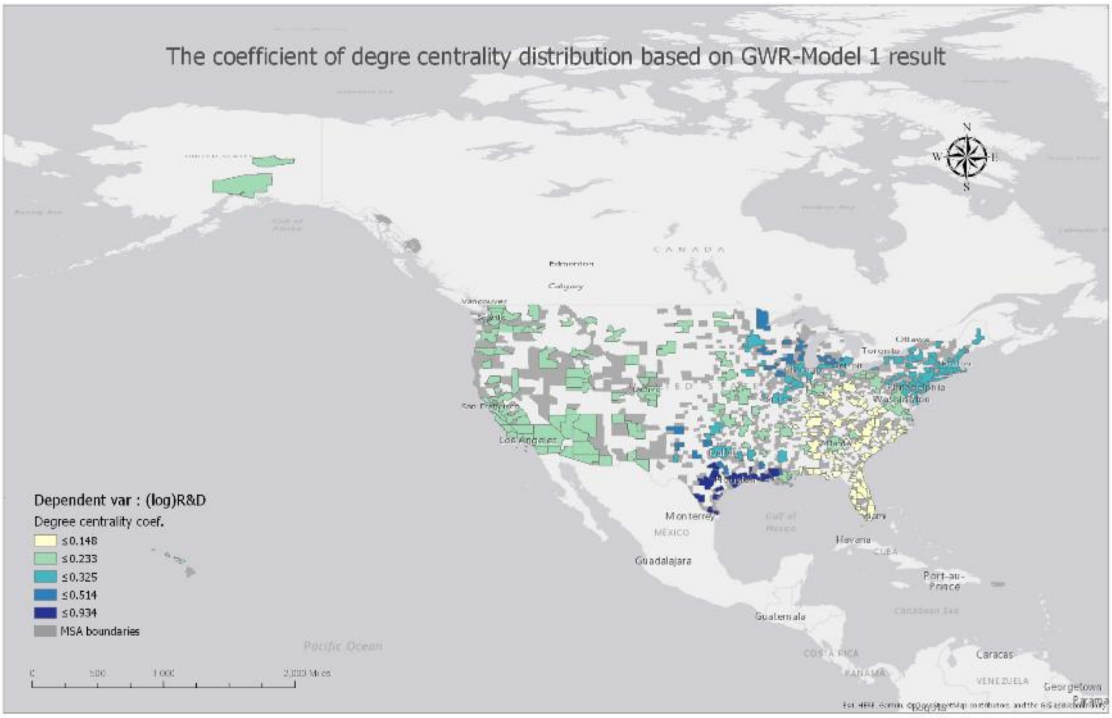
Appendix 13 Moran's I Auto Correlation Test for Local R^2 in GWR-Model 1



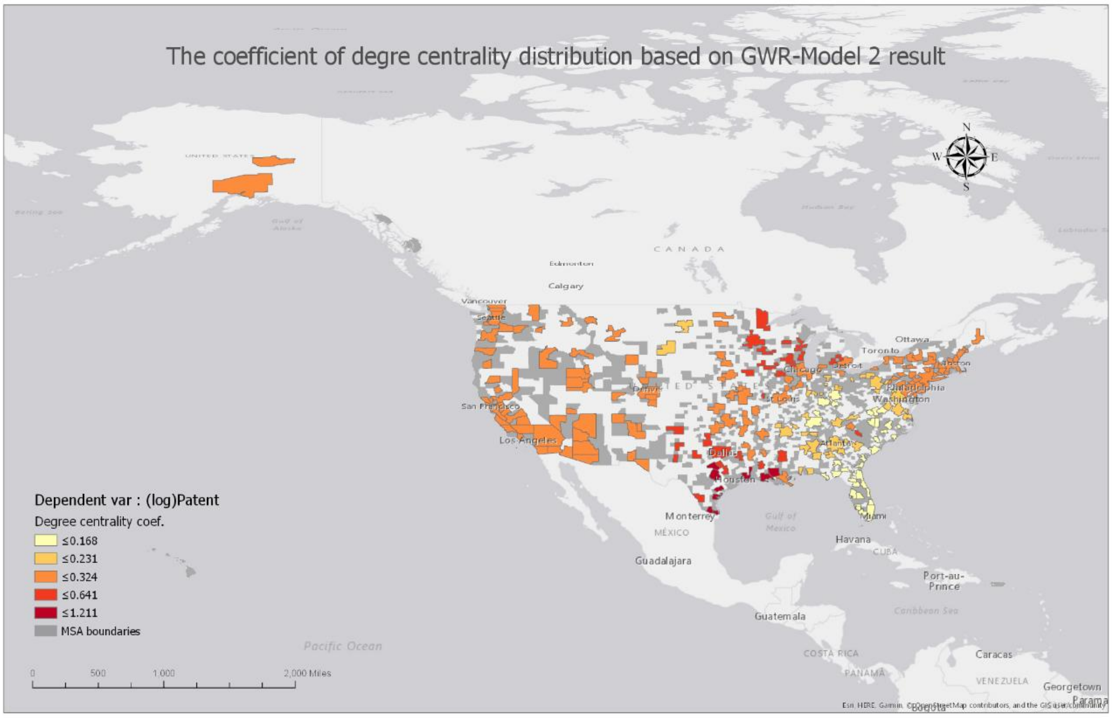
Appendix 14 Moran's I Auto Correlation Test for Local R^2 in GWR-Model 2



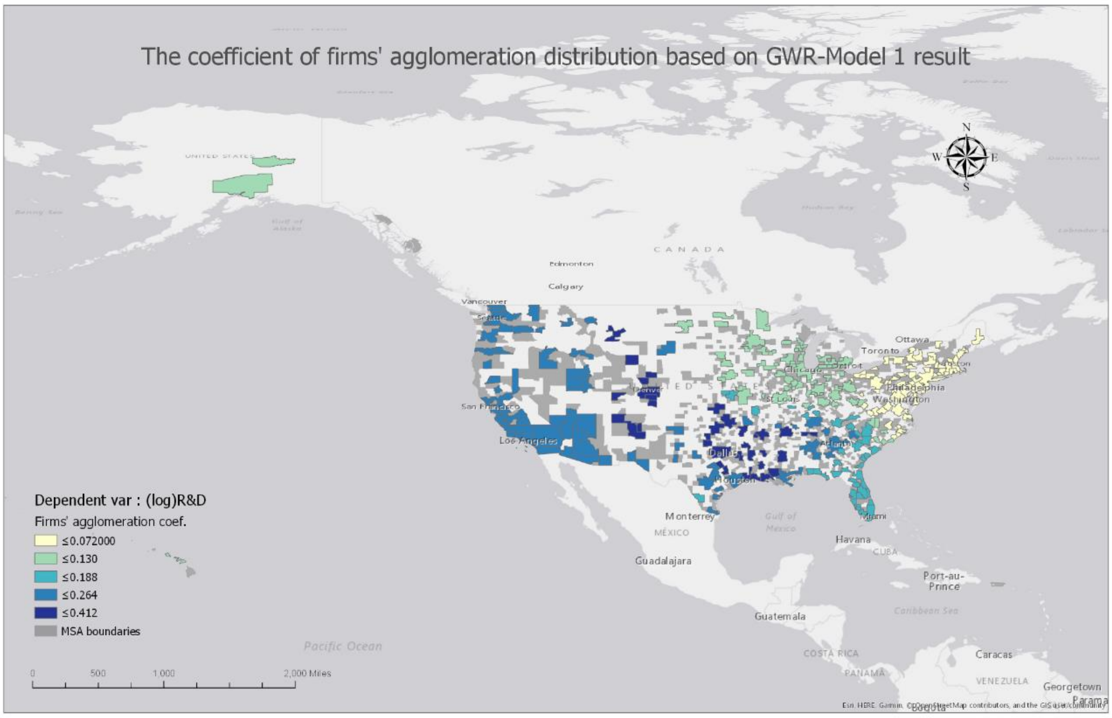
Appendix 15 The β distribution of degree centrality on GWR-Model 1 (dependent variable: logR&D)



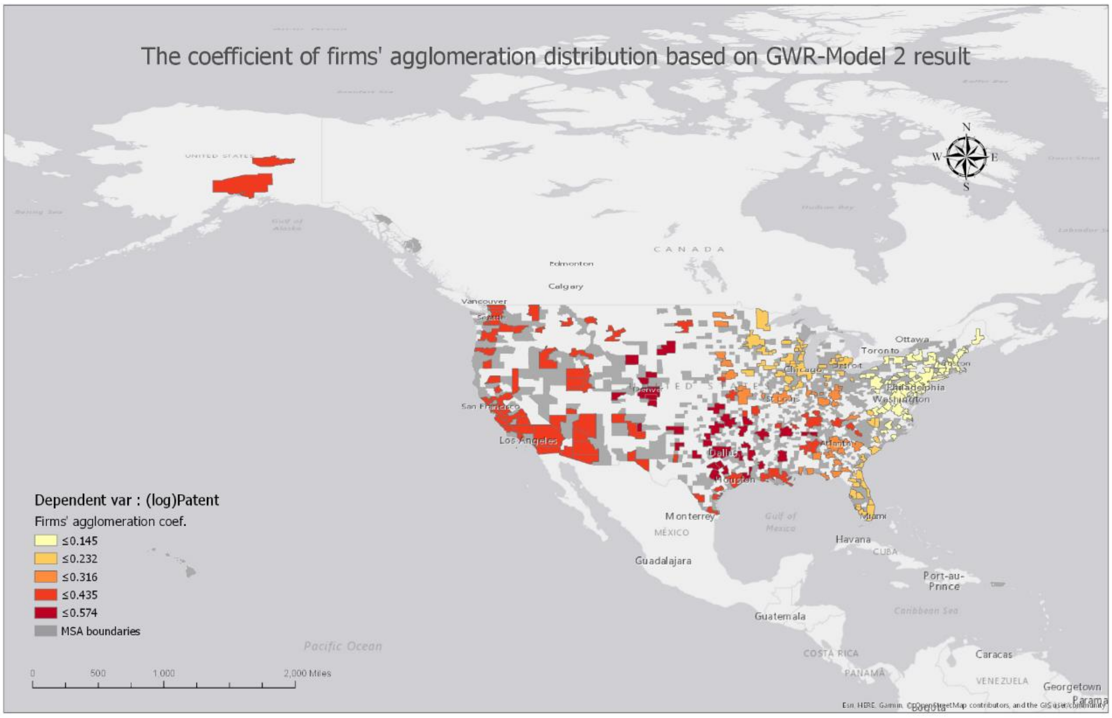
Appendix 16 Appendix 15 The β distribution of degree centrality on GWR-Model 2 (dependent variable: logPatent)



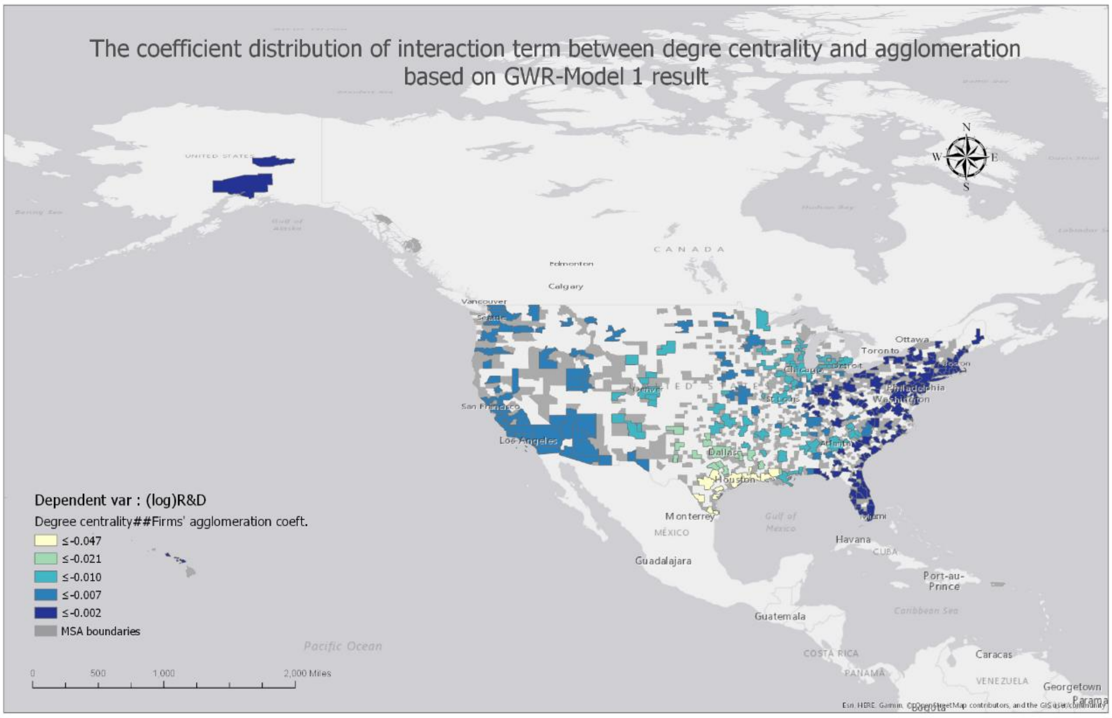
Appendix 17 The β distribution of firms agglomeration on GWR-Model 1 (dependent variable: logR&D)



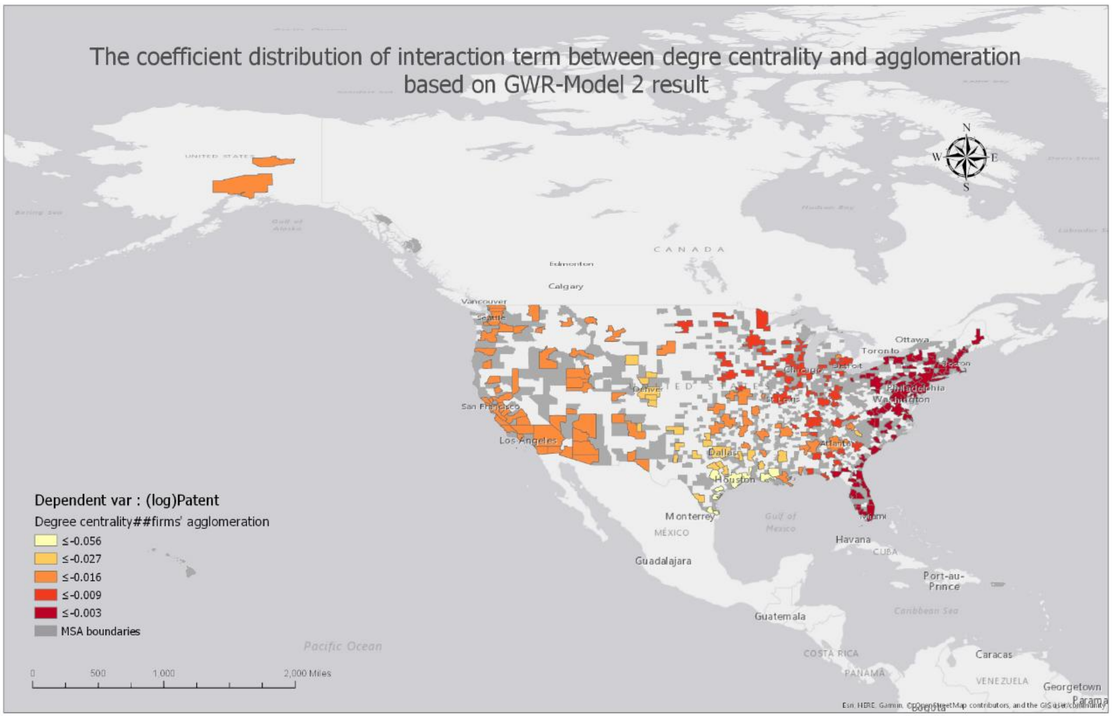
Appendix 18 The β distribution of firms agglomeration on GWR-Model 2 (dependent variable: logPatent)



Appendix 19 The β distribution of interaction term between air network and firms' agglomeration on GWR-Model 1 (dependent variable: logR&D)



Appendix 20 The β distribution of interaction term between air network and firms' agglomeration on GWR-Model 1 (dependent variable: logPatent)



Annex : IHS copyright form

In order to allow the IHS Research Committee to select and publish the best UMD theses, participants need to sign and hand in this copy right form to the course bureau together with their final thesis.

Criteria for publishing:

A summary of 300 to 500 words should be included in the thesis.

The number of pages for the thesis is about 60.

The thesis should be edited.

Please be aware of the length restrictions of the thesis. The Research Committee may choose not to publish very long and badly written theses.

By signing this form you are indicating that you are the sole author(s) of the work and that you have the right to transfer copyright to IHS, except for items cited or quoted in your work that are clearly indicated.

I grant IHS, or its successors, all copyrights to the work listed above, so that IHS may publish the work in *The IHS thesis series*, on the IHS web site, in an electronic publication or in any other medium.

IHS is granted the right to approve reprinting.

The author(s) retain the rights to create derivative works and to distribute the work cited above within the institution that employs the author.

Please note that IHS copyrighted material from *The IHS thesis series* may be reproduced, up to ten copies for educational (excluding course packs purchased by students), non-commercial purposes, providing full acknowledgements and a copyright notice appear on all reproductions.

Thank you for your contribution to IHS.

Date : September 5th 2018

Your Name(s) : Fitri Nurramdani Solehuddin

Your Signature(s) : ____

Please direct this form and all questions regarding this form or IHS copyright policy to:

The Chairman, IHS Research Committee Burg. Oudlaan 50, T-Building 14th floor, 3062 PA Rotterdam, The Netherlands	j.edelenbos@ihs.nl Tel. +31 10 4089851