EMPIRICAL ANALYSES OF REGIONAL INNOVATION AND ECONOMIC GROWTH IN THE UNITED STATES

By

Giri Raj Aryal

A DISSERTATION

Submitted to
Michigan State University
in partial fulfilment of the requirements
for the degree of

Agricultural, Food, and Resource Economics – Doctor of Philosophy

2018

ABSTRACT

EMPIRICAL ANALYSES OF REGIONAL INNOVATION AND ECONOMIC GROWTH IN THE UNITED STATES

By

Giri Raj Aryal

Much of the innovation creation literature is focused on urban firms and areas or relies heavily on data based on these; less studied are rural firms or rural areas in this regard. The goal of this dissertation is to explore the drivers of rural-urban innovation gap and the link between regional innovation and economic growth and propose policies to mitigate regional innovation ecosystem deficiencies and impediments that contribute to the gap.

In my first essay, I analyze heterogeneity in inventiveness across urban and rural counties is using a spatial autoregressive negative binomial regression model, considering spatial spillover effects, creative class population, industry characteristics, human capital, and other regional factors influencing innovation. Results indicate that drivers of invention, namely a college-educated labor force and diversity of high-tech industries are common across all counties types, but urban inventive advantage persists due to agglomeration economies, higher number of universities, and higher shares of high-tech firms, professional services and immigrants. Consistent with the creative class hypothesis, population share of college graduates in creative disciplines also positively contributes to inventive output in urban counties. However, the effects of spatial spillovers and mobile phone technology penetration are stronger for rural counties, suggesting that policies promoting rural centers of innovation, technological diversity, and communication infrastructure in rural counties could help mitigate the urban-rural innovation gap.

My second essay explores the interdependence between regional innovation and economic growth by accounting for possible endogenous relationships among regional innovation, income growth, employment and population. It draws on data for 3,038 counties in the 48 contiguous states of the United States collected from several publicly available sources for 2009-13. Endogeneity tests using instrumental variable regressions show that regional innovation and economic growth have endogenous relationships. Considering the endogeneity and estimating the system of simultaneous equations for regional innovation and economic growth using three stage least squares (3SLS) method, I find that innovation belongs to system of regional growth. Further, reduced form estimates of the 3SLS results suggest that policies promoting regional clusters of high-tech firms and capitalizing on the knowledge potential of the immigrants are likely to reinforce both regional innovation rates and economic growth.

My third essay analyzes the characteristics that potentially influence innovation creation across rural and urban firms employing a survey dataset from 2014 National Survey of Business Competitiveness combined with secondary data reflecting the regional business and innovative environments where these firms operate. The number of patent applications filed by these firms measures their innovation creation, and the paper employs a negative binomial regression estimation for analysis. The findings of this essay show that, after controlling for industry, county and state factors, rural and urban firms differ in their innovation creation characteristics and behaviors, suggesting that urban firms capitalize on their resources better than rural firms. Other major findings of the essay provide evidence that (i) for rural firms, the influence of university R&D is relevant to innovation creation, but their perception of university provided information is not significant and (ii) rural firms that are willing to try, but fail, in terms of innovation creation have a slight advantage over other rural firms less willing to take on the risk.

To
my parents, Prem Nath and Kaushila Aryal
my brother, Sudarshan Aryal
my sisters, Laxmi and Saraswati Aryal

ACKNOWLEDGEMENTS

First and foremost, I would like to express my profound gratitude for the incredible mentorship, guidance, inspiration, and motivation I received from my major advisor Dr. Satish Vasudev Joshi and my research supervisor Dr. John Thomas Mann II, who also jointly funded the most part of my PhD study at Michigan State. Thank you for believing in me since the very beginning and insisting that I made progress even when I hit roadblocks as part of any long-term research. Also, sincere thanks to the rest of my advisory committee – Dr. Scott Loveridge, who also funded a part of my PhD study, and Dr. Mark Skidmore for precious comments on my research and technical as well as personal advice throughout my PhD career.

I also extend my gratitude to the faculty and staff of the Department of Agricultural, Food and Resource Economics (AFRE), Michigan State University, especially the Graduate Secretaries, Ms. Debbie Conway and Ashleigh Booth, Travel Arranger, Nancy Creed, and the Graduate Chairs Dr. Scott Swinton and Dr. Robert Myers for consistent support and assistance while navigating the PhD study.

I am forever indebted to my AFRE colleagues, Jina Yu, John Olwande, Jong-Woo Kim, Jungmin Lim, Mayuko Kondo, Miguel Castro, Nahid Sattar, among others, for the lasting bonds created through sharing struggles and common experience. I am also grateful to all my friends in East Lansing for their support by celebrating my successes and commiserating the failures during my doctoral program.

Last but most importantly, I would like to a acknowledge my family who encouraged me to accept my offer of admission into the PhD program at MSU and supported me throughout the whole program. Specially, I want to thank my remarkable parents, Prem Nath and Kaushila

Aryal, for all the sacrifices they made for my education and provided forever inspiration to work hard and follow fearlessly my dreams. I also wish to acknowledge my brothers, especially Sudarshan Aryal, who always encouraged and supported me to hold onto my aspirations throughout the tough times of my academic and personal life. Finally, I would like to thank my dearest and caring sisters, Laxmi and Saraswati Aryal, for their unending love that relentlessly helped me move forward through the challenging situations and cherish every small accomplishment.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

ACS	American Community Survey
BEA	Bureau of Economic Analysis
BLS	Bureau of Labor Statistics
CBP	Community Business Patterns
ERS	Economic Research Service
FCC	Federal Communication Commission
KPF	Knowledge Production Function
NAICS	North American Industry Classification System
NSBC	National Survey of Business Competitiveness
NSF	National Science Foundation
NUTS	Nomenclature of Territorial Units for Statistics
OECD	Organization for Economic Co-operation and Development
R&D	Research and Development
RENB	Random Effects Negative Binomial
RKPF	Regional Knowledge Production Function
SBA	Small Business Administration
SBIR	Small Business Innovation Research
SD	Standard Deviation
USDA	United States Department of Agriculture
USPTO	United States Patents and Trademark Office

ESSAY 1: DRIVERS OF DIFFERENCES IN INVENTIVENESS ACROSS URBAN AND RURAL REGIONS

1.1. Introduction

Innovation is central to economic competitiveness. Prior research identified and analyzed economic and non-economic factors driving innovation, and variations in innovation-related outputs across time and regions. Key drivers include population densities, critical mass of educated and high-skilled employees, research and development (R&D) expenditures by universities and private industries, innovation and communication infrastructure, and network externalities (Acs, et al., 2002; Anselin, Varga, & Acs, 1997; Audretsch & Feldman, 2004; Barkley et al., 2006; Carlino, et al., 2007). It is no surprise that urban regions are more conducive to innovation due to scale economies, network externalities and knowledge spillovers, i.e., the agglomeration effect (Carlino, et al., 2001; Carlino et al., 2007; Feldman & Florida, 1994). When considering broader regions, however, questions about appropriate geographic units arise since the benefits of knowledge spillovers attenuate with distance (Rosenthal & Strange, 2004).

Many prior studies focused on larger geographic units, such as state or metropolitan statistical area (MSA), which are likely to "… obscure the spatial (innovation) processes that occur within a region or across its regional boundaries" (Feldman & Florida, 1994, p. 216). Further, evidence suggests that spillover effects are likely more pronounced using smaller geographic units such as the county (Jaffe, et al., 1993). On the other hand, more granular level studies may only consider smaller regions limiting the analysis of knowledge spillovers (Monchuk & Miranowski, 2010; Stephens et al., 2013), or do not explicitly analyze rural-urban differences in the rates of innovation (Zheng & Slaper, 2016). Distinguishing innovation rates by county types may be relevant as urban or proximate to the urban counties fare better in terms of innovation and economic growth (Monchuk & Miranowski, 2010; McGranahan, et al., 2010;

Henderson & Executive, 2007; Henderson & Weiler, 2010; Henderson & Abraham, 2004; Stephens, et al., 2013).

To address these gaps, I empirically analyze rural-urban gaps in innovation, focusing on differences in spillover effects and drivers of innovation among rural and urban counties in the U.S. This study contributes to the literature by analyzing regional heterogeneity of inventiveness, measured as patents per capita of inventive class population, across urban, metro-adjacent rural, and rural remote areas, considering the spatial spillover effects, creative class population, industry characteristics, human capital, communication access, and other state-level factors influencing innovation. I use a comprehensive county-level data set spanning the entire U.S., and empirically account for the spatial spillover (and spatial error correlation) and count nature of the dependent variable, by estimating a spatial autoregressive negative binomial regression model with county-level random effects.

I find that the spillover effect of regional inventiveness is stronger for rural counties than for urban counties, implying externalities arising from innovative climate in their neighboring areas have larger influence in rural counties. Second, 'the creative class hypothesis,' that the population share of college graduates in creative disciplines positively influences invention rates, is empirically supported only in urban areas but not in rural areas. This points to another source of rural disadvantage. I also find that the industry mix, in terms of professional services and manufacturing, positively influences inventiveness only in urban and metro-adjacent areas. Similarly, the influence from the presence of 4-year colleges and universities, share of high-tech firms, and new immigrants were statistically significant only for urban areas, likely reflecting the benefits of agglomeration economies. However, inventiveness in rural areas is positively associated with higher levels of mobile/cellular access compared to broadband services (via cable or landline) in urban areas, suggesting that cellular services are substitute sources of knowledge and information in rural areas. Additionally, the share of college-educated labor force and the diversity of high-tech industries influence inventiveness across all regions (urban, metro-adjacent rural, and rural remote). Finally, I do not find significant associations between tax burden, unemployment rate, and state-level venture capital on patenting rates in my study.

1.2. Review of Literature

Innovation is a key driver of economic growth and regional development as the manifestation of new ideas and knowledge (e.g., in improved products and processes) provide entrepreneurial opportunities leading to regional prosperity (Acs, et al., 2002; Feldman & Florida, 1994). Earlier research focused on the firm or industry unit of analysis, and found innovation, measured by patents, was positively associated with higher productivity and profit (Bound et al., 1984; Griliches, 1990; Hall et al., 1986; Hausman et al., 1984; Pakes & Griliches, 1980; Scherer, 1965a, 1965b; Cincera; 1997). Later research extended the Griliches (1979) knowledge production function (KPF) approach to study geographically mediated knowledge spillovers, for example, between universities and the private sector (Jaffe, 1989; Audretsch & Feldman, 2004). Increasingly, regions came to be considered more appropriate units for analyzing the innovation process as my understanding of knowledge spillovers and agglomeration economies across firms and industries evolved (Audretsch & Feldman, 1996; Rosenthal & Strange, 2004, Florida, et al., 2016). Within the regional dimension of innovation, large cities and metro regions received greater scholarly attention since the co-location of firms and knowledge workers in clusters of similar industries were assumed to facilitate spillovers of tacit knowledge due to proximity (Glaeser et al., 1992; Henderson, 2003).

A major challenge for researchers analyzing innovation is identifying appropriate measures of the multi-faceted innovation process (Acs et al., 2002; Cameron, 1996; Cohen & Levin, 1989; Mann & Shideler, 2015). Typically used proxies to capture the different stages of innovation include: R&D expenditures for inputs, number of patents for invention output, and new product introductions for final innovative outputs (Aghion & Howitt, 1990; Acs & Audretsch, 2005a). At the same time, no single proxy can adequately capture the multi-dimensional and stochastic concept of innovation (Mann & Shideler, 2015). For example, R&D expenditures are often directed toward imitation or technology adoption, in addition to generating inventions/patents (Mansfield, 1984; Kleinknecht,1987; Kleinknecht & Verspagen,1989). Reliable and comprehensive data on direct measures of innovative outputs such as new product or service announcements are difficult to obtain (Acs, et al., 2002; Huang et al., 2010). Patent statistics as innovation proxies are criticized because neither all inventions are patented nor do all patents lead to commercialized final products (Griliches, 1990; Nagaokaet al., 2010). Additionally, the implicit assumption of homogeneity of any chosen proxy measure in terms of relative contributions to actual technological change or economic value generated is inconsistent with reality (Acs & Audretsch, 2005b; Cohen & Levin, 1989; Pakes & Griliches, 1980). In fact, Capello & Lenzi (2014), in their analyses of the nexus between innovation and economic growth in 27 European counties, make the distinction between invention (e.g., patents) and innovation (e.g., commercialized output), and argue that less knowledge-intensive regions can achieve economic growth, as some regions may benefit more from new knowledge creation while other may benefit more from innovation commercialization.

Despite these limitations, patents remain a popular output indicator of the innovation process due, in part, to data availability and their consistent correlations with other proxies

(Autant-Bernard, 2001; Acs, et al., 2002; Czarnitzki et al., 2009; Pakes & Griliches 1980; Parent & Lesage, 2008). For example, Acs et al. (2002) found that patent applications performed similarly to new product announcements. Parker et al. (2017) compared 40 potential measures of innovation and found that patent applications were statistically similar in performance to the other 39 measures.

Another related challenge, especially when analyzing relative innovation performance of regions, is the choice of the appropriate scaling when estimating innovation rates. Wojan et al. (2015) highlight concerns about using patents per capita which assumes an inaccurate level of homogeneity across regions. For example, retirement communities or tourist towns cannot reasonably have the same innovation potential as equally populated technology/industrial cities or university towns. They show that urban areas appear to be more inventive when patents are scaled per capita, but patenting rates scaled by the inventor class (science, engineering and technical professionals) are more equally distributed across urban and rural regions[1]. Along a similar vein, Florida (2002) argued that the creative class, consisting of artists, musicians, architects, etc., also contributes directly and indirectly to innovation by allowing more creative collaborations and technology adaptation to meet creative, non-technical professional needs. A number of regional scientists since then have explored the relationship between entrepreneurship and innovation production and the creative class.

The rural-urban innovation gap can be explained in terms of the drivers of urban innovation, specifically that urban firms have better access to innovation inputs such as human capital, physical capital, knowledge stock, infrastructure, support services, and output markets (Barkley et al., 2006; Henderson & Weiler, 2010; McGranahan, et al., 2010; Monchuk &

[1] I refer the interested reader in a more detailed discussion and presentation of patenting rates across rural and urban regions to Wojan, Dotzel, & Low (2015), who include a number of helpful and informative figures.

Miranowski, 2010; Orlando & Verba, 2005). Small populations and thin markets limit the ability of rural firms to capitalize on economies of scale. Further, higher population density and the concentration and diversity of industries provide more opportunities for communication between innovators. This leads to more synergistic knowledge spillovers and agglomeration economies in urban settings, the benefits of which are difficult, if not impossible, to replicate in rural areas. Rural regions, however, are also not homogeneous. Empirical analysis suggests that spillovers arising from entrepreneurship and innovation creation are stronger in counties that have denser population and are more proximate to metro counties (Stephens, et al., 2013; Henderson & Weiler, 2010; Monchuk & Miranowski, 2010; Henderson & Executive, 2007; Feser & Isserman, 2006). Other studies posit that rural entrepreneurship is driven more by necessity than by innovative opportunity, which often leads to abandonment when better paying jobs arise (Acs, 2006; Henderson, 2002). Further, some of the behavioral factors analyzed include rural ownership characteristics such as multi-generational ownership and risk aversion (Renski & Wallace, 2012; Markley, 2001), and such factors may be less attractive to equity and venture capital investors.

Although the extant literature taken together, identifies a large set of potential influencing factors driving the rural-urban innovation gap, individual studies suffer from one or more of the following limitations: limited geographical coverage focusing mostly on urban areas or sub-regions; relatively large units of analyses (states or metropolitan areas); confounded innovation output measures due to normalization by population; and inadequate consideration of the count nature of patents, spatial spillover effects, correlated spatial errors and potential creative class contribution in model specifications. My study attempts to address these limitations by building on a recent working paper by Zheng & Slaper (2016). Using a similar comprehensive county-level dataset, I analyze spillovers using a distance decay function. However, the focus of Zheng & Slaper (2016) is mainly on spillovers from University R&D expenditures and the sensitivity of these spillovers to distance. Instead, I turn my attention to analyzing urban-rural gaps in invention rates, considering three county types, urban, metro-adjacent and remote rural. I normalize my output variable by the inventive class population and control for the potential creative class contribution. Finally, my econometric approach controls for the count nature of the dependent variable and spatial correlation; whereas, Zheng & Slaper (2016) relied on linear ordinary least square (OLS) estimations.

1.3. Empirical Methods

The regional knowledge production function (RKPF) is central to a number of empirical studies of regional innovation and knowledge spillovers and can include region-specific factors that may influence regional innovative outputs (Charlot et al., 2014; Buesaet al., 2010; Ponds et al., 2009; Varga, 2000). I assess the rural-urban differences in innovation using the extended RKPF framework where the dependent variable, a measure of inventive output in rural and urban US counties, is modeled as a function of inventive inputs, county-level regional characteristics, and state-level fixed-effects. My dependent variable, inventions per inventive class, follows the spirit of Wojan et al. (2015) and is operationalized as the number of patent applications originating in a county normalized by the inventive class population (where inventive class is measured as the number of degree holders in science and engineering fields excluding social sciences). The empirical model also includes indicators classifying counties as metro, metro-adjacent rural, and remote rural, to explore rural-urban differences, as well as state and temporal fixed-effects.

The RKPF model using patent applications per 1,000 inventive class population makes OLS assumptions inappropriate due to the count nature of patent applications (Hausman et al., 1984). Under these conditions, OLS estimates are likely to be biased and inconsistent (Greene, 2003). On the other hand, count models are often estimated using a Poisson distribution. However, the conditional mean and variance of Poisson models are assumed equal. Thus, when the dependent variable is over-dispersed, this assumption is violated leading to underestimated standard errors of coefficient estimates and spuriously high statistical significance (Hilbe, 2011). Regional patent data are essentially the counts and are right skewed with large portions of probability mass centered around zero.[2] To address this, I use a negative-binomial estimation procedure, as it can account for such over-dispersion and skewness by allowing the variance to be different than the mean (Hilbe, 2011).

Following Hausman, et al., (1984), Hall et al., (1986), and Griliches (1990), I also employ a county-level random effects model instead of a fixed-effects model. This helps address the large number of counties with zero patent output during the study period and relatively short panel data. A fixed-effect regression model excludes these counties with time-invariant zero patents from analysis, which may introduce potential sample selection problems (Hall et al., 1986). My negative binomial regression model with county-level random-effects takes the form:

$$\log(Pat_{it}) = \boldsymbol{x}'_{it}\boldsymbol{\beta} + \boldsymbol{z}'_{it}\boldsymbol{\gamma} + v_{it} \text{ , and} \qquad v_{it} = (\alpha_{it} + u_{it}) \tag{1}$$

where Pat_{it} are annual utility patent applications per inventive population in county i in year t; x_{it} represents the vector of innovation inputs; and z_{it} represents the vector of other relevant regional

[2] The smaller the geographical units of observations are, the higher number of zero patent observations are likely. In this county-level study as well, nearly two-thirds 64%, 37%, 77%, and 85% of the probability mass of the dependent variable (patent applications per inventive class) is centered at zero for combined, metro, metro-adjacent rural, and remote rural data sets respectively. I do not include the figure on the distribution of patents to save space for remaining analysis, but it is available from authors upon request.

factors. The county-level random effects parameter α_i is assumed to follow a beta distribution (Hilbe, 2011).

Regional patenting has been found to exhibit spatial dependence (Autant-Bernard, 2001; Parent & Lesage, 2008; Florida, 2014), that is, inventive activities in a region may have spillover effects on patenting rates in neighboring regions, and such spillover effects are likely to be more prevalent with more granular geographical units of analyses such as counties. These spillovers can arise from increased proximity, mobility and interaction possibilities, and shared infrastructure and amenities that create an innovative climate. Hence, I hypothesize that innovation rates measured by patenting rates in one county influences patenting rates in its neighboring counties. I include a spatially lagged dependent variable to help estimate these spillover effects. Therefore, my estimated empirical model takes the form of the spatial autoregressive (also known as "mixed-regressive" Anselin, 1988) negative binomial model as shown in equation (2).[3] This model also accounts for spatial error correlation and is specified as:

$$\log(Pat_{it}) = \rho \boldsymbol{W_t} Pat_{it} + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \boldsymbol{z}'_{it}\boldsymbol{\gamma} + v_{it} \qquad (2)$$

where $W_t Pat_{it}$ is the spatially lagged dependent variable, and ρ is the spillover, or spatial dependence, coefficient to be estimated. When the neighborhood of each county does not change, as in my application, W_t is identical each year. Thus, for any year W= w,

[3] The spatial spillover effects arising from mobility of human capital, R&D activities of colleges and universities, and the network of high-tech firms in surrounding regions may also play an important role in determining regional inventiveness. Recognizing this possibility, I initially considered a Spatial Durbin Model (SDM) where spillover effects of the independent variables, namely the high-tech variety, the number of 4-year colleges and universities, and the share of bachelor's or higher degree holders, are included with the spatially lagged dependent variable in equation (2). However, I found that the spatially lagged independent variables displayed a very high degree of collinearity. Instead, I decided to employ an SAR model where the aggregate spatial spillover effects of the dependent variable (inventive output) are modeled. Thus, my model assumes that the potential spillover effects from neighboring counties are adequately captured by their respective patenting rates.

$$\boldsymbol{W} = \begin{bmatrix} 0 & w_{1,2} & w_{1,3} & \dots & w_{1,n} \\ w_{2,1} & 0 & w_{2,3} & \dots & w_{2,n} \\ w_{3,1} & w_{3,2} & 0 & \dots & w_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{n,1} & w_{n,2} & w_{n,3} & \dots & 0 \end{bmatrix}, \text{ and } \boldsymbol{WPat_i} = \begin{pmatrix} \sum_{j=1}^{n} w_{1,j} Pat_{jt} \\ \sum_{j=1}^{n} w_{2,j} Pat_{jt} \\ \sum_{j=1}^{n} w_{3,j} Pat_{jt} \\ \vdots \\ \sum_{j=}^{n} w_{n,j} Pat_{jt} \end{pmatrix}$$

is a nx1 vector whose j^{th} element is a scalar resulting from a linear combination (weighted average) of patenting rates per inventive class in counties neighboring to county *i*. Weights take their values as:

$$w_{i,j} = \begin{cases} \frac{1}{\sum_{j=1}^{n} d_{i,j}} & if\ d_{i,j} < d\ miles, i \neq j \\ 0, & Otherwise \end{cases} \qquad \text{for all } i\text{=1 to n}$$

where $d_{i,j}$ is the geographical distance between centroids of counties *i* and j, and *d* is the threshold distance beyond which spatial dependence in terms of patenting rate is assumed to be zero. I empirically estimate equation (2) using maximum likelihood estimation procedures (Hilbe 2011), with *xtnbreg* command with *random effects* option in Stata® software.

1.4. Data and Variables

My county-level dataset comprises several secondary sources and includes the 48 contiguous states in the U.S. for the period 2009-13.[4] Table 1 lists key variables and their data sources. My dependent variable is patent applications per inventive class population, i.e. the number of college graduates in science and engineering in the county. Additionally, I classified counties into three groups using the 2013 USDA Rural-Urban Continuum Codes (RUCC). I classified counties with RUCC codes of 1, 2, and 3 as metro counties, 4, 6, and 8 as metro-adjacent rural counties, and 5, 7, and 9 as remote-rural counties.

[4] My sample includes 2833 counties as some were dropped due to missing observations across the different secondary data used.

Table 1.1 Variables Definition and Data Source

Variables	Description	Source	Year
Patents per inventive class	Utility patent applications per 1,000 four-year or higher college degree holders in selected S&E fields	USPTO, ACS	2009-13; 2013
Splag patents per inventive class	Spatial lag of patents per inventive class with distance decaying within 50 miles from county centroids	USPTO, ACS	2009-13; 2013
High-tech variety	Number of 3-digit NAICS high-tech industries	CBP	2009-13
High-tech share	Share of high-tech establishments in total business establishments	CBP	2009-13
Universities/colleges	Number of four-year or higher college degree awarding institutions in a county	NSF	2009-13
College plus education	Share of 25 years and older people with bachelor's or higher degree in county population	ACS	2009-13
Arts share	Share of four-year or higher college degree holders in selected Arts fields in total population	ACS	2013
Foreign-born-non-citizen population	Share of foreign-born people who are "not US citizens" in total population	ACS	2009-13
Unemployment change	Change in current year unemployment rate from previous year's rate	BLS	2009-13
Tax burden	Percent of personal income paid in tax	Census of govts.	2007, 2012
High-speed broadband penetration	Dummy (=1(high), if at least 60% of households had high-speed broadband connection; 0(low) otherwise)	FCC	2009-13
Cellphone service penetration	Dummy (=1 (high), if number of cellphone service providers is one s.d. above sample mean; 0(low) otherwise)	FCC	2009-13
Manufacturing intensity	Dummy (=1 (high), if share of population 16 years and older employed in manufacturing industries is one s.d. above sample mean; 0(low) otherwise)	ACS	2009-13
Professionals service intensity	Dummy (=1(high), if share of population 16 years and older employed in professional service industries is one s.d. above sample mean; 0(low) otherwise)	ACS	2009-13
Average venture capital	Average venture capital financing per business establishment, at state level	NSF, CBP	2009-13
County types	Classification of counties based on urban population and commuting patterns	ERS	2003

ACS= American Community Survey: Census Bureau; BLS= Bureau of Labor Statistics; CBP= County Business Patterns; ERS= Economic Research Service-USDA; FCC= Federal Communications Commission; NSF= National Science Foundation; USPTO= U.S. Patents and Trademarks Office

Prior literature shows that human capital is strongly associated with invention rates (Charlot et al., 2014; Buesa et al., 2010; Ponds et al., 2009; Varga, 2000). I use the share of people 25 years or older with a bachelor's or higher in the county population as my human capital measure, and I expect it to positively influence the regional rates of innovation. According to the "creative class hypothesis" (Florida 2002, 2014; McGranahan & Wojan, 2007) individuals in creative occupations including artists and designers positively influence invention rates. Thus, I include the population share of college graduates in selected arts fields and hypothesized it to have a positive coefficient.

Academic institutions act as centers of research, expertise and knowledge-based activities, and train highly-skilled labor force that facilitate inventive activities in other firms including small firms (Acs et al., 1994). Further, universities are increasingly encouraging patenting by their faculty members (Czarnitzki et al., 2009). I include the number of private and public 4-year colleges and universities as a control variable. I hypothesized it to have a positive coefficient indicating positive influence on patenting rates.

To provide additional controls for the entrepreneurial and innovative environment in a county, I include variables for the share of high-tech establishments, the variety of high-tech industries represented, manufacturing intensity, professional service intensity, and venture capital per firm. Note that I limit the industry focus to manufacturing and professional services, as these two industries account for a high level of patenting, and innovation rates in rural and urban areas were found to be more similar in manufacturing intensive areas (Wojan & Parker, 2017). The high-tech variety variable is based on NSF designated high-tech firms, and is the number of four-digit and six-digit high-tech NAICS industries (out of the maximum 45) operating in the county (National Science Foundation, 2017). The share of high-tech establishments is calculated as the total number of these firms divided by the total number of all establishments in a county. Both High-tech variety and share of high-tech establishments are expected to positively influence innovation due to synergistic cross-fertilization of ideas across related and growing industries. The manufacturing intensity and professional service intensity variables are binary variables which are coded as 1 if the shares of the potential labor force (16 years and older) employed in these sectors is at least one standard deviation above the sample mean (based on the combined geography types, metro, metro-adjacent, and remote rural) or "0" otherwise. Both manufacturing intensity and professional service intensity are expected to positively influence innovation production. I use state-level venture capital financing data from the NSF as a proxy for private investment, since private investment data are not readily available at the county level. The variable is normalized as venture capital investment per firm in thousands of dollars, and it is expected to be positively associated with patenting rates. Additionally, a growing body of literature examines the influence of immigrants on innovation production (Kerr, 2013; Kerr & Lincoln 2010; Niebuhr, 2010). I include the variable Share foreign-born non-citizen population, as studies find that it is mainly recent immigrants that positively influence innovation creation.

Communication infrastructure facilitates innovation and may be especially relevant for rural inventors where opportunities for face-to-face communications are less frequent with other innovators (Conley & Whitacre 2016). I use two indicator variables, one for high-speed broadband penetration and another for cellphone/mobile service penetration drawing on Federal Communication Commission (FCC) data. Following Conley & Whitacre (2016), high-speed broadband penetration is coded as "1" if more than 60% of households in a county have a high-speed connection and "0" otherwise. The second variable, cellphone service penetration which has not been used on the prior literature to the best of my knowledge, is coded as “1” if the county penetration is more than one standard deviation above the sample mean (by county type) and “0” otherwise. Although broadband access is more likely ubiquitous in metro areas, I hypothesize a positive association between invention rates and broadband access and cellphone penetration, especially in rural counties.

A variable measuring local total tax burden, defined as ratio of per capita total local tax to per capita personal income, is included to assess the impact of taxes on innovation production in a region. The tax burden data are from the Census of Governments for the years 2007 and 2012.[5] While higher levels of local government services (e.g. education, roads, law enforcement, parks, etc.) are expected to facilitate innovation, it has also been argued that higher taxes inhibit innovation by reducing private resources and incentives for innovation efforts (Bartik 1991; Mukherjee et al., 2017). I conjecture that the facilitation effect of public services will dominate the negative effects, and hypothesize a positive association between tax burden and rates of inventive outputs.

Finally, I use the spatial lag of the dependent variable (derived from the spatial weighting method described in the methods section) to examine the spillover effects of the inventive outputs in neighboring counties. I use the threshold of 50 miles[6] so that the spillover is assumed to occur across county boundaries if the county centroids are located within the distance. Given the discussion of prior literature above, I anticipate this measure to be positively associated with the innovation production in a county.

[5] Since I have tax burden data only for 2007 and 2012, I used the 2007 tax burden for years 2009-2011, and the 2012 tax burden for years 2012-2013.
[6] Following Zheng & Slaper (2016), I tested the threshold of 100 miles, but it did not greatly affect the results, as the magnitudes, signs and significance of the coefficient estimates did not change. For the sake of parsimony, I only include shorter distance in my modeling. The 100-mile distance results are available on request.

1.5. Empirical Results

1.5.1. Summary Statistics

Table 2 shows the summary statistics for all variables in the combined sample as well as each county type. Simple comparison of unconditional means of the number of patent applications per inventive class population across urban and rural counties indicates that metro (urban) counties innovate significantly more than both types of rural counties; metro-adjacent rural counties innovate relatively more than the remote rural counties on average. However, all three county types display large heterogeneity (or dispersion) in the rates of patenting within their groups.[7]

Simple correlation analysis (not shown, but available on request) also support the hypothesized associations between invention rates and various explanatory factors. Table 2 also shows differences among the three county types, with respect to the means of several explanatory variables expected to influence innovation rates. For example, means of patenting spillover, high-tech variety, 4-year colleges and universities, which are hypothesized to positively influence the regional patenting rates, all show highest values for metro counties, followed by metro-adjacent and then by remote-rural counties. The means of people 25-year old or older with bachelor's degree or higher, foreign-born population, share of arts degrees, and high-speed broadband penetration are highest in metro areas. Interestingly, they are higher in remote rural areas compared to metro-adjacent rural areas. On the other hand, tax burden and cellphone service penetration, on average, are the highest in remote rural areas; while, the metro-adjacent counties are more manufacturing intensive among the three county types.

[7] Coefficient of variation (dispersion) for metro counties is (47.617/8.117)*100% =586%, for metro-adjacent counties it is (11.026/0.819)*100% =635%, and for remote rural counties it is (5.753/0.819)*100% =702%

Table 1.2 Summary Statistics

Variables	Combined		Metro		Metro-adj. Rural		Remote Rural	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Patents per inventive class	3.9	30.5	8.1	47.6	1.7	11.0	0.8	5.8
Splag patents per inventive class	8.2	17.7	14.6	23.2	6.2	12.9	1.9	9.1
High-tech variety	14.1	9.5	21.1	10.5	10.9	5.3	8.4	5.2
High-tech share (% points)	4.5	3.6	5.8	3.6	3.7	2.0	3.8	4.
Universities/colleges	0.6	2.4	1.4	3.7	0.1	0.4	0.1	0.4
College plus education (% points)	13.0	5.4	15.5	5.8	10.8	3.8	12.3	4.9
Arts share (% points)	1.7	1.3	2.2	1.5	1.4	1.0	1.5	1.3
Foreign-born-non-citizen population (% points)	2.8	3.5	3.6	3.8	2.2	2.8	2.3	3.5
Unemployment change (in % points)	7.8	28.8	8.1	28.7	7.7	29.1	7.5	28.5
Tax burden (% points)	3.7	2.0	3.7	1.6	3.6	1.9	4.0	2.6
Manufacturing intensity (1=high; 0=low)	3%	N/A*	2%	N/A	5%	N/A	3%	N/A
Professionals service intensity (1=high; 0=low)	3%	N/A	6%	N/A	1%	N/A	1%	N/A
High-speed Broadband penetration (1=high; 0=low)	41%	N/A	60%	N/A	26%	N/A	32%	N/A
Cellphone service penetration (1=high; 0=low)	20%	N/A	16%	N/A	19%	N/A	27%	N/A
Average venture capital, state level, (in $1,000)	0.3	0.5	0.4	0.6	0.3	0.4	0.2	0.4
Observations	14165		5435		4760		3970	

* N/A = Not applicable (for dummy variables; the means refer to frequencies in percent)

These summary statistics suggest that the urban advantage in production of inventive outputs is driven by the higher levels of the factors that are found to be positively associated with regional innovation. However, regression analyses using empirical model (2) was carried out to test if this urban advantage persists after controlling for the differences in the levels of these innovation drivers between urban and rural areas.

1.5.2. Regression Estimation Results

This section analyzes the advantage of urban areas in creating inventive outputs compared to the rural areas using the estimation results of the SAR model in equation (2). For this analysis, I first estimate my empirical model for the combined data set, with county level random-effects, state and temporal fixed-effects. I also include indicator variables for metro-adjacent and remote-rural county types, with metro-counties serving as the reference category. The first column of coefficient estimates in table 3 show the maximum likelihood estimation results from my empirical model in equation (2) for the combined dataset.

Coefficient estimates in the first column of table 3 suggest that urban areas are more inventive than both types of rural areas, with the rates of inventive outputs shown to occur 48% [= ($e^{0.391}$-1)*100%] and 104% [= ($e^{0.715}$-1)*100%][8] more frequently in urban areas than the metro-adjacent and remote rural counties respectively. The positive statistically significant coefficient (ρ) estimate of the spatially lagged dependent variable, indicates significant spillover effects of patenting rates in neighboring counties on the county patenting rate. That is, patenting in a county is positively influenced by conditions that favor patenting in neighboring counties. The coefficient estimates of the explanatory variables indicate statistically significant positive association (at 5% or higher level) of patenting rates with several variables including high-tech variety, colleges and universities, college plus education, professional service intensity indicator, and spillover effects. Foreign-born population is significant only at 10% level. The positive sign on the coefficient of "arts share" supports the hypothesis of synergistic effects of creative class on patenting productivity of the inventive class. However, the variables related to the innovation

[8] See Hilbe (2011) for interpretation of the coefficients. Essentially, incident rate ratio=exp(coefficient estimate) and IRR is interpreted as the rate ratio for a unit change in independent variable of interest.

infrastructure (tax burden, high-speed broadband penetration, and cellphone service penetration) and manufacturing intensity do not show a statistically significant association.

1.5.3. Rural Urban Comparative Advantage in Innovation

The results from the combined sample generally support previous findings regarding the urban innovative advantage and drivers of innovation albeit with a more comprehensive dataset and a more refined count model estimation. However, in view of the persistent urban advantage in patenting rates and significant differences in levels of explanatory variables across county types revealed by the summary statistics, I estimate the regression model shown in equation (2) separately for the three subsamples by county type. The goal of these subsample estimations is to empirically explore differences in the innovative capacities (patenting rates) of rural and urban counties that were otherwise identical within their subsamples. I conduct likelihood ratio test, suggested by Brooks and Lusk (2010), where null hypothesis is framed as various rural and urban regions can be represented by common drivers of their innovative capacity (i.e., use of the combined model for analysis is appropriate). The alternative hypothesis is that the parameters of various drivers across the three county types (coefficient estimates columns 2-4 of table 3) are dissimilar. The results from the likelihood ratio test rejected the null hypothesis in favor of the alternative hypothesis[9], suggesting that there are some differences between rural and urban counties in terms of potential drivers of their patenting intensity. Thus, use of separate estimation models is justified.

[9] The test statistics for the likelihood test is computed as two times the difference between the log-likelihood of model 1 and the sum of the log-likelihoods of models 2-4 in tables 4. For example, the test statistics is 466 {2*[-15283-(-8694-4045-2311)] = 2*-233= 466}. The chi-square critical value with 60 degrees of freedom and 99% confidence level (88.4) is less than the test statistic.

Table 1.3 Results from RENB-SAR Model on Full sample and Sub-samples by County Types

Dep. Var.: Patents per inventive class	Coefficient Estimates			
	Combined	Metro	Metro-adj. Rural	Remote Rural
Metro-adj. Rural	-0.391***	-	-	-
	(0.082)			
Remote Rural	-0.715***	-	-	-
	(0.097)			
Splag patents per inventive class	0.005***	0.003*	0.009**	0.010*
	(0.001)	(0.002)	(0.004)	(0.006)
High-speed broadband penetration (1=high; 0=low)	0.049	0.054	0.080	0.014
	(0.035)	(0.037)	(0.097)	(0.128)
Cellphone service penetration (1=high; 0=low)	0.025	0.005	0.149*	0.240*
	(0.026)	(0.026)	(0.083)	(0.133)
Manufacturing intensity (1=high; 0=low)	0.021	-0.083	0.377**	-0.522
	(0.083)	(0.094)	(0.149)	(0.412)
Professional service intensity (1=high; 0=low)	0.141***	0.074*	0.496**	0.116
	(0.039)	(0.041)	(0.200)	(0.331)
Arts share (in % points)	0.083**	0.100**	-0.013	0.112
	(0.042)	(0.048)	(0.109)	(0.110)
High-tech variety	0.085***	0.057***	0.130***	0.183***
	(0.004)	(0.005)	(0.012)	(0.016)
High-tech share (in % points)	0.008	0.022**	0.002	-0.079*
	(0.009)	(0.011)	(0.029)	(0.040)
Universities/colleges	0.063***	0.086***	0.020	-0.311
	(0.018)	(0.019)	(0.145)	(0.223)
College plus education (in % points)	0.100***	0.092***	0.100***	0.112***
	(0.011)	(0.014)	(0.028)	(0.031)
Foreign-born-non-citizen population (in % points)	0.023*	0.059***	0.041	0.001
	(0.012)	(0.017)	(0.026)	(0.026)
Unemployment change (in % points)	-0.000	0.000	-0.002	-0.001
	(0.001)	(0.001)	(0.002)	(0.003)
Tax burden (in %points)	-0.000	0.023	-0.049	0.033
	(0.015)	(0.019)	(0.038)	(0.032)
Average venture capital, state level (in $1,000)	-0.011	-0.029	0.072	-0.093
	(0.042)	(0.040)	(0.225)	(0.373)
Constant	-1.875***	-1.718***	-0.433***	-1.903
	(0.296)	(0.354)	(0.090)	(1.368)
Time fixed effects	Yes	Yes	Yes	Yes
State-level fixed effects	Yes	Yes	Yes	Yes
Observations	14,165	5,435	4,760	3,970
Log likelihood	-15283	-8694	-4045	-2311
Model DF	67	65	62	60
AIC	30706	17523	8220	4749
BIC	31235	17972	8640	5145

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

My results in columns 2-4 of table 3 show that having high cellphone service penetration is important in rural areas (statistically significant coefficients at 10% level). However, high-speed broadband (via cable/landline) penetration does not appear to influence patenting rates. I find that the spatial spillover effects of patenting in neighboring counties are significant in all county types, but higher in rural counties (at 5% significance level in metro-adjacent counties and 10% in metro and remote rural counties). This suggests that the spillovers from inventive activities in surrounding counties likely have greater influence on invention rates in rural counties. I also find that the concentration of creative population, measured with the share of arts college degree holders in the county population, is positively associated with patenting rates in the urban and metro-adjacent areas, but not statistically significant in rural remote areas. My finding supports the prior findings in the literature on the important role of creative class in the urban innovation and indicates why such a role is neglected in rural settings.

The coefficient estimates of the manufacturing intensity and professional service intensity (positive and statistically significant) in column (3) of table 3 show that patenting is a more frequent phenomenon in both manufacturing and profession service industries located in metro-adjacent rural counties. In urban areas, only the professional service intensity is statistically significant but small in comparison. This provides weak support for the notion that differences in patenting rates between rural and urban areas may also be linked to differences in industry make up. I also find the foreign-born-non-citizen population contributes to the innovation rate, but only in urban areas.

The results (table 3) show that association of high-tech share with patenting rates is opposite in urban and rural counties, with the association being positive in metro counties but negative in remote rural counties. For metro-adjacent counties, it is positive but not statistically significant. These suggest that increases in share of high-tech firms in total firms by a unit percentage point is likely to increase the patenting rates in metro counties by 2.2% [= ($e^{0.022}$-1)*100%], but the patenting rates in remote rural counties respond to the change with 7.6% [= ($e^{-0.079}$-1)*100%] lower patenting rates. On the other hand, the coefficient of high-tech variety is positive and statistically significant in all three county types. Compared to both types of rural counties, the influence of high-tech variety on patenting rate is lower in urban counties. These results suggest that diversity of high-tech industries rather than the absolute number of high-tech firms spurs innovation in remote rural counties, similar to metro-adjacent counties. However, metro counties benefit from both number and diversity of high-tech firms in terms of patent generation. Finally, I find that the share of population with a 4-year degree or higher is a major driver of patenting rates in all three county types, but its influence appears to be stronger in the remote rural counties than in the metro-adjacent or metro counties. The coefficient estimates of the *universities/colleges* in columns (2)-(4) of table 3 show that the number of 4-year private or public academic institutions hosted by a county has statistically significant positive association with patenting rates in metro counties only.

1.6. Summary and Conclusion

I analyzed regional heterogeneity in innovation rates and the drivers of such heterogeneity, using a comprehensive county level dataset covering the period 2009-13. I compare the rates of inventive outputs among the urban areas and two types of rural areas, metro-adjacent rural and remote-rural areas of the U.S. using patent applications per 1,000 inventive class population (measured by the number of science and engineering graduates with four-year college or higher degree). I account for effects on rates of creation of inventive outputs in a county arising from the spatial spillover effects of inventiveness in surrounding counties by using the spatially lagged variable of my measure of inventive output and including high-speed broadband penetration and cellphone service penetration, which are likely to help facilitate the spatial spillover. I also account for patenting heterogeneity across industries by controlling the intensity of manufacturing activities and professional service provision. Additionally, I control other factors that are commonly found in the literature to be important in regional innovation such as advanced educational attainment, presence of colleges and universities, and high-tech firms. I conduct my econometric analysis using spatial autoregressive negative binomial count models of RKPF, aimed at identifying the drivers of differences in inventiveness across urban and rural regions.

I find that urban areas on average are more inventive than the rural areas, even after accounting for the spatial spillover effects, industry effects, and other common factors related to regional innovation. My results show that patenting is characterized by spatial spillover in urban and both types of rural regions, and the spillover effects are stronger in rural regions. This suggests that spatial spillover effects from inventive activities in neighboring areas are important in both urban and rural regions, but the rural communities may be more dependent on ideas and knowledge from adjacent areas, thus receiving higher spillover externalities compared to urban areas. Further, I find that higher penetration of cellphone service is likely to support inventions in rural areas. Lack of evidence on the supportive role of access to high-speed broadband connection does not suggest that it is not important in facilitating regional inventions and its spillover. Instead, urban areas may already have provided such access or internet access; whereas, cellphone service providers may be substituting the access of high-speed broadband connection in rural areas. Additionally, the apparently smaller gap in patenting rates between metro and metro-adjacent counties (48%) compared to those between metro and remote rural counties (104%) is likely driven by the larger share of manufacturing industries that are expected to produce significantly more patents in metro-adjacent counties.

Other results from this study confirm the prior findings in the literature that the diversity of high-tech industries and the percentage of population with advanced educational attainment are important contributors of inventiveness in all three county types- metro, metro-adjacent and remote rural counties. This suggests that the urban advantage in inventions is likely to arise from larger number of high-tech firms, advance educational institutions, and larger population share of new immigrants and creative class individuals. I do not find evidence of influence of tax burden, unemployment rate, and state-level venture capital on patenting rates in my study.

My results provide two policy insights regarding regional inventiveness and economic development. First, the results suggest that the policies intending to mitigate the rural-urban inventiveness gap should focus on building strong communication infrastructures in the rural regions, as these infrastructures are likely to generate stronger spillover effects in the rural regions. Second, the key drivers of inventiveness such as advanced education and diversity of high-tech industries play important roles in driving the rates of inventive outputs in urban and rural counties, with the importance (of both the variables) being more critical in rural areas. But, on average, these rural areas are less diverse and have lower levels of population share with advanced education. So, the similar policies promoting investments in education and attracting more diverse high-tech industry are effective in both urban and rural areas, but relatively larger investments may be needed in rural areas compared to urban areas, due to agglomeration-related dis-economies in rural areas. More important but related to the second implication, an increase in the number of high-tech firms within the existing diversity is expected to spur the rate of invention output in metro areas but it is expected to have opposite influence in remote rural counties. Therefore, the policies to promoting high-tech firms may further amplify the inventiveness gap between urban and remote rural areas if such policies fail to attract the firms from diverse high-tech industries.

Finally, some caveats are in order in interpreting the findings of this study. I use patents/inventive population as an overall indicator of inventive capacity and productivity of a region, which is subject to criticism as discussed in the literature review. Further, the patent data I obtained from USPTO contain the residential address of the inventor(s). I used the county of residence of the first inventor, in case of multiple inventors, to match the patent data with other county level data. This may lead to bias in comparison as the place of work of the (first) inventor might be different from his/her place of residence. Additionally, my patent dataset does not distinguish between product or process innovations and this distinction might have implications for the growth effects of innovation, nor does the patents data provide information on whether the patent represented an incremental innovation or a radical invention.

As the prior findings suggest that the incremental efficiency improving/cost reducing inventions are likely to occur more frequently in rural areas, but the radical inventions are more likely in the urban areas (e.g., Orlando & Verba, 2005). Such information may also reveal potential chain-patenting suspected to be occurring. Due to lack of more detailed data, I assume that these effects are random. Finally, in my construction of spatial weight matrix, I do not distinguish among the spillovers across county types, for example if spillovers are prevalent more from urban to rural areas or vice versa, which might be an interesting topic for future research.

REFERENCES

REFERENCES

Acs, Z. (2006). How is entrepreneurship good for economic growth? *Innovations: technology, governance, globalization*, *1*(1), 97-107.

Acs, Z. J., Anselin, L., & Varga, A. (2002). Patents and Innovation Counts as Measures of Regional Production of New Knowledge. *Research policy*, *31*(7), 1069-1085.

Acs, Z. J., Audretsch, D. B., & Feldman, M. P. (1994). R & D spillovers and recipient firm size. *The review of Economics and Statistics*, 336-340.

Acs, Z. J., & Audretsch, D. B. (2005)a. Entrepreneurship, innovation, and technological change. *Foundations and Trends in Entrepreneurship, 1*(4), 149-195.

Acs, Z.J., & Audretsch, D.B. (2005)b. Entrepreneurship and innovation. Discussion Papers on *Entrepreneurship, Growth and Public Policy*, 2005-21, Max Planck Institute of Economics, Group for Entrepreneurship, Growth and Public Policy.

Aghion, P., & Howitt, P. (1990). *A model of growth through creative destruction* (No. w3223). National Bureau of Economic Research.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers.

Anselin, L., Varga, A., & Acs, Z. (1997). Local geographic spillovers between university research and high technology innovations. *Journal of urban economics, 42*(3), 422-448.

Audretsch, D. B., & Feldman, M. P. (2004). Knowledge spillovers and the geography of innovation. *Handbook of regional and urban economics, 4*, 2713-2739.

Audretsch, D. B., & Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production. *The American economic review, 86*(3), 630-640.

Autant-Bernard, C. (2001). The geography of knowledge spillovers and technological proximity. *Economics of Innovation and New Technology, 10*(4), 237-254.

Barkley, D. L., Henry, M. S., & Lee, D. (2006). Innovative Activity in Rural Areas: The Importance of Local and Regional Characteristics. *Community Development Investment Review, 2*(3), 1-14.

Bartik, T. J. (1991). Who benefits from state and local economic development policies? Retrievied from http://research.upjohn.org/cgi/viewcontent.cgi?article=1093&context=up_press

Bound, J., Cummins, C., Griliches, Z., Hall, B.H., & Jaffe, A. (1984). Who Does R & D and Who Patents? In Griliches, Z. (Ed.), *R & D, Patents, and Productivity*, 21-54, Chicago, IL: University of Chicago Press.

Brooks, K., & Lusk, J. L. (2010). Stated and revealed preferences for organic and cloned milk: combining choice experiment and scanner data. *American Journal of Agricultural Economics*, *92*(4), 1229-1241.

Buesa, M., Heijs, J., & Baumert, T. (2010). The determinants of regional innovation in Europe: A combined factorial and regression knowledge production function approach. *Research Policy*, *39*(6), 722-735.

Cameron, G. (1996). *Innovation and Economic Growth*. Centre for Economic Performance, London School of Economics and Political Science.

Capello, R., & Lenzi, C. (2014). Spatial heterogeneity in knowledge, innovation, and economic growth nexus: conceptual reflections and empirical evidence. *Journal of Regional Science*, *54*(2), 186-214.

Carlino, G. A., Chatterjee, S., & Hunt, R. M. (2001). *Knowledge Spillovers and the New Economy of Cities*. Economic Research Division, Federal Reserve Bank of Philadelphia.

Carlino, G. A., Chatterjee, S., & Hunt, R. M. (2007). Urban Density and the Rate of Invention. *Journal of Urban Economics*, *61*(3), 389-419.

Charlot, S., Crescenzi, R., & Musolesi, A. (2014). Econometric modelling of the regional knowledge production function in Europe. *Journal of Economic Geography*, *15*(6), 1227-1259.

Cincera, M. (1997). Patents, R&D, and technological spillovers at the firm level: some evidence from econometric count models for panel data. *Journal of Applied Econometrics*, *12*(3), 265-280.

Cohen, W. M., & Levin, R. C. (1989). Empirical studies of innovation and market structure. *Handbook of industrial organization, 2*, 1059-1107.

Conley, K., & Whitacre, B.E. (2016). Does Broadband Matter for Rural Entrepreneurs and Creative Class Employees? *The Review of Regional Studies, 46*(2), 171-190.

Czarnitzki, D., Kraft, K., & Thorwarth, S. (2009). The knowledge production of 'R' and 'D'. *Economics Letters*, *105*(1), 141-143.

Czarnitzki, D., Glänzel, W., & Hussinger, K. (2009). Heterogeneity of patenting activity and its implications for scientific research. *Research Policy*, *38*(1), 26-34.

Feldman, M. P., & 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), 210-229.

Feser, E., & Isserman, A. (2006). Harnessing growth spillovers for rural development: The effects of regional spatial structure. *Report to USDA Rural Development, University of Illinois at Urbana-Champaign*.

Florida, R. (2002). The rise of the creative class. *The Washington Monthly*, *34*(5), 15-25.

Florida, R. (2014). *The rise of the creative class-revisited: Revised and expanded*. New York, NY: Basic Books.

Florida, R., Adler, P., & Mellander, C. (2016). The city as innovation machine. *Regional Studies*, 1-11.

Glaeser,E.L., Kallal, H.D., Scheinkman, J.A., & Shleifer, A. (1992). Growth in Cities. *Journal of Political Economy*, 100, 1126–1153

Greene, W. (2003). *Econometric Analysis*. New Jersey: Prentice Hall.

Griliches, Z. (1979). Issues in assessing the contribution of research and development to productivity growth. *The Bell Journal of Economics*, 92-116.

Griliches, Z. (1990). Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature*, *28*(4), 1661-1707.

Hall, B., Griliches, Z., & Hausman, J. (1986). Patents and R&D: Is There a Lag? *International Economic Review*, *27*(2), 265-283.

Hausman, J., Hall, B. H., & Griliches, Z. (1984). Econometric Models for Count Data with an Application to the Patents R&D Relationship. *Econometrica*, 52, 909-38.

Henderson, J. (2002). Building the rural economy with high-growth entrepreneurs. *Economic Review-Federal Reserve Bank of Kansas City*, *87*(3), 45.

Henderson, J. V. (2003). Marshall's scale economies. *Journal of urban economics*, *53*(1), 1-28.

Henderson, J., & Abraham, B. (2004). Can Rural America Support a Knowledge Economy? *Economic Review-Federal Reserve Bank of Kansas City*, *89*(3), 71.

Henderson, J., & Executive, B. (2007). Understanding Rural Entrepreneurs at the County Level: Data Challenges. *Federal Reserve Bank of Kansas City–Omaha Branch*.

Henderson, J., & Weiler, S. (2010). Entrepreneurs and Job Growth: Probing the Boundaries of Time and Space. *Economic Development Quarterly*.

Hilbe, J. M. (2011). *Negative binomial regression* (2nd ed.). New York: Cambridge University Press.

Huang, C., Arundel, A., & Hollanders, H. (2010). How firms innovate: R&D, non-R&D, and technology adoption. *UNU-MERIT Working Paper 2010-2027.*

Jaffe, A. B. (1989). Real effects of academic research. *The American Economic Review*, 957-970.

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *The Quarterly journal of Economics*, 577-598.

Kerr, W. R. (2013). *US high-skilled immigration, innovation, and entrepreneurship: Empirical approaches and evidence* (No. w19377). National Bureau of Economic Research.

Kerr, W. R., & Lincoln, W. F. (2010). The supply side of innovation: H-1B visa reforms and US ethnic invention. *Journal of Labor Economics*, *28*(3), 473-508.

Kleinknecht, A. (1987). Measuring R & D in small firms: How much are I missing? *The Journal of Industrial Economics*, 253-256.

Kleinknecht, A., & Verspagen, B. (1989). R&D and market structure: The impact of measurement and aggregation problems. *Small Business Economics*, *1*(4), 297-301.

Mann, J., & Shideler, D. (2015). Measuring Schumpeterian Activity Using a Composite Indicator. *Journal of Entrepreneurship and Public Policy*, *4*(1), 57-84.

Mansfield, Edwin. (1984). Comment on Using Linked Patent and R&D Data to Measure Interindustry Technology Flows. In Griliches, Z. (ed.), *R&D, Patents, and Productivity*, Chicago, IL: University of Chicago Press, 462-464.

Markley, D. (2001). Financing the new rural economy. *Exploring Policy Options for a New Rural America*, *69*, 80.

McGranahan, D., & Wojan, T. (2007). Recasting the creative class to examine growth processes in rural and urban counties. *Regional studies*, *41*(2), 197-216.

McGranahan, D. A., Wojan, T. R., & Lambert, D. M. (2010). The rural growth trifecta: outdoor amenities, creative class and entrepreneurial context. *Journal of Economic Geography*, 11(3), 529-557.

Monchuk, D. C., & Miranowski, J. A. (2010). The Impacts of Local Innovation and Innovative Spillovers on Employment and Population Growth in the US Midwest. *Journal of Regional Analysis & Policy*, *40*(1), 61.

Mukherjee, A., Singh, M., & Žaldokas, A. (2017). Do corporate taxes hinder innovation? *Journal of Financial Economics*, *124*(1), 195-221.

Nagaoka, S., Motohashi, K., & Goto, A. (2010). Patent statistics as an innovation indicator. *Handbook of the Economics of Innovation*, *2*, 1083-1127.

National Science Foundation. (2017). Science & Engineering Indicators 2016. Retrieved from https://www.nsf.gov/statistics/2016/nsb20161/#/

Niebuhr, A. (2010). Migration and innovation: Does cultural diversity matter for regional R&D activity? *Papers in Regional Science*, *89*(3), 563-585.

Orlando, M. J., & Verba, M. (2005). Do Only Big Cities Innovate? Technological Maturity and the Location of Innovation. *Economic Review-Federal Reserve Bank of Kansas City*, *90*(2), 31.

Pakes, A., & Griliches, Z. (1980). Patents and R&D at the Firm Level: A Firs Report. *Economics letters*, *5*(4), 377-381.

Parent, O., & LeSage, J. P. (2008). Using the variance structure of the conditional autoregressive spatial specification to model knowledge spillovers. *Journal of Applied Econometrics*, *23*(2), 235-256.

Parker, J., Mann, J. & Loveridge, S. (2017). Rural V. Urban: A National Survey on Determinants of Business Innovation Activities. Midwest Economics Association, Cincinnati, Ohio, March 31-April 2, 2017.

Ponds, R., Oort, F. V., & Frenken, K. (2009). Innovation, spillovers and university–industry collaboration: an extended knowledge production function approach. *Journal of Economic Geography*, *10*(2), 231-255.

Renski, H., & Wallace, R. (2012). Entrepreneurship in Rural America. *Financing Economic Development in the 21st Century*, *245*.

Rosenthal, S. S., & Strange, W. C. (2004). Evidence on the nature and sources of agglomeration economies. *Handbook of regional and urban economics, 4*, 2119-2171.

Scherer, F. (1965)a. Corporate Inventive Output, Profits, and Growth. *Journal of Political Economy, 73*(3), 290-297

Scherer, F. (1965)b. Firm Size, Market Structure, Opportunity, and the Output of Patented Inventions. *The American Economic Review, 55*(5), 1097-1125.

Stephens, H. M., Partridge, M. D., & Faggian, A. (2013). Innovation, Entrepreneurship and Economic Growth in Lagging Regions. *Journal of Regional Science*, *53*(5), 778-812.

Varga, A. (2000). Local academic knowledge transfers and the concentration of economic activity. *Journal of Regional Science*, *40*(2), 289-309.

Wojan, T., & Parker, T. (2017). Innovation in the Rural Nonfarm Economy: Its Effect on Job and Earnings Growth, 2010-2014. USDA-ERS Report, ERR-238.

Wojan, T. R., Dotzel, K. R., & Low, S. A. (2015). Decomposing Regional Patenting Rates: How the Composition Factor Confounds the Rate Factor. *Regional Studies, Regional Science, 2*(1), 535-551.

Zheng, P., & Slaper, T. (2016). University Knowledge Spillovers, Geographic Proximity and Innovation: An Analysis of Patent Filings Across U.S. Counties. Kelley School of Business Research Paper No. 16-76. Retrieved from SSRN: https://ssrn.com/abstract=2857130 or http://dx.doi.org/10.2139/ssrn.2857130

ESSAY 2: EXPLORING THE INTERDEPENDENCE OF INNOVATION AND REGIONAL ECONOMIC GROWTH: A COUNTY LEVEL ANALYSIS

2.1. Introduction

Regional economic growth models have commonly accounted for the interdependency between the choice of households for their place of residence and the decision of firms to locate their business by modeling the equilibrium levels of population and employment induced by those decisions as being simultaneously determined. Carlino & Mills (1987) and Steinnis & Fisher (1974) used a two-equation system in which equilibrium levels of population and employment are determined simultaneously, influenced by several region-specific exogenous factors that the firms and households consider in their location decisions. It is possible within such partial equilibrium framework that some of these variables, mainly income, are simultaneously determined along with population and employment. Arguing that the quality and number of jobs that move to a region determine the equilibrium level of income in a region, Deller et al. (2001) account for this possibility of interdependent relationship of income with population and employment by modeling a system of three simultaneous equations, and by treating income as endogenous variable in the system. However, the above studies ignore the potential interdependent effect of innovation on regional economic growth.

The role of innovation and innovative entrepreneurs in regional growth has been well documented in several empirical studies (e.g. Adelaja et al., 2009; Feser & Isserman, 2006; Henderson & Weiler, 2010; McGranahan et al., 2010; Monchuk & Miranowski, 2010; Stephens et al., 2013; Young at al., 2014; 2013) and they have largely confirmed the expected positive association of population, employment, and income growth with innovation. Improvements in the innovation-led productivity of input factors including labor productivity result in income and employment growth (Ezell & Atkinson, 2010). Innovative industries are found to generate higher overall employment accompanied with Schumpeterian “creative destruction” – displacement of lower-wage and lower-productivity jobs with better-paid and more productive ones (Helpman, 2004; OECD, 1994).

The extant literature mentioned above has considered unidirectional positive influence of innovation rates on regional economic growth as measured by income, employment and population. These studies treat innovation as an exogenous driver of economic growth. However, I posit that innovation rates themselves are in turn influenced by economic opportunities and growth. Higher incomes and employment opportunities attract more technically qualified and creative class of people to the region and encourage synergistic creative cooperation resulting in higher innovation rates. In other words, I argue that innovation plays an important role in regional economic growth, and equilibrium innovation rates are in turn simultaneously determined with levels of income, employment and population.

Empirically, I consider the role of innovation in regional population, employment, and income growth by expanding Deller et al.’s (2001) model. I extend their three-equation system to include innovation growth as an endogenous variable simultaneously determined along with population, employment, and income growth, wherein the growth in county level patent applications between 2009-13 serves as a measure of innovation, Additionally, I analyze the gaps in economic growth and innovation rates between leading and lagging regions including rural and urban regions, and factors influencing economic growth and innovation.

The remainder of the article is organized as follows. In the next section, I review the literature on the relationship between innovation and economic growth and their determinants. The conceptual and empirical models are laid out in section three. Section four provides the description of the data and variables. In section five, I present and discuss my findings. Section six concludes with policy implications and scope for further research.

2.2. Literature Review

This section reviews the extant literature on the role of innovation in economic growth. It also contains a review of the factors influencing regional innovation and economic growth, which facilitates the choice of variables used in this study.

2.2.1. Innovation and Economic Growth

Innovation is arguably a phenomenon as old as human history, as it is inherently human to produce new ideas for something better. However, the role of innovation in economic growth lacked scholarly attention until the emergence of Marxian-Schumpeterian theoretical perspective (Fagerberg, 2003). According to this perspective, innovations in firms occur as they strive to survive technological competition, which is the major form of competition in the “capitalist” economy. These innovations spur possibilities for new businesses and further innovations, thus setting stage for long-run economic growth. In other words, it implies that innovation stimulates growth and the latter leads to more innovations as firms continuously seek to find new combination of resources to remain technologically competitive. Neither Marx nor Schumpeter applied their perspectives to explain the macro-level implications of innovation in terms of cross-country differences in economic growth, but several analyses from 1960s onwards suggested innovation as the major explanatory factor (Fagerberg, 1994, 2002; Fagerberg & Verspagen, 2002; Posner, 1961).

Innovation and technological change are regarded important factors also in the “classical” approach to explain growth (e.g. Solow, 1956), but only were associated to the part of the growth that cannot be explained by the contributions of the accumulated traditional factors of production such as labor and capital[10]. In other words, the classical approach assumed the rates of innovations as exogenously determined in the economic system rather than explaining the mechanism that results to the technological change.

Marxian-Schumpeterian perspective received renewed attention among the economists after the emergence of "new growth theory" (e.g. Romer, 1986; Lucas, 1988) that introduced knowledge capital as an intermediary input in the firm's production function. In the new growth models, innovation and technical change result from decisions of the profit maximizing agents to produce knowledge and utilize it as an input to production of final outputs. Thus, conceptually parallel to the Marxian-Schumpeterian early growth models, these models, also known as "endogenous growth models", emphasize growth as a result determined endogenously by the rational actions of (innovative) entrepreneurs, who identify opportunities, introduce innovations, create competition that demands further innovations, and offer more business opportunities for new entrepreneurs (Acs et al., 2012; Carree & Thurik, 2003; Wennekers & Thurik, 1999; Wong et al., 2005).

Theoretically, the relationship between innovation and economic growth and development has been acknowledged by endogenous growth theorists for a long time beginning with Kuznets (1971) and Marshall (1990). Although perspectives differ on the mechanisms by which innovation influences economic growth, that is, whether knowledge is endogenous or exogenous in the system of economic growth, it is commonly recognized that change in knowledge and economic growth affect each other directly and indirectly (Howells, 2005). In the real world as well, innovation has been the central driver of prosperity and economic growth.

[10] See Nadiri (1993) for a summary of the summary of studies using this approach.

(Feldman & Florida, 1994). Innovation-led productivity has been estimated to account for nearly half of the U.S. GDP growth in past 50 years following the World War II (Mandel, 2004) and more than two-thirds of the income growth (Ezell & Atkinson, 2010). Improvement in labor productivity in innovative industries is associated with creation of better-paid jobs by displacing lower-paid ones[11] (Helpman, 2004; OECD, 1994). The influence of new ideas and products on people's lives has been growing at ever-increasing rate. For example, it took 35 years for the telephone to reach a quarter of the American population, but just 13 years for the cellphones, and seven years for the internet (Federal Reserve Bank of Dallas, 1996).

America remains as the global economic leader due to its competitive advantage arising from its early adoption of an open and inclusive market economy that attracted talented workforce and innovative entrepreneurs from across the world (Ezell, 2009). Now that several countries have learned and adopted the lessons from America's open market experience and can compete on traditional cost and quality terms, the ability to constantly create new and better products and services will confer the major competitive edge in the 21st century. The fact that more than three dozen of countries have formulated innovation-led strategies for their economic growth and development in the first decade of the 21st century (Ezell, 2009) suggests that innovation will become more powerful driver of economic growth in the future than ever before.

2.2.2. Innovation and Regional Economic Growth

The models of endogenous growth recognize that knowledge is only partially excludable, as the firms producing it are unable to fully appropriate it. This knowledge spills over to the regional knowledge stock that benefits the surrounding firms (Acs et al., 2012). The implications

[11] See Capello & Lenzi (2013) for the discussion on potentially different roles of product and process innovations in employment growth. Data limitations do not allow us to analyze the difference in this analysis.

of this theoretical recognition combined with those from the Marxian-Schumpeterian perspective discussed above is that regional innovation and economic growth are interrelated, or they drive each other to continuous change.

The knowledge spillover from R&D activities of surrounding institutions such as universities and other firms served as an explanation for the source of knowledge input for small enterprises that were innovative but lacked their own R&D commitments (Acs & Audretsch, 2005; Jaffe, 1989). Subsequently, the findings from a series of studies that knowledge spillover tends to be bounded by geographical distance (Anselin et al., 1997; Audretsch & Feldman, 1996; Jaffe et al., 1993) marked a major advance in inquiries into the role of geography in innovation and growth.

The economic growth of regions after the end of the World War II was focused more on the internal competition within a broad economic region than global competition. Various states in the U.S., and similarly European nations, began to compete among each other to attract firms by investing in physical infrastructure (Ezell & Atkinson, 2010). This strategy provided the firms with efficiency gains through reduction in costs, especially transportation costs, and extended local markets. However, with increasing global economic integration and the advances in information and communication technologies, the regions are getting more interconnected and innovations are affecting broader geographies, and spreading faster globally. For example, nearly half of U.S. productivity growth comes from improvements in technology in foreign countries (Eaton & Kortum, 1995). Therefore, the competitiveness of the regions that are integrating economically at global scale at ever-increasing pace lies more in the capability of firms and industries to create constantly new value in the global marketplace, which depends on regional 'technological infrastructures' (Feldman & Florida, 1994).

2.2.3. Drivers of Regional Innovation

Regional innovation has been receiving increasing attention of academic scholars and innovation policy makers in the past two decades (e.g. Audretsch, 1998; Feldman, 1994; Jaffe, 1989Krugman, 1991) with special focus on knowledge spillovers in geographically bounded areas (e.g. Acs et al., 1992; Jaffe et al., 1993) and on innovative activities of spatially concentrated industries (e.g. Audretsch & Feldman, 1996; Brenner, 2004; Porter, 1990). Feldman & Florida (1994) addressed the influence of local factors on regional innovativeness using the regional ‘technological infrastructures’. Following the work of Cooke (1992), ‘regional innovation system’ emerged as a different concept to understand regional innovation by combining the systemic nature of innovation within a geographic context. The empirical literature in regional innovation differs on the notion of agglomeration economies. One strand of literature argues that creation of new regional knowledge is based on the spillovers arising from the concentration of firms in an industry, also known as ‘Marshall-Arrow-Romer’ externalities (see Rosenthal & Strange, 2001) and the other explains it in terms of spillovers among industries, also known as ‘Jacobian’ externalities (see Frenken et al., 2007). However, the entire literature on regional innovation is in consensus with the major role of human capital, research and development (R&D) expenditures by private firms and universities, population and employment density, and industrial diversity (or concentration) as factors affecting regional innovativeness.

2.2.4. Drivers of Regional Growth

To explain the observed regional population and employment patterns, early researchers sought to understand the roles of factor prices, markets, fiscal characteristics (Bartik, 1985; Helms, 1985; Plaut & Pluta, 1983; Romans & Subrahmanyam, 1979; Wheat, 1986), agglomeration economies (Carlino, 1985), and public policy (Bartik, 1991). Most of the early models of regional growth considered population and employment separately. Carlino & Mills (1987) developed lagged adjustment models for county-level growth in which employment and populations are simultaneously determined. Their findings suggested that income was an important determinant of employment and population growth, while climatic conditions and local tax policies were important for population growth. Clark & Murphy (1996) applied the Carlino & Mills model to sectoral employments and population growth at the county-level U.S. data and found evidence of simultaneous feedback between population density and employment density. Their findings showed that climate variables such as temperature and sunshine and local government expenditures had minor effects on both short-run and long-run growth in population. They also found stronger influence of employment density on population density than that of the latter on the former.

Deller et al. (2001) extended the Carlino & Mills model to include changes in regional income levels as simultaneously determined along with changes in population and employment levels. Using the model to study the economic growth in nonmetropolitan U.S. counties, they found that a range of factors including natural resource amenities, property tax, education levels, distribution of income, and age played major role in regional economic growth. Their results of the negative relationships between the initial levels and growth in population, employment and income suggested that rural areas were catching up to offer increased economic opportunities.

Using the conditional convergence model, where the rate at which the poorer regions catch up the richer regions is assumed to be conditioned on several regional structural factors[12], to study the county level income growth, Rupasingha et al. (2002) consider the role of social and institutions factors – social capital, income inequality, ethnic diversity in addition to other

[12] See Yeager (1999) for some examples of structural factors affecting economic convergence.

factors. They find that higher social capital, lower income inequality, and higher diversity were associated with higher income growth. The other variables they found to have positive association with income growth include higher level of human capital, lower local tax, and higher highway expenditure. Using the similar model of conditional convergence, Pede (2013) found that economic diversity, measured by the distribution of employment across sectors in the U.S. counties through several entropy indices, was positively associated with county level per capita income growth. He also found positive association with respect to other variables such as percentage of bachelor's degree holders, and age composition of the population, and metropolitan indicator.

Frenken et al. (2007) studied the role of industrial variety in regional employment growth in the Netherlands. They found that the industrial variety, as a measure of 'Jacobian externalities' was positively associated with the employment growth at NUTS-3 level[13].

In their analysis of the determinants of income growth across U.S. labor markets using production function approach, Hammond & Thompson (2008) found little support for the role of public capital investment. Human capital investment had larger income growth in metropolitan regions than non-metro, but manufacturing investment had larger effect in the latter regions. They also found that regions with colleges and universities, lower tax rates, and higher level of household amenities accumulated larger pool of human capital.

Komarek & Loveridge (2015) investigate the role of firm size distribution on the county-level growth in the US between 1990 and 2000 and find that the larger share of small firms had positive impact on employment growth but negative effect on income growth; medium sized firms affected the income growth positively. They suggest that small businesses are net job

[13] NUTS stands for Nomenclature of Territorial Units for Statistics, a geocode standard, and the Netherlands has 40 NUTS-3 units.

creators but pay lower wages. Additionally, they find larger pool of highly educated population, and urban counties were positively associated with the employment and income growth.

The empirical literature on the relationship between innovation and regional economic growth at the regional level is relatively sparse compared to analyses at national level. Investigating the effects of urban to rural spillovers on regional economic growth, Feser and Isserman (2006) used a cross-section of 3,079 US counties and measured innovative activities with average utility patents over the 1990-95 period in region around a county in their set of explanatory variables. From their two-stage least square (2SLS) estimation, they found that high level of patenting was associated with high level of employment growth during 1990-2000. They also found that growth spilled over more to the rural counties proximate to the highly-urbanized counties than those proximate to the less urbanized counties.

Monchuk & Miranowski (2010) found that innovation, as measured by utility patents, were positively associated with employment and population growth in the Midwest regions during 1990-2005. Additionally, they found that increased rurality makes growth slower, although innovation positively affected the growth in rural regions. Adelaja (2009) analyzed a sample of 3,023 US counties by estimating a linear system of simultaneous equations. He found positive relationship of average patents held in a county during 1990-1993 with both its employment and per capita income growth during 1990-2000.

In a test of the empirical relationship between technological change and employment growth at regional level of NUTS2 European countries, Capello & Lenzi (2013) report that product innovations lead to job growth in regions specialized in production sector, whereas process innovations dampen the job growth in regions with large cities.

Stephens et al. (2013) studied the role of several factors including knowledge-based factors - proximity to universities, patenting rates, college graduates, creative workers, and high-tech employment share - in the growth of wage and salary employment in the economically lagging region of Appalachia. They didn't find significant association between employment growth and these factors except the creative workers. Self-employment, as a proxy for entrepreneurship, was found to have very strong association with wage and salary employment growth. In an analysis of the effect of the SBA lending on the growth of US counties between 1990 and 2008, Young et al. (2014) used citation-weighted patents per capita as one of the several control variables and reported negative association between patenting and county growth.

In a study on the role of proximity to the nearest urban centers in the regional economic growth, Partridge et al. (2008) examine the U.S. county level employment growth by differentiating distance effects for several tiers in the American urban hierarchy. They find that the regions more proximate to the urban centers grow faster than the distant regions and conclude that distance effects are stronger over time.

In summary, the extant literature in regional economics has identified and empirically analyzed a number of drivers of regional economic growth measured by income, employment and population. A separate stream of literature has analyzed factors influencing regional innovation, including R&D inputs, spillover effects, and other socioeconomic drivers, and the unidirectional contribution of innovation to economic growth. Influential analyses (e.g. Deller et al., 2001) have analyzed the interdependence between income, population and employment. Although endogenous growth theorists have recognized and analyzed interdependence of innovation and national economic growth, surprisingly no research has empirically analyzed the interdependencies between income, employment, population and innovation at the regional level using the general equilibrium framework.

2.3. Modeling and Estimation

This section presents the conceptual model of regional economic growth based on the general equilibrium framework, specifies the empirical regression model for this study, proposes the hypotheses to be tested, and discusses the estimation method.

2.3.1. Regional Growth Model

Profit maximizing firms choose their location based on the factors that affect their production and distribution costs. The production cost depends on the supply of the inputs such as labor, capital, and land, and the distribution cost depends on the distance to the output markets. Capital input typically refers to physical capital. I allow knowledge capital to be a component of firms' capital. Therefore, firms' location choices are also dependent on factors such as the population with higher education and the opportunity to collaborate with universities. According to endogenous growth theory, innovative firms intentionally decide to invest in innovation inputs such as human and R&D capital. So, the firms decide their extent of innovation jointly with other traditional decisions. On the consumer side, utility maximizing consumers derive their utility from the purchased goods and services that the firms provide, so their residential location decision depends on the supply of such goods and services. Equilibrium population and employment are determined by a host of factors that affect the location decision of the firms and the consumers. Carlino & Mills (1987) assume that the equilibrium levels of population and employment are simultaneously determined while all other factors affecting them are exogenous. Deller et al. (2001) model simultaneous determination of income together with population and employment. They argued their approach helps capture the job quality and understand the regional growth process. In other words, people that make migration decisions consider the quality of life communities can support through the income levels that are determined by the opportunities to get existing work or start a new business. Innovative firms are likely to generate growth in overall income and employment through improved factor productivity (Ezell & Atkinson, 2010) and displacement of low-wage jobs with better-paid jobs (Helpman, 2004; OECD, 1994). From a regional perspective, the regions are likely to vary in terms of their entrepreneurial culture leading to varying rates of innovation and job creation among regions. I therefore posit that innovation is endogenous in the system of regional growth. My model enables us to examine the role of innovation in economic growth, specifically whether it drives the growth of county-level regional economies, or is led by the regional economic growth, or is determined simultaneously along the growth process.

I build upon the Carlino & Mills (1987) and Deller et al. (2001) model for simultaneous system of regional growth. Following Deller et al. (2001), who add income to the two-equation system of Carlino & Mill's (1987) model, I add innovation to their three-equation system. With endogenous innovation, the general structural model expands to following system of four linear equations:

$$Pop^{*} = \alpha_{0P} + \beta_{1P}Emp^{*} + \beta_{2P}Inc^{*} + \beta_{3P}Innov^{*} + \sum \delta_{IP}\Omega^{P0p} \tag{1}$$

$$Emp^{*} = \alpha_{0E} + \beta_{1E}Pop^{*} + \beta_{2E}Inc^{*} + \beta_{3E}Innov^{*} + \sum \delta_{IE}\Omega^{Emp} \tag{2}$$

$$Inc^{*} = \alpha_{0Inc} + \beta_{1Inc}Pop^{*} + \beta_{2Inc}Emp^{*} + \beta_{3Inc}Innov^{*} + \sum \delta_{IInc}\Omega^{Inc} \tag{3}$$

$$Innov^{*} = \alpha_{0Innov} + \beta_{1Innov}Pop^{*} + \beta_{2Innov}Emp^{*} + \beta_{3Innov}Inc^{*} + \sum \delta_{IInnov}\Omega^{Innov} \tag{4}$$

where Pop^{*}, Inc^{*}, Emp^{*}, and $Innov^{*}$ represent equilibrium levels of the endogenous variables population, employment, personal income per capita, and innovation, and Ω^{Pop}, Ω^{Emp}, Ω^{Inc}, and

Ω^{Innov} contain the set of variables representing initial conditions of the dependent variables and the exogenous regional characteristics. The subscripts on the parameters and the superscripts on the set of exogenous variables identify their association with their respective dependent variables.

Following Mills & Price (1984), Carlino & Mills (1987), and Deller et al. (2001), the population, employment, income, and innovation adjust to their equilibrium levels through a distributed-lag adjustment process as follows:

$$Pop_t = Pop_{t-1} + \lambda_{Pop}(Pop^* - Pop_{t-1}) \tag{5}$$

$$Emp_t = Emp_{t-1} + \lambda_{Emp}(Emp^* - Emp_{t-1}) \tag{6}$$

$$Inc_t = Inc_{t-1} + \lambda_{Inc}(Inc^* - Inc_{t-1}) \tag{7}$$

$$Innov_t = Innov_{t-1} + \lambda_{Innov}(Innov^* - Innov_{t-1}) \tag{8}$$

where the subscripts t and $t-1$ represent the values of the variables at a time and its one period lag (five years in my study) respectively, and λ's are the speeds of adjustment to equilibrium levels of their respective variables with $0 \le \lambda_{Pop}, \lambda_{Emp}, \lambda_{Inc}, \lambda_{Innov} \le 1$.

Rearranging equations 5-8 and substituting their equilibrium values from equations 1-4, the following system of equations can be derived:

$$\Delta Pop = \alpha_{0P} + \beta_{1P}Pop_{t-1} + \beta_{2P}Emp_{t-1} + \beta_{3P}Inc_{t-1} + \beta_{4P}Innov_{t-1} + \gamma_{1P}\Delta Emp + \gamma_{2P}\Delta Inc + \gamma_{3P}\Delta Innov + \sum \delta_{IP}\Omega^P \tag{9}$$

$$Emp = \alpha_{0E} + \beta_{1E}Pop_{t-1} + \beta_{2E}Emp_{t-1} + \beta_{3E}Inc_{t-1} + \beta_{4E}Innov_{t-1} + \gamma_{1E}\Delta Pop + \gamma_{2E}\Delta Inc + \gamma_{3E}\Delta Innov + \sum \delta_{IE}\Omega^E \tag{10}$$

$$Inc = \alpha_{0Inc} + \beta_{1Inc}Pop_{t-1} + \beta_{2Inc}Emp_{t-1} + \beta_{3Inc}Inc_{t-1} + \ + \ \gamma_{1Inc}\Delta Pop + \gamma_{2Inc}\Delta Emp + \gamma_{3Inc}\Delta Innov + \sum \delta_{IInc}\Omega^{Inc} \tag{11}$$

$$\Delta Innov = \alpha_{0Innov} + \beta_{1Innov}Pop_{t-1} + \beta_{2Innov}Emp_{t-1} + \beta_{3Innov}Inc_{t-1} + \beta_{4Innov}Innov_{t-1} + \gamma_{1Innov}\Delta Pop + \gamma_{2Innov}\Delta Emp + \gamma_{3Innov}\Delta Inc + \sum \delta_{IInnov}\Omega^{Innov} \quad (12)$$

where $\Delta Pop = Pop_t - Pop_{t-1}$, $\Delta Emp = Emp_t - Emp_{t-1}$, $\Delta Inc = Inc_t - Inc_{t-1}$, and $\Delta Innov = Innov_t - Innov_{t-1}$. Note that the α and β's in this system of equations are different from those in the system of equations 5-8 absorbing the speeds of adjustment, λ's.

The dependent variables ΔPop, ΔEmp, ΔInc, and $\Delta Innov$ in equations 9-12 are the change in the population, employment, per capita personal income, and innovation as measured by the number of patent applications per capita, between 2009-2013. The vectors Ω^P, Ω^E, Ω^{Inc}, and Ω^{Innov}contain several exogenous variables that represent county-level characteristics at the initial period - year 2009 for all the exogenous variables except tax, revenue, and highway expenditures, which correspond to the year 2007. I follow Carlino & Mills (1987), Deller et al. (2001), Monchuk Miranowski (2010), Komarek & Loveridge (2015), and Rupasingha et al. (2002) to design these vectors that include different sets of regional characteristics. I classify the various regional characteristics broadly into four types:

Demand characteristics: Ethnic diversity, location- metro or nonmetro counties, income inequality, and share of expenditure in the construction and maintenance of highways.

Supply characteristics: Percent of population between 25 and 44 years; number of non-farm proprietors; firm size (distinguished between percent of firms with less than 100 employees and greater than 100 employees); concentration of high-tech firms – measured by the share of high-tech firms in total number of firms and the varieties of high-tech firms.

Government characteristics: Total tax per capita and the share of total county revenue earned from local, state, and federal governments

Innovation characteristics: Share of college educated population; expenditures in research and development by universities located in own and neighboring counties; and small business innovation research (SBIR) awards received by small firms.

The above grouping of regional characteristics by no means is assumed to contain mutually exclusive set of variables. For example, the share of college educated population is likely to determine the ability of firms to innovate and equally their ability to supply the goods and service demands in a region. Similarly, highway expenditure is as likely to represent regional demand as regional supply because improved transportation network connects the consumers to the broader regional markets and provides potential access to the substitutes to the goods produced by local firms while it might increase the supply efficiency of the local firms through reduction in transportation cost. But I assume that it affects more the ability of firms to satisfy demands than the regional demand within today's increasing trend of online shopping. Government characteristics and innovation characteristics are also equally likely to overlap with the demand and supply characteristics.

2.3.2. Hypotheses

The non-rivalrous nature of technology implied by the models of endogenous growth models (Aghion & Howitt, 1992; Grossman & Helpman, 1993) imply a link between the innovation and population growth. As the cost of invention is independent of the people benefitting from it (Arrow, 1962; Romer, 1990), growth in population implies technological progress (at the constant cost). On the other hand, the macroeconomic implication of Malthusian model in relation to technology and population (Galor & Weil, 2000; Malthus, 1959) is that the growth in population is limited by the level and growth of technology. Combining the implications of both these models, Kremer (1993) develops a model of population growth and empirically finds that initial level of population is directly proportional to the population growth and technological change. This background provides a basis for the test of my first hypothesis:

Hypothesis 1*: Regional growth in population and patenting rates positively influence each other.*

The link between innovation and employment is not always clear. It greatly depends on the nature of the technology. The innovation in labor-saving process technology of a firm instantly reduces its labor demand but the compensating effects may arise due to transfer of the improved productivity to the consumers in the form of reduced output prices thereby stimulating demand (Harrison et al., 2008). This is expected to generate positive employment effects in other firms in ancillary industries due to increase in their level of activities but negative effects in competing industries if the firms fail to survive from the technological competition (Spezia & Vivarelli, 2000). But, conceptualizing this process as Schumpeter's "creative destruction", I expect that exit of incompetent firms would set the stage for the entry of new innovative firms in the market, thus generating net positive employment effects. On the other hand, the innovation in the product side is expected to induce positive employment effects due to increase in demand for improved products. This effect might be weakened if the new products substitute the existing products in the market (Harrison et al., 2008). Also, similar compensating effects as in process innovation are likely to arise if the new product requires change in production methods. In this way, I expect the increase in innovation rates to generate higher level of employment. The growth in employment may lead to more economic activities, more competition, and the need for more innovations. This process cannot be perpetual but constrained by the growth stage of the economy, implying that higher employment may not necessarily lead to more innovation rates.

However, as my study covers the economic recovery period, I expect the growth in employment to positively drive growth in innovation rates. Accordingly, my second hypothesis is:

Hypothesis 2*: Regional growth in employment and patenting rates positively influence each other.*

Although commercialization of an innovation, in the long-run could be skill-saving as well as skill-biased, I argue that the innovation creation, as measured by patenting rates within a five-year period in my study, is mostly skill-biased. My argument is inspired again by the endogenous growth models where human capital in the form of educated and skilled people are needed to generate new economic knowledge. On the other hand, the wage inequality between college-educated workers and non-college workers, also known as "college premium", has been observed to have increased in the US[14] in the recent years. It is also observed that the within group wage inequality has also increased historically, Aghion (2002) argues in his model based on the Schumpeterian growth theory that the inequality is generated by the additional wage premium that is due to the reduced technological distance between the previous and current job of the group who get opportunities to learn by doing in innovative jobs. The combination of the idea from the endogenous growth theory and the observed wage premium for educated workers implies that workers in innovative firms enjoy wage premium over those in non-innovative firms. Consistent with this implication, improved labor productivity in innovative industries is argued to displace lower-paid unskilled jobs with better-paid ones (Harrison et al., 2008; Helpman, 2004; OECD, 1994). Building on these ideas, I expect that the growth in regional innovation enhances living standard of the regional population by means of income growth. I

[14] Autor et al., (1998) show that the ratio of number of "college-equivalent" and "non-college equivalent" workers grew from an average rate of 2.5% during 1940-1970 to 3.05% during 1970-1995. In the meantime, the ratio of weekly wage rates of these groups fell by 0.11% during 1940-70 but increased by 25% during 1970-1995.

also expect that the regions with higher income level provide more business opportunities to the innovative firms to through higher demand for improved goods and potentially through higher source of financial capital. Thus, my third hypothesis is:

Hypothesis 3: *Regional growth in patenting rates and per capita personal income positively influence each other.*

From the results on the hypothesis tests 1-3, I will infer whether the innovation belongs in the three-equation system of regional economic growth as modeled by Deller et al. (2001).

Combining the implications of the Malthusian model and the endogenous growth models discussed for laying out the first hypothesis, Kremer (1993) develops a model of population growth and empirically finds that initial level of population is directly proportional to the population growth and technological change. Combining this finding and the expected positive association of innovation growth with population, employment, and income growth in the hypotheses 1-3, I also expect the initial levels of population, employment, income, and innovation to have positive relation with their respective growth, for which reason the fourth hypothesis to be tested in this study is:

Hypothesis 4: *Initial levels of population, employment, income, and patenting rates are positively influence their respective growth rates.*

These hypotheses on the role of initial conditions provide the tests of regional convergence (or divergence) in terms of the measures of innovation and economic growth. Negative association means convergence, implying reducing regional gaps but positive association would suggest growing gaps.

Further, I test the predictions of the endogenous growth theory regarding the roles of human capital and knowledge spillover from the universities and the clustering of high-tech industries. The positive role of the human capital in innovation is straightforward from the implications of the models of endogenous growth and so is the role of universities in knowledge spillover (Jaffe, 1989; Mansfield, 1991). The positive role of proximity among firms in promoting knowledge spillover and innovation is undebated as knowledge spillover is defined as "working on similar things and hence benefiting much from each other's research" (Griliches, 1992). However, it is debatable whether such proximity refers to the firms within the same industry (specialization) or across firms in different industries (diversity)[15]. Building on Jacob's (1969) concept, I expect the diversity of high-tech industries to be conducive to innovation and economic growth. These lead to my next three hypotheses.

Hypothesis 5*: The share of regional population with college or higher education is positively related to innovation and economic growth.*

Hypothesis 6*: The expenditure in R&D by universities in a region is positively related to innovation and economic growth.*

Hypothesis 7*: The diversity of high-tech industries is positively related to the regional innovation and economic growth.*

2.3.3. Estimation

First, I estimated individual equations 9-12 separately using instrumental variable regression[16] to test endogeneity of the variables in each equation. I conducted an endogeneity test using Durbin and Wu-Hausman tests, where the null hypotheses are that the variables modeled as endogenous can be treated as exogenous. Failure to reject the null hypothesis implies that the

[15] See Glaeser et al. (1992) for the discussion of the concept on the role of industrial specialization and Jacob (1969) on the role industrial diversity in facilitating knowledge spillover and technological progress.

[16] I used Stata's ivregress command to run instrumental variable regression, and post-estimation command *estat endogenous* to conduct endogeneity tests (https://www.stata.com/manuals13/rivregresspostestimation.pdf)

ordinary least square (OLS) estimator of the equations provides consistent estimates. Alternatively, the rejection implies that the OLS estimates are inconsistent due to the correlation between the endogenous variables and the disturbances in the equations (Greene, 2003), and instrumental variable techniques are required to account such correlation.

Following the evidence of endogeneity in each equation, which I will discuss in the following results section, I estimate a structural model of county growth represented by the system of equations 9–12 using three-stage least square regression (3SLS)[17] to analyze the interdependence among the innovation and economic growth variables (hypothesis tests 1-3). The 3SLS estimator also improves the efficiency of the parameters across the equations, which are likely to be correlated through some unobservables in the equations (Wooldridge, 2010) such as the propensity to patent or the perception of the businesses about the potential of the regions for market growth, or the risk-seeking entrepreneurial cultures of the regions. The correlation may arise, for example, by the simultaneous effects of the unobservable variable representing the entrepreneurial culture on the employment growth and innovation growth.

In estimating 3SLS regression models, the variables that are excluded from each equation are so chosen as to get an identified system of equations satisfying the exclusion restrictions[18], using the Sargan-Hansen test[19]. The choice of the excluded variables is made based on these variables' higher correlation with the variables which they serve as instruments for but are likely to be uncorrelated with the disturbance terms. The validity of the instruments or the overidentifying exclusion restrictions are tested using the Hansen's (1996) test for

[17] I used Stata's *reg3* for my analysis in this study (https://www.stata.com/manuals13/rreg3.pdf)

[18] For an (over-identified) idendified model, the number of variables excluded from an equation should be (greater than) equal to the number of endogenous variables (see Wooldridge, 2010).

[19] In Sargan-Hansen test, null hypothesis is that the instruments are valid instruments (uncorrelated with the disturbance term) and the excluded instruments are correctly excluded from the estimated equations.

heteroskedastic disturbances. Under the joint null hypothesis that the excluded instruments are valid (uncorrelated with disturbances) or the excluded instruments are correctly excluded, the Hansen's Statistics is chi-squared distributed. The failure to reject the null hypothesis satisfies the overidentifying restrictions.

Finally, I estimate reduced forms of the structural coefficient estimates from 3SLS method to test the remaining hypotheses 4-7. The reduced form estimates are obtained by regressing each dependent variable in the equations 9-12 on the set of exogenous variables in the system, or equivalently they are the coefficient estimates from the first stage estimation of 3SLS. The reduced forms of the structural coefficients include both the direct effects and indirect effects arising from interdependence among the endogenous variables (Carlino & Mills, 1987).

2.4. Data

The empirical model is estimated using data for a sample of 3,038 counties in the 48 contiguous states of the United States. Secondary data are collected from several sources for the period during 2009-13. Table 1 provides the specific sources of data for the variables used in this study, their definition, and summary statistics. The counties are classified into metro and nonmetro categories according to the Rural-Urban Continuum Codes (RUCC), 2013 developed by the Economic Research Service of USDA[20]. For the analysis in this study, metro counties represent the "urban" regions and the non-metro counties the "rural" regions.

The number of the domestic utility patent applications per 10K population serves as my measure of the rate of innovation in the US counties. I aggregated the patent applications

[20] 2013 RUCC are accessible at: https://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx
Metro areas include all counties containing one or more urbanized areas: high-density urban areas containing 50,000 people or more; metro areas also include outlying counties that are economically tied to the central counties, as measured by the share of workers commuting on daily basis to the central counties. Non-metro counties are outside the boundaries of metro areas and have no cities with 50,000 residents or more.

originating from residential zip codes of the primary inventors to derive the county level patent applications. The university R&D expenditures data that come from NSF are available at city level. I matched the university cities with their associated counties. For example, if either a county does not have any city with college or university or the institutions of higher learning do not spend in R&D activities, I assume in this study that the county has zero university R&D. I also account for the spillover effects of the university R&D from the counties hosting the university/colleges to their neighboring counties by constructing a spatially lagged university R&D variable based on a distance decay function within 100 miles from the county centroids.

I employed the firms-related data including the variety of high-tech industries from US Census Bureau's Community Business Patterns (CBP). The number of all kinds of business establishments and their employments were derived by summing these variables at the three-digit level of industry codes across the 2012 North American Industry Classification System (NAICS), 2012. I derived the number of high-tech establishments by summing at the six-digit level of industry codes across the 2012 NAICS codes that constitute high-tech industries, as defined by NSF. The data on foreign-born population and the population with college or higher degree come from the five-year estimates for 2009 of American Community Survey (ACS)[21]. I combine the population of "naturalized US citizen" and "not a US citizen" to derive the foreign-born population variable.

[21] ACS surveys 295,000 households randomly each year with no repeated household in five years and reports the estimates from data collected in five years.
https://www.census.gov/content/dam/Census/programs-surveys/acs/about/ACS_Information_Guide.pdf

Table 2.1 Variables Definition, Summary Statistics, and Data Source

Variable Code	Description	Mean	SD	Source
nonmetro	RUCC 2003 county type (0=metro; 1=non-metro)	N/A	N/A	ERS, USDA
Δpopulation	Change in population (1k), 2009-13	3	14.51	BEA
Δemployment	Change in employment (1k), 2009-13	2.54	14.39	BEA
Δincome	Change in per capita personal income ($1k), 2009-13	5.81	5.17	BEA
Δinnovation	Change in patents per 10k population, 2009-13	0.32	3.96	USPTO
lagged_population	Initial population (1k), 2009	96	308.01	BEA
lagged_innovation	Initial patents per 10k Population, 2009	2.3	30.1	USPTO
lagged_employment	Initial Employment (1k), 2009	54.72	191.13	BEA
lagged_PCPI	Initial Per Capita Personal Income ($1k), 2009	32.43	7.65	BEA
nfarm_propri	Number of Non-Farm Proprietors (1k), 2009	11.16	41.06	BEA
urd	University R&D expenditures ($1M), 2009	13.17	83.72	NSF
splag_urd	Spatial lag of university R&D expenditures ($1M), 2009	14.9	60.44	NSF
sbir	SBIR awards ($1M), 2009	0.61	5.2	SBIR
pct_ht_estabs	Percent of total establishments in hi-tech industries, 2009	5.07	2.71	CBP
ht_variety	Variety of Hi-tech Industries, 2009	13.39	9.64	CBP
ethnic_diversity	Ethnic Diversity by Race and Ethnicity, 2009	0.46	0.27	MCDC
taxpercap	Total taxes per capita ($1k), 2007	1.31	1.24	US Census
pct_ig_rev	Percent of total revenue from federal, state, and local governments, 2007	41.79	13.77	US Census
pct_highway_expend	Percent of total expenditure in highways, 2007	6.09	4.42	US Census
pct_collegeplus	Percent of population with college or higher degree, 2009	8.01	3.41	ACS
pct_foreign_born	Percent of foreign-born population, 2009	4	5.02	ACS
income_ineq	Gini Index of income distribution, 2010	0.43	0.04	ACS
estct_pct_nlt100	Percent of total establishments with 1-19 employees, 2009	89.09	3.75	CBP
est_pct_ngt100	Percent of total establishments with greater or equal to 100 employees, 2009	4.12	10.12	CBP
pct_pop_25to44	Percent of total population aged 25-44, 2009	0.22	0.94	ACS

The derivation of concentration measure of high-tech industries similar to HHI measure, which is widely adopted in the literature for representing industrial concentration (see Carlino et al., 2007), does not allow us to distinguish between the counties with zero employment and those with a completely specialized industry, as several counties in this study sample have zero high-tech employment. So, I created a high-tech variety variable that measures the number of three-digit level NAICs high-tech industry categories. Ranging in its value from 1 to 45, this variable essentially represents the concentration of high-tech industries after controlling the share of high-tech industries in total industrial employment and avoiding the zero-employment problem.

I constructed the ethnic diversity variable is by creating an entropy index similar to Theil Index (Audretsch et al., 2010):

$$Ethnic\_Diversity_{i,t} = -\sum_{r=1}^{R_i} s_{irt} \ln(s_{irt})$$

where s_{rit} is the share of population identified as race r in region i at a time t, where $R_i \in$ {White, Black, Asian & Pacific Islander, American Indian, Other}.

The entropy index could reach its maximum ($\ln(R_i)$ or $\ln(5)$ $in\ our\ case$) at s_{irt} = $1/R_i$and its minimum (0) at s_{irt}=1, that is when a single race forms the entire population of a region. The index measures the share and the variety of the races in the population. I used income inequality data represented by Gini Index from ACS for the census year 2010.

2.5. Results

From the Durbin and Wu-Hausman endogeneity tests following the instrumental regression of equation 9, I found that patents growth and employment growth have an endogenous relationship with population growth, but I do not find statistical evidence of endogeneity of PCPI growth. From the similar tests in equation 10, I found that population growth was endogenous with employment growth, while I did not find the statistical evidence of the

Table 2.2 3SLS Results of the Estimation of the County Growth Model

Variables	Coefficient Estimates			
	ΔPopulation	ΔEmployment	ΔPCPI	ΔPatents
ΔPopulation (1k people)		0.134**	-0.224***	0.156***
		[2.39]	[4.27]	[2.95]
ΔEmployment (1k jobs)	0.162*		0.597***	-0.131***
	[1.79]		[5.12]	[3.73]
ΔPCPI ($1k)	-0.052	0.236***		-0.058
	[0.07]	[2.82]		[0.90]
ΔPatents (per 10k pop.)	-1.830***	-0.005	-0.104	
	[4.61]	[0.04]	[0.80]	
Lagged population (1k people)	-0.008*	-0.028***	0.019***	-0.007***
	[1.75]	[14.85]	[5.36]	[5.77]
Lagged employment (1k jobs)	0.086***	0.015**	0.005	0.002
	[7.41]	[2.45]	[0.89]	[0.47]
Lagged PCPI ($1k)	-0.086	-0.093***	0.155***	0.039*
	[0.83]	[3.75]	[8.47]	[1.95]
Lagged patents (per 10k pop.)	-0.102***	-0.003	-0.01	-0.056***
	[4.68]	[0.35]	[1.32]	[24.54]
Nonmetro (1=yes; 0=metro)	-1.313	-0.4	0.932***	-0.216
	[1.44]	[1.57]	[4.66]	[1.13]
Percent collegeplus (% points)	0.417***	0.009	-0.056	0.036
	[3.06]	[0.20]	[1.39]	[1.02]
Percent foreign born (% points)	0.419***	0.058**	0.096***	-0.014
	[5.87]	[1.98]	[3.83]	[0.62]
Tax per capita ($)	-1.936***			
	[5.19]			
Pct. highway expend. (% points)	0.105		0.186***	
	[0.70]		[10.35]	
Pct. intergov. Revenue (% points)	-0.081**		-0.051***	
	[2.17]		[8.37]	
Pct. est. w/ <100 emp (% points)	-0.712**	0.045	0.141	0.281***
	[2.04]	[0.35]	[1.31]	[2.66]
Pct. est. w/ >100 emp (% points)	-0.412***	0.246***	-0.241***	0.220***
	[4.08]	[4.59]	[4.47]	[4.89]
Pct. pop. age25-44 (% points)	1.167***	-0.097	0.077	0.005
	[3.81]	[1.04]	[0.97]	[0.07]
Income inequality (Ginni Index)	-1.783	-1.962		
	[0.37]	[0.83]		
Pct. high-tech firms (% points)	0.621***	0.039	0.202***	0.03
	[4.37]	[0.80]	[4.96]	[0.85]
High-tech variety	-0.074	-0.078***	-0.108***	-0.004
	[0.57]	[3.52]	[5.32]	[0.24]
Ethnic Diversity (Theil Index)		-0.562		
		[1.53]		

Table 2.2 (cont'd)

Non-farm proprietors (1k)		0.384***	-0.242***	
		[22.37]	[5.20]	
SBIR Awards		-0.062***	0.012	-0.028*
		[3.06]	[0.61]	[1.81]
University R&D ($1M)				0.010***
				[8.13]
SPLAG univ R&D ($1M)				0.002
				[1.31]
constant	73.579**	-1.105	-15.803	-29.088***
	[2.10]	[0.08]	[1.47]	[2.77]
R^2	0.55	0.9	0.07	0.2
N	3,038	3,038	3,038	3,038

endogeneity of the growth in PCPI and patenting rates. In equation 11, I found that population, employment, and patents growth were jointly endogenous, but not individually, with PCPI growth. Similarly, in equation 12, I found that population, employment, and PCPI growth were jointly endogenous, but not individually, with patent growth.

2.5.1. Regional Innovation and Economic Growth Interdependence

To test the hypotheses 1-3, I analyze the 3SLS estimation results on the structural coefficients in table 2. The positive and significant coefficient of *ΔPopulation variable* in *ΔPatents* equation (column 4), and negative and significant coefficient of *ΔPatents* variable in *ΔPopulation* equation (column 1) show that relationship between regional growth in population and patenting rates is highly interactive. These findings suggest that growth in population leads to growth in regional innovation, but less innovative regions may experience relatively faster growth in population.

Turning to my hypothesis 2, I find that employment growth negatively influences regional patenting rate, but I do not find statistical evidence of influence of patenting growth on regional employment. These findings suggest decreasing marginal patent productivity of employees, or more frequent patenting by smaller firms, and offsetting of the job displacement effects of the labor-saving innovations with the job creation effects of the product improvement innovations. However, I do not find statistical evidence for the support of the hypothesis 3 regarding either the influence of innovation growth on income growth or the influence of income growth on innovation growth.

Combined, these findings show that innovation belongs to the regional growth ecosystem and the results from a growth study that does not account for this innovation effect is likely to suffer from specification bias (omitted variable bias in case of missing innovation variable and endogeneity in system of equations method). Beside my focus on my principal research questions, the results in table 2 show that population growth is directly proportional to employment growth, suggesting that people move to the regions with more employment opportunities and the increase in labor supply stimulates business growth. But the increasing supply of labor (also the skilled and educated workforce) is likely to have downward pressure on the PCPI growth due from lower wages.

2.5.2. Exogenous Factors of Regional Innovation and Economic Growth

Turning to the test of hypothesis 4, my findings from the reduced forms of the structural coefficients presented in table 3 show that the lower initial level of patenting rates is associated with higher patenting growth, suggesting that regional gap in inventive activities is shrinking. Also, I find that lower initial regional patenting rate is associated with higher growth in employment and PCPI. Combined, I find that less innovative regions experienced higher growth rates in population, inventive activities, and employment.

On the other hand, my results show that higher initial levels of PCPI lead to higher growth in PCPI but lower growth in population and employment. These results suggest that prosperous regions generated relatively fewer number of jobs (but likely high-paying) and lower inflow of people. Combinations of these implications with the possible convergence in regional innovation rates (preceding paragraph) and better-paid jobs in innovative firms suggest the possibility of wage discrimination by the innovative firms based on the condition of regional prosperity.

Turning to my hypothesis 5, I find that an increase in the share of population with four-year bachelor or higher degree is positively associated with the growth in population and inventive activities. Interestingly, the variable is negatively associated with the growth in PCPI. It is likely that some regions were not likely in situations to fully absorb the supply of fresh graduates during the recovery period following the great depression thus putting downward pressure on the wages earned by those fresh degree holders. I find that the expenditure by universities in R&D is positively associated with patenting rates (positive and statistically significant coefficient of *University R&D* in ΔPatents equation, column 4), thus supporting my hypothesis 6 and suggesting that university research plays an important role in accelerating innovative activities.

The positive and statistically significant coefficient of the *Percent high-tech firms* in all four columns shows that high-tech businesses play a major role in generating innovations and economic growth. Further the negative and statistically significant coefficient of *high-tech variety* in the *ΔEmployment* and *ΔPCPI* equations, columns 2 and 3 respectively) suggests that the externalities generated from the R&D and other knowledge spillover are higher in regions with more specialized industries and are manifested in growth of regional employment and income. However, I do not find any evidence for the significant influence of such externality on growth in inventive activities.

Table 2.3 Reduced Form Estimates of the Parameters in the County Growth Model

	Coefficient Estimates			
Variables	ΔPopulation	ΔEmployment	ΔPCPI	ΔPatents
Lagged population (1k people)	-0.004	-0.027***	0.004***	-0.005
	[0.27]	[2.88]	[2.63]	[1.00]
Lagged employment (1k jobs)	0.073**	0.026	0.002	0.01
	[2.01]	[1.36]	[0.76]	[1.05]
Lagged PCPI ($1k)	-0.134***	-0.070**	0.138***	0.02
	[4.00]	[2.36]	[3.45]	[0.89]
Lagged patents (per 10k pop.)	-0.001	-0.005**	-0.007***	-0.056***
	[0.18]	[2.10]	[4.57]	[3.83]
Nonmetro (1=yes; 0=metro)	-0.716***	-0.27	0.929***	-0.376**
	[2.94]	[1.53]	[6.08]	[2.03]
Percent collegeplus (% points)	0.280***	0.025	-0.117***	0.080**
	[4.54]	[0.52]	[2.59]	[2.38]
Ethnic Diversity (Theil Index)	-0.028	-0.736**	-0.720***	-0.317*
	[0.05]	[2.18]	[2.67]	[1.70]
Percent foreign born (% points)	0.388***	0.141***	0.093***	0.036*
	[4.59]	[2.63]	[4.25]	[1.80]
Tax per capita ($)	-1.504***	-0.085	0.299**	-0.262
	[3.75]	[0.37]	[2.08]	[1.54]
Pct. highway expend. (% points)	0.079***	0.019	0.191***	0.006
	[3.41]	[1.34]	[5.92]	[0.44]
Pct. intergov. Revenue (% points)	-0.072***	-0.018	-0.046***	-0.001
	[4.24]	[1.63]	[6.64]	[0.18]
Income inequality (Ginni Index)	-4.616	-4.254*	1.432	1.716
	[1.22]	[1.74]	[0.62]	[0.94]
Pct. est. w/ <100 emp (% points)	-0.944**	-0.014	0.339***	0.113
	[2.58]	[0.06]	[4.10]	[0.92]
Pct. est. w/ >100 emp (% points)	-0.593**	0.141	-0.028	0.111*
	[2.22]	[0.80]	[1.10]	[1.71]
Pct. pop. age25-44 (% points)	0.866***	-0.043	-0.144	0.179
	[3.49]	[0.29]	[0.43]	[1.55]
Pct. high-tech firms (% points)	0.490***	0.142**	0.172***	0.091**
	[5.06]	[2.34]	[3.68]	[2.27]
High-tech variety	-0.102	-0.131***	-0.164***	0.008
	[1.62]	[2.97]	[10.53]	[0.56]
Non-farm proprietors (1k)	0.125	0.396***	-0.032***	-0.031
	[0.98]	[5.71]	[3.51]	[1.34]
University R&D ($1M)	-0.014	0.0001	0.001*	0.008***
	[1.56]	[0.03]	[1.94]	[3.15]
SPLAG univ R&D ($1M)	-0.007	-0.010**	-0.002**	0.002
	[1.16]	[2.09]	[2.23]	[1.23]
SBIR Awards	-0.031	-0.089	-0.023*	-0.033
	[0.20]	[0.89]	[1.92]	[0.90]

Table 2.3 (cont'd)

constant	99.863***	5.994	-35.033***	-13.015
	[2.73]	[0.24]	[4.26]	[1.06]
R^2	0.73	0.88	0.28	0.25
N	3,038	3,038	3,038	3,038

The t-stats in the square brackets are based on the robust standard errors; * p<0.1; ** p<0.05; *** p<0.01

Additional Findings on Regional Growth Drivers: The regions with higher racial diversity are associated with statistically significant decline in the number of jobs, PCPI, and inventive activities. My findings in terms of employment growth are consistent with those of Deller et al. (2001) but contrast with those of Carlino & Mills (1987) in that the former study found positive association between the percent black population, as a proxy to racial diversity, and employment but the latter study found negative association. However, in terms of income growth, my results contrast with Deller et al.'s (2001) finding of positive association between the percent black population and the income growth.

The findings of this study show that the share of foreign born population is positively associated with the growth in population, employment, PCPI, and inventive activities. I also find that the regions with higher total taxes are found to hinder population growth but support the income growth, probably through the externalities due to the higher spending on public goods. In terms of income growth, my findings are consistent with both Carlino and Mills' (1987) and Deller et al.'s (2001) findings in that the former study found the local taxes per capita to be negatively associated with population growth and the latter study found negative association between the property tax and population growth. However, my finding of positive association with income growth contrasts with Deller et al.'s finding of negative association.

The findings from table 6 show that the higher expenditures in highways and road networks attract more population and increase income levels likely because of the increased access to the larger urban areas and labor markets. Further, I find that the share of revenues received from government agencies is found to have negative association with population growth and income growth.

My findings on the relationship of the firm size show that the increase in the number of businesses, whether small or large, are associated with decline in population level, implying people's preference to live in places farther from the industrial areas. However, the larger share of small businesses is associated with increase in PCPI. The places with larger share of prime working age population are found to attract more population. Further, non-farm proprietors are found to create more but low-paying jobs, as shown by the significantly positive association of the variable with the employment growth but negative with the PCPI growth.

2.6. Summary and Conclusion

Theoretical and empirical consensus shows that innovation enhances economic growth at various geographical levels. It is equally likely that firms in growing regions are likely to have more resources out of higher profits to expend to the innovative activities and employ innovative workforce and such regions might provide better access for these firms to the financial resources and collaborations needed for innovation. So, it is likely that growing regions enhance innovative activities. As existing literature presents evidence of simultaneity among population, employment, and income growth, it could be possible that innovation growth occurs simultaneously with one or more of the economic growth variables. To investigate this possibility, I extended the three-equation simultaneous equation model in the literature for population, employment, and income to four-equation model, where innovation is endogenously determined in the regional economy along with economic growth variables. The extended model includes several variables identified in the literature to be related to economic growth and innovation.

Using the Durbin and Wu-Hausman tests, I found that regional innovation and economic growth exhibit an endogenous relationship. From the 3SLS estimation results, I found that growth in population leads to growth in regional innovation, but less innovative regions may experience relatively faster growth in population.   I also found that employment growth negatively influences regional patenting rate, but there is no statistical evidence of influence of patenting growth on regional employment, suggesting decreasing marginal patent productivity of employees.

My findings from the reduced form coefficients of the 3SLS estimates indicate that the lower initial level of inventive activities, measured by the patents per capita, are associated with higher growth in such activities, suggesting convergence between the leading and lagging regions in terms of inventive activities in the longer run. Compared to metro regions, growth in population and number of patent applications is significantly lower in the non-metro regions but the growth in income levels is higher.

My major findings show that foreign-born population and high-tech firms in higher regionally concentrated industries are associated positively with both innovation and economic growth. Combined with the findings on the simultaneous relationship, these results support the idea that that policies to promote regional clusters of high-tech firms and capitalize on the knowledge potential of the immigrants is likely to reinforce regional innovation and economic growth.

REFERENCES

REFERENCES

Acs, Z. J., & Audretsch, D. B. (2005). *Entrepreneurship and innovation* (No. 2105). Papers on Entrepreneurship, Growth and Public Policy.

Acs, Z. J., Audretsch, D. B., & Feldman, M. P. (1992). Real effects of academic research: comment. *The American Economic Review*, *82*(1), 363-367.

Acs, Z. J., Audretsch, D. B., Braunerhjelm, P., & Carlsson, B. (2012). Growth and entrepreneurship. *Small Business Economics*, *39*(2), 289-300.

Adelaja, S., Hailu, Y. G., & Abdulla, M. (2009). New Economy Growth Decomposition in the US. In Selected paper prepared for presentation at the American Agricultural Economics Association Annual Meeting.

Aghion, P. (2002). Schumpeterian growth theory and the dynamics of income inequality. *Econometrica*, *70*(3), 855-882.

Aghion, P., & Howitt, P. (1990). *A model of growth through creative destruction* (No. w3223). National Bureau of Economic Research.

Anselin, L., Varga, A., & Acs, Z. (1997). Local geographic spillovers between university research and high technology innovations. *Journal of urban economics*, *42*(3), 42448.

Arrow, K. J. (1972). Economic welfare and the allocation of resources for invention. In *Readings in Industrial Economics* (pp. 219-236). Palgrave, London.

Audretsch, B. (1998). Agglomeration and the location of innovative activity. *Oxford review of economic policy*, 14(2), 18-29.

Audretsch, D. B., & Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production. *The American economic review*, *86*(3), 630-640.

Audretsch, D., Dohse, D., & Niebuhr, A. (2010). Cultural diversity and entrepreneurship: a regional analysis for Germany. *The annals of regional science*, *45*(1), 59-125.

Bartik, T. J. (1985). Business location decisions in the United States: Estimates of the effects of unionization, taxes, and other characteristics of states. *Journal of Business & Economic Statistics*, *3*(1), 14-22.

Bartik, T. J. (1991). *Who benefits from state and local economic development policies?* Michigan: W.E. Upjohn Institute for Employment Research.

Brenner T. (2004). *Local Industrial Clusters: Existence, Emergence, and Evolution*. Routledge, London.

Capello, R., & Lenzi, C. (2013). Innovation and employment dynamics in European regions. *International Regional Science Review*, 36(3), 32353.

Carlino, G. A. (1985). Declining city productivity and the growth of rural regions: a test of alternative explanations. *Journal of Urban Economics*, *18*(1), 11-27.

Carlino, G. A., & Mills, E. S. (1987). The determinants of county growth. Journal of Regional Science, 27(1), 39-54.

Carlino, G. A., Chatterjee, S., & Hunt, R. M. (2007). Urban Density and the Rate of Invention. *Journal of Urban Economics, 61*(3), 389-419.

Carree, M. A., & Thurik, R. (2003). The Impact of Entrepreneurship on Economic Growth. In Audretsch, D.B, & Acs, Z.J. (Eds.), *Handbook of Entrepreneurship Research*, (437–471). Boston/Dordrecht: Kluwer-Academic Publishers.

Clark, D. E., & Murphy, C. A. (1996). Countywide employment and population growth: An analysis of the 1980s. *Journal of Regional Science*, *36*(2), 235-256.

Cooke, P. (1992). Regional innovation systems: competitive regulation in the new Europe. *Geoforum*, *23*(3), 365-382.

Deller, S. C., Tsai, T. H., Marcouiller, D. W., & English, D. B. (2001). The role of amenities and quality of life in rural economic growth. *American Journal of Agricultural Economics*, *83*(2), 35365.

Eaton, J., & Kortum, S. (1997). Engines of growth: Domestic and foreign sources of innovation. *Japan and the World Economy*, *9*(2), 235-259.

Ezell, S. J., & Atkinson, R. D. (2010). The good, the bad, and the ugly (and the self-destructive) of innovation policy: A policymaker's guide to crafting effective innovation policy. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1722845. Accessed 02/27/2017.

Ezell, S.J., (2009). America and the World—I're No. 40! Democracy, A Journal of Ideas, Issue 14. Available at http://democracyjournal.org/magazine/14/america-and-the-world-were-40/. Accessed 06/01/2017.

Fagerberg, J. (1994). Technology and international differences in growth rates. *Journal of economic Literature*, *32*(3), 1147-1175.

Fagerberg, J. (2002). *Technology, Growth and Competitiveness: Selected Essays*. Cheltenham: Edward Elgar.

Fagerberg, J. (2003). Schumpeter and the revival of evolutionary economics: an appraisal of the literature. *Journal of evolutionary economics*, *13*(2), 125-159.

Fagerberg, J. (2006). Innovation: A guide to literature. In Fagerberg, J., Mowery, D.C., & Fagerberg, J., & Verspagen, B. (2002). Technology-gaps, innovation-diffusion and transformation: an evolutionary interpretation. *Research policy*, *31*(8-9), 1291-1304.

Federal Reserve Bank of Dallas (1996). *Annual Report, 1996*. Available at https://www.dallasfed.org/fed/annual. Accessed 04/12/2017.

Feldman, M. P. (1994). *The Geography of Innovation*. Boston: Kluwer Academic Publishers.

Feldman, M. P., & 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), 210-229.

Feser, E., & Isserman, A. (2006). Harnessing growth spillovers for rural development: The effects of regional spatial structure. *Report to USDA Rural Development, University of Illinois at Urbana-Champaign*.

Frenken, K., van Oort, F., & Verburg T. (2007). Related variety, unrelated variety, and regional economic growth. *Regional Studies*, 41, 685–697.

Galor, O., & Weil, D. N. (2000). Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond. *American economic review*, *90*(4), 806-828.

Glaeser, E.L., Kallal, H.D., Scheinkman, J.A., & Shleifer, A. (1992). Growth of cities. *Journal of Political Economy*, 100, 1126-1152.

Griliches, Z. (1992). The search for R&D spillovers. *Scandinavian Journal of Economics*, 94, 29-47.

Grossman, G. M., & Helpman, E. (1993). *Innovation and growth in the global economy*. MIT press.

Greene, W. (2003). *Econometric Analysis*. New Jersey: Prentice Hall.

Hammond, G. W., & Thompson, E. C. (2008). Determinants of income growth in metropolitan and nonmetropolitan labor markets. *American Journal of Agricultural Economics*, *90*(3), 783-793.

Hansen, B. E. (1996). Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica: Journal of the econometric society*, 413-430.

Harrison, R., Jaumandreu, J., Mairesse, J., & Peters, B. (2008). Does Innovation Stimulate Employment? A Firm-Level Analysis Using Comparable Micro-Data from Four European Countries. ZEW Discussion Papers, No. 08-111.

Helms, L. J. (1985). The Effect of State and Local Taxes on Economic Growth: A Time Series--Cross Section Approach. *The Review of Economics and Statistics*, 574-582.

Helpman, E. (2004). *The Mystery of Economic Growth.* Cambridge: Belknap Press.

Henderson, J., & Weiler, S. (2010). Entrepreneurs and job growth: Probing the boundaries of time and space. *Economic Development Quarterly*, *24*(1), 23-32.

Howells, J. (2005). Innovation and regional economic development: A matter of perspective? *Research policy*, *34*(8), 1220-1234.

Jacobs, J. (1969). *The Economy of Cities*. New York: Random House.

Jaffe, A. B. (1989). Real effects of academic research. *The American economic review*, 957-970.

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *the Quarterly journal of Economics*, *108*(3), 577-598.

Komarek, T., & Loveridge, S. (2015). Firm sizes and economic development: Estimating long-term effects on US county growth, 1990–2000. *Journal of Regional Science*, *55*(2), 26279.

Kremer, M. (1993). Population growth and technological change: One million BC to 1990. *The Quarterly Journal of Economics*, *108*(3), 681-716.

Krugman, P. (1991). *Geography and Trade*. Cambridge: MIT Press.

Kuznets, S. (1971). *Economic Growth of Nations: Total Output and Production Structure*. Harvard: Belknap.

Lucas Jr, R. E. (1988). On the mechanics of economic development. *Journal of monetary economics*, *22*(1), 3-42.

Malthus, T. R. (1959). *Population: the first essay* (Vol. 31). Ann Arbor: University of Michigan Press.

Mandel, M.J. (2004). *Rational Exuberance: silencing the enemies of growth and why the future is better than you think*. New York: HarperBusiness.

Mansfield, E. (1991). Academic research and industrial innovation. *Research policy*, *20*(1), 1-12.

Marshall, A. (1990). *Principle of Economics*. London: Macmillan.

McGranahan, D. A., Wojan, T. R., & Lambert, D. M. (2010). The rural growth trifecta: outdoor amenities, creative class and entrepreneurial context. *Journal of Economic Geography*, *11*(3), 529-557.

Mills, E. S., & Price, R. (1984). Metropolitan suburbanization and central city problems. *Journal of Urban Economics*, *15*(1), 1-17.

Monchuk, D. C., & Miranowski, J. A. (2010). The impacts of local innovation and innovative spillovers on employment and population growth in the US Midwest. *Journal of Regional Analysis & Policy*, *40*(1), 61.

Nadiri, M. I. (1993). *Innovations and technological spillovers* (No. w4423). National Bureau of Economic Research.

OECD (1994). *The OECD Jobs Study: Facts, Analysis, Strategy*. Available at http://www.oecd.org/dataoecd/42/51/1941679.pdf. Accessed 06/24/2017.

Partridge, M. D., Rickman, D. S., Ali, K., & Olfert, M. R. (2008). Employment growth in the American urban hierarchy: long live distance. *The BE Journal of Macroeconomics*, *8*(1).

Pede, V. O. (2013). Diversity and regional economic growth: Evidence from US counties. *Journal of Economic Development*, *38*(3), 111.

Plaut, T. R., & Pluta, J. E. (1983). Business climate, taxes and expenditures, and state industrial growth in the United States. *Southern Economic Journal*, 99-119.

Porter, M. E. (1990). *The Competitive Advantage of Nations*. New York: Free Press.

Posner, M. V. (1961). International trade and technical change. *Oxford economic papers*, *13*(3), 323-341.

Romans, T., & Subrahmanyam, G. (1979). State and local taxes, transfers and regional economic growth. *Southern Economic Journal*, 435-444.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, 94, 1002–1037.

Romer, P. M. (1990). Endogenous technological change. *Journal of political Economy*, *98*(5, Part 2), S71-S102.

Rosenthal, S. S. & Strange, W. C. (2001). The determinants of agglomeration. *Journal of Urban Economics*, 50, 191–229.

Rupasingha, A., Goetz, S. J., & Freshwater, D. (2002). Social and institutional factors as determinants of economic growth: Evidence from the United States counties. *Papers in regional Science*, *81*(2), 139-155.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The quarterly journal of economics*, *70*(1), 65-94.

Spezia, M., & Vivarelli, M. (2000). US high-skilled immigration, innovation, and entrepreneurship: Empirical approaches and evidence. In Vivarelli, M., and Pianta, M.(Eds), *The Employment Impact of Innovation – Evidence and Policy* (pp. 12–25). Routledge.

Steinnes, D. N., & Fisher, W. D. (1974). An econometric model of intraurban location. *Journal of Regional Science*, *14*(1), 65-80.

Stephens, H. M., Partridge, M. D., & Faggian, A. (2013). Innovation, entrepreneurship, and economic growth in lagging regions. *Journal of Regional Science*, *53*(5), 778-812.

Wennekers, S., & Thurik, R. (1999). Linking entrepreneurship and economic growth. *Small business economics*, *13*(1), 27-56.

Wheat, L. F. (1986). The determinants of 1963–77 regional manufacturing growth: Why the South and West grow. *Journal of Regional Science*, *26*(4), 635-659.

Wong, P. K., Ho, Y. P., & Autio, E. (2005). Entrepreneurship, innovation and economic growth: Evidence from GEM data. *Small business economics*, *24*(3), 335-350.

Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data. MIT press.

Yeager, T. (2018). Institutions, transition economies, and economic development. Routledge.

Young, A. T., Higgins, M. J., Lacombe, D. J., & Sell, B. (2014). The direct and indirect effects of Small Business Administration lending on growth: Evidence from US county-level data (No. w20543). National Bureau of Economic Research.

ESSAY 3: EXPLORING INNOVATION CREATION ACROSS RURAL AND URBAN FIRMS: ANALYSIS OF THE NATIONAL SURVEY OF BUSINESS COMPETITIVENESS

3.1. Introduction

Innovative firms are essential to sustain economic growth. Established firms must continuously innovate to survive the forces of creative destruction in the face of new and disruptive technologies. Innovation also serves as a mechanism for new firm entry into emerging markets and enables these new entrants to compete with existing firms as well as other new entrants (Christensen, 2013; Schumpeter, 1942). The literature on regional innovation is primarily focused on urban innovation and based on firm data from urban areas that foster innovation creation and adoption; less studied is rural innovation and potential differences in innovation drivers between rural and urban areas (Dabson, 2011). Of the studies comparing rural and urban innovation, many conclude that rural America lags in its innovation performance (Orlando & Verba, 2005; Porter et al., 2004; Wojan et al., 2015).

Economies need innovation-based entrepreneurship to achieve and sustain growth (Mann & Shideler, 2015), and competitiveness of the overall US economy builds on the rural-urban interdependency (Dabson, 2007, 2011). The rural-urban innovation gap has long-term consequences. For example, lower education rates and fewer economic opportunities for youth lead to sluggish wealth creation which in turn contribute to the persistence of rural poverty (Lyons et al., 2018; Orlando & Verba, 2005; Porter et al., 2004; Ratner & Markley, 2014).

Innovation in urban areas is generally explained in terms of the agglomeration effect supported by the higher population density as well as higher concentrations and diversity of firms and industries in these areas (Carlino et al., 2007; Glaeser et. al., 1992; Orlando & Verba, 2005). Urban agglomeration facilitates the urban firms' opportunities to capitalize on their scale economies through enhanced communication and knowledge spillovers among innovative firms and industries, better supply of critical innovation resources such as human capital, extended buyer and supplier networks, and financial and professional support services (Aryal, et al., 2018; Orlando & Verba, 2005). On the other hand, scattered populations and less developed markets in rural areas restrict the opportunities for innovation by rural firms.

Related to the locational obstacles to innovation, rural firms have lower levels of skilled managers, professionals, and technicians and rural entrepreneurs are more likely to start new businesses based on necessity rather than opportunity, which frequently leads to a non-innovative enterprise that may be abandoned when better paying jobs arise (Acs, 2006; Henderson, 2002). Rural firms are also less likely to be growth-oriented, which may be attributed to owner characteristics such as embracing of the multi-generational business ownership models or the tendency to avoid the risk associated with adopting and/or creating innovation (Knickel et al., 2009; Renski & Wallace, 2012). Such business models are less likely to attract equity and venture capital due to reduced interest or flexibility in potential exit strategies, a necessity for innovative startups (Markley, 2001).

In terms of policy obstacles, rural economies are often framed as primarily agriculture-dependent, with substantial public resources focused on cost-saving technologies for agriculture production (Mowery et al., 2010; Stauber, 2001). While these kinds of innovations are important for the growth and development of the US agricultural sector, dependence of rural economies on agriculture significantly declined after the industrial revolution, and this gave rise to a new diversity of rural industries. Thus, when policy makers overlook rural industrial diversity, this oversight likely negatively impacts non-agriculture related innovation in rural areas through missed opportunities for new firms and reduced competition for existing firms (Stauber, 2001).

To provide guidance for policy makers that helps mitigate the negative effects of the challenges highlighted above, it remains necessary to continue expanding my understanding of the obstacles faced by rural firms in terms of innovation creation and adoption (Chatterji, et al., 2014; Fortunato, 2014). This is the underlying motivation for this study. I develop my analytical models from firm-level data provided by the 2014 National Survey on Business Competitiveness (NSBC). The NSBC data were made available by United States Department of Agriculture (USDA) for confidential access. It is a unique survey of the US firms, containing 257 variables from questions covering topics such as Research and Development (R&D) activities, innovation outputs, failed innovations, patents, other intellectual property protection, employee education levels, affiliated industry, location factors including local amenities, market share, location-based barriers, local government impact, among others. I combine selected innovation related firm variables from the NSBC data with county-level secondary data from the Bureau of Economic Analysis (BEA), Bureau of Labor Statistics (BLS), Small Business Administration (SBA), and US Census Bureau, to capture the external business and innovation environment in which these firms operate.

I use the number of patent applications as my measure for innovation creation and employ negative binomial regression models to empirically test firm-level characteristics that influence innovation creation. Three models are estimated, one for combined sample of rural and urban firms (N = 4,351), one for urban firms only (N = 1,117), and one for rural firms only (N = 3,234). In my empirical analysis, I first test whether there is a difference between the innovation-related characteristics of rural and urban firms once external regional factors are controlled. Further, I examine potential firm-level characteristics and behaviors that drive innovation across urban and rural firms. I find that there are differences between rural and urban firms in terms of influencers of their patenting activities, with urban firms exhibiting better capability to capitalize on their resources compared to rural firms. Additionally, I find evidence that the influence of university R&D is relevant to innovation creation in rural firms, but their perception of university provided information may not be significant.

The remainder of this paper is organized as follows. The next section provides a review of literature on firm innovation and its measure. The third section describes my research methodology including a description of the data, selection of variables, and the empirical model. Results are presented and discussed in section four, and section five concludes with summary and policy recommendations.

3.2. Literature Review

During the 1950s through 1970s, innovation studies as well as policies directed towards improving innovation mainly focused on the role large firms play in driving the innovation process (Chandler, 1977; Schumpeter, 1942). The belief was that large firms, through scale economies, were best suited to bear the risk of R&D investment necessary to create new innovations. The role of small businesses in this process was viewed as minimal as they were argued to be handicapped by a lack of financial, physical, and human capital needed to innovate and commercialize new technologies (Galbraith, 1956). With the emergence of new technologies and as new evidence showing that small firms played a major role in job creation, this belief evolved. Scholars recognized that scale economies also occurred through geographic proximity to a large number of small firms and this is as important to the innovation process as the scale economies of large enterprises. This new understanding recognized the importance of the links between entrepreneurship, small and large firms, and innovation creation in terms of driving technological progress (Acs & Audretsch, 2005). From this view and based on the knowledge production function framework formalized by Griliches (1979), business formation is a key starting point in the innovation process. While new firms are created exogenously, the innovation and technological change occurs through the performance of these firms endogenously as they pursue knowledge creation for the purpose of improving the firm's standing (Arrow, 1962; Cohen & Levin, 1989; Scherer, 1984). Therefore, R&D efforts are considered as the most important inputs to innovation creation, but these efforts remain relevant to both new and established firms regarding innovation creation (Cohen & Klepper, 1991; 1992).

The focus on the firm as a central unit of the innovation process shifted to a broader geographical level of analysis, following Jaffe's (1989) modification of Griliches' (1979) knowledge production function to study the spillover of knowledge between universities and private industries (Audretsch & Feldman, 2004). Over the years, a wide theoretical consensus has emerged showing that knowledge spillover is an important source of innovation in urban areas (Audretsch & Feldman, 1996; Rosenthal & Strange, 2004). This led to wider scholarly interest in urban innovation studies as densely co-located firms tend to facilitate face-to-face interaction among knowledge workers and give rise to greater extent of spillover tacit knowledge spillover in urban areas (Glaeser et al., 1992; Henderson, 2003).

On the empirical side, difficulty in measuring innovation and technological progress, generally arising from data availability, made estimation of the knowledge production function challenging (Cohen & Levin, 1989; Kuznets, 1962). The available measures act as proxies reflecting one or more aspect of the innovation process. Firm- and regional-level innovation metrics are typically categorized as: (1) the inputs into the innovative process, such as R&D expenditures and the share of R&D employees in workforce; (2) an intermediate output, such as the number of patent applications; or (3) direct measure of innovative output such as new product announcements (Aghion & Howitt, 1990).

Each category, and respective measure, has limitations, and this fact is well reflected in the literature. For example, tangible innovation creation appears lumpy relative to the levels of inputs such as R&D expenditures. (Kleinknecht,1987; Kleinknecht & Verspagen,1989). Additionally, formal R&D budgets are not necessarily solely directed toward innovation creation; instead, they may include activities such as imitation and technology transfer (Mansfield, 1984). Similarly, patent applications and awards data, often used as a measure of innovation, are frequently criticized in the literature. For example, using number of patents as a measure of innovation suffers from the implicit assumption of homogeneity regarding the innovation's economic value both in terms of market value and total R&D investment (Cohen & Levin, 1989; Pakes & Griliches, 1980). Further, not all innovations are likely to result in patents nor are all patents likely to be used for a final innovative output, for example, they may be used as leverage for financing or held as defense against competing products (Nagaoka et al., 2010). Challenges also arise in the use of direct measures of innovative output such as new product or service launches in a market. Most notably, new product or service launches and similar output measures are expensive and labor intensive to measure (Acs, et al., 2002; Huang et al., 2010)

While patents as innovation measures have limitations, the literature maintains that patents remain a reliable metric for innovation creation (Acs, et al., 2002; Czarnitzki et al., 2009; Pakes & Griliches 1980). For example, Acs et al., (2002) compared patent applications to an SBA data set constructed from information in trade and technical journals on new products and reported that patents performed as well as this alternative innovation creation measure. Similarly, comparing 40 different potential innovation measures constructed from 2014 NSBC data (the same data set as used in the current study), Parker et. al., (2017) showed that patent applications were significantly correlated with the other 39 innovation measures. Additionally, current availability and the historical use of patent data makes it a popular measure in terms of examining changes over time and for comparing different levels of aggregation (e.g., influencers at the firm-level versus the regional-level).

3.3. Data

Two types of variables are included in my model - firm and county level. Firm-level data are from the 2014 NSBC conducted by the USDA. Respondents (N=10,929) are comprised of US establishments with more than five employees in the tradable sectors that include mining, manufacturing, wholesale trade, transportation and warehousing, information, finance and insurance, professional/scientific/technical services, arts, and management of business. In total, there are 257 potential variables from survey questions covering topics such as R&D activities, innovation output (sales from new or improved products or services), failed innovations, patents, other intellectual property protection, employee education levels, affiliated industry, business founder's conceptualization, location factors including local amenities, effects of the 2008-2009 recession, market share, location-based barriers, and local government impact[22]. However, only a portion of the total 2014 NSBC observations were included in this study due to incomplete responses (1927), observations with missing location of the firm in terms of county FIPS (350), observations for which the respondents reported either they were "not familiar" or "slightly familiar" with how innovation carried out in their businesses (704), and observations with missing responses relevant variables in my analysis (3404).

[22] For details on the survey, please refer to Wojan (2015) available at https://www.oecd.org/sti/193%20-%20SelfReportedInnovationSurveys_IncreasingReliability_ClearedManuscript.pdf

County-level data are intended to represent the regional business climate and include variables for university R&D, human capital, industrial structure, and selected demographic, and fiscal characteristics. The county-level data cover the 48 contiguous states not including the District of Columbia[23]. These data come from the US Census Community Business Patterns (CBP) and American Community Survey (ACS), the US Bureau of Economic Analysis (BEA), and National Science Foundation (NSF). When matching Firm- and county-level data by the county FIPS codes, 193 additional observations were dropped due to missing values within the county-level dataset. Thus, my effective sample for this study includes 4,351 establishments. Table 1 below includes variable names, a brief description, and source, and the next few subsections provide details about variables selection and construction.

Dependent Variable: Firm Innovation Creation. The 2014 NSBC included three questions about patenting that occurred between 2011 and 2013, including whether or not the firm applied for one or more patents (binary), the number of patent applications filed (count), and the number of patents awarded (count). Part of my motivation for use of the self-reported patent counts is to make these study results comparable to prior work, and patent counts allow for more modeling flexibility relative to a binary response as it includes magnitude (Acs et al., 2002; Czarnitzki et al., 2009; Griliches, 1990; Trajtenberg, 1987). Of the two patent count options from the survey (applications and awards), patent applications are frequently used in the literature as they reflect the most recent level of firm inputs.[24] Further, the patent award date relative to when the application is filed can occur in the same year or even decades from the application date (Hall et

[23] Alaska and Hawaii were excluded because of missing observations for several counties; District of Columbia was also excluded as it has a single county and I control for the state-level fixed effects using state dummies in my analysis. From the 48 included states, I also eliminated the counties with missing values for county-level variables

[24] I considered scaling patents by the number of persons in the inventive class (engineers and other scientists), but the “number of professionals” which reflects to some degree this value also includes a wide range of other nonscientific fields such as accountants. Thus, self-reported firm-level patent application counts were used.

Table 3.1 Variables Description and Data Source

Variables	Definition	Source (Year)
Firm-level		
Patent applications	Total number of patent applications during 2011-13	NSBC (2014)
Rural	Located in a non-metro county (1=yes; 0=no)	
Academic information	Academia as valuable source of new ideas (not at all valuable=0, somewhat valuable=1, very valuable=2)	
Bachelor's degree	Employs individuals with at least bachelor education (1=yes; 0=no)	
Difficulty hiring	Difficulty finding qualified applicants (0=very difficult; 1=somewhat or not difficult)	
High-tech (NSF def.)	Firm belonging to high-tech industry (1=yes; 0=no)	
Firm size	Establishment size (total number of employees)	
Firm age	Establishment age (years in operation until 2013)	
Percent man. and prof.	Management and professional employees as percent of full and part time employees on payroll (percentage points)	
Final innovative output	Introduced innovation in product, service, production, or distribution method in past 3 years (1=yes; 0=no)	
Other IP activity	Involved in other forms of IP protection than patents in past 3 years (1=yes; 0=no)	
Abandoned innovation	Any improvement or innovation activities abandoned in past 3 years (1=yes; 0=no)	
R&D activity	Conducted internally or hired, R&D and design services in past 3 years (1=yes; 0=no)	
Angel/venture funding	Received some venture or angel capital financing in past 3 years (1=yes; 0=no)	
Rejected for loan	Tried to borrow but received none from financial institutions in past 3 years (1=yes; 0=no)	
Green tech	Production or service provision to any green energy sector (1=yes; 0=no)	
Internet sales	Sold products or services over the internet (1=yes; 0=no)	
Export products	Exported products/services internationally (1=yes; 0=no)	
Industry-level Fixed Effects	Industry indicators at two-digit level NAICS (NAICS 21, 31, 32, 33, 42, 48, 51, 52, 54, 55, and 71)	
County-level		
Univ. R&D per cap.	University R&D per capita	NSF (2010)
SPLAG univ. R&D per cap.	University R&D per capita in neighboring counties	NSF (2010)
Percent pop. bach. degree	Bachelor or higher degree holders as percent of population 25 years and over	ACS (2010)
High-tech variety	High-tech Variety	CBP (2010)
Percent foreign born	Foreign-born population as percent of total county population	ACS (2010)
Percent prof., sc., and tech. employment	Employment in professional, scientific, and technological industries as percent of civilian employed population 16 years and over	ACS (2010)
Unemp. rate	Unemployment rate	BEA (2010)
Total tax per capita	Total taxes per capita	Census of Govts. (2012)

al., 2005). This makes patent applications more consistent with other variables generated from the 2014 NSBC as they reflect input levels in the nearby period as when new applications were filed but not necessarily for those applications leading to awards if the application-award lags were many years down later. Further, while patent applications reflect the firm's effectiveness in endogenous innovation efforts, patent awards depend on whether other firms/individuals were first movers with a similar patent application. Therefore, I selected patent applications as a better indicator of innovation output and use it as the dependent variable in my estimations.

Independent Variables. I include a range of firm specific characteristics in my model including: location (rural/urban county), innovation creation actions and behaviors, and perceptions and characteristics related to human capital. First, firms located in rural counties are distinguished from those in urban counties. County classification is based on the 2013 Rural Urban Continuum Codes (RUCC). In the combined model (discussed more in the methods section), an indicator for rural is included. The other two models include urban-only or rural-only firms. It is important to note that a goal of the 2014 NSBC was to collect data allowing for detailed analysis of rural firms while also making comparison of results to urban firms possible. To achieve this, the survey over-sampled rural firms relative to urban firms. Thus, the rural-only model includes about 3 times the number of firms as the urban-only model does, and the combined model is heavily weighted towards rural firms. Similar to prior literature, I anticipate that the rural parameter in the combined model is negative, and/or that differences in urban and rural firm models necessitate separate models, that is, one for rural and one for urban firms.

Second, I consider specific firm behaviors and activities in the innovation creation process. Aghion & Howitt (1990) identified three categories or stages of innovation development, Research and Development input (R&D expenditure), intermediate R&D output (patent), and final innovative output (new product or process), and the most innovative firms were active in each. I broaden their description of each category to include other ways in which these activities may occur and based on 2014 NSBC responses. Within my modeling framework, each category is represented with an indicator variable. The first category (R&D input) includes in-house R&D, purchased external R&D, design activities, and design services. The second (intermediate R&D output) is made up of forms of IP protection other than patents (the dependent variable) and includes, industrial design, trademark, copyrights, trade secrets, and first mover's advantage. The third (final innovative output) was expanded to include producing any new or significantly improved goods or services, introduction of new or significantly improved methods of manufacturing, and use of new logistics, delivery and distribution methods for inputs, goods or services. Additionally, firms may choose to abandon an innovation at some stage of development, and I include an indicator for this decision. Lin et al., (2013) showed that innovative firms with mixed and complimentary IP strategy (form example, using multiple forms of IP protection) tent to be more successful. Additionally, and keeping within the tradition of the framework described by Aghion & Howitt (1990), I identified firms as "high-tech" if it operated in an NSF-designated high-tech industry based on the 4- and 6-digit NAICS codes of firms provided by the 2014 NSBC (NSF, 2016). I expect that all these parameters to be positively associated with patenting activity.

Third, I include a number of indicator variables based on activities that may influence innovation creation. Many businesses collaborate with academic institutions in conducting research activities. However, Howells, et al. (2012) showed that while these collaborations benefited the firms, the firms did not necessarily acknowledge this benefit. It may be that the firm-level variables for academic obtaining academic information are negative and the county-level controls for university R&D (discussed below) are positive, supporting Howells et al., (2012) finding. Similarly, the research findings and the extension outreach programs of universities can benefit firms by introducing them to new knowledge (Lyons et al., 2018). These results may be similar or different from what Howells et al., (2012) found.

Firms may also get access to angel or venture funding to help further develop and scale up an innovation, or they may be limited to pursuing more traditional forms of financing such as loans from financial institutions (Renski & Wallace 2012). I expect the former to be positively associated with patenting, and the latter, which is framed as rejection for private financing (rejected for loan), to be negatively associated with patenting. I also include indicators for firms that said they sold their products or services via internet, exported their products or services, and produced products or provided services in any of the five "green" sectors (production of renewal energy, increasing energy efficiency, conservation of natural resources, prevention, reduction, and cleaning up of pollution, and production of clean transportation fuels). I anticipate these indicators for broader market access and new markets (green tech) are positively associated with patenting.

Fourth, the NSBC survey provides information about different aspects of human capital choices and perceptions. I include an indicator for firms that required individuals with at least bachelor's degree for any of their occupational categories, and an indicator for firms that reported having difficulty in finding qualified applicants for their positions in the labor market. Following Aghion & Howitt (1990), I anticipate the first (bachelor's degree) to be positively correlated with patenting, while the second (difficulty hiring) to be negatively associated with patenting. I include the share of management and professionals to total employees at the firms, a measure of establishment size (total number of employees), and the age of the firm. Based on the finding in the literature (Aghion & Howitt, 1990; Henderson, 2003) I expect that these final variables are positively associated with patenting.

Industry controls. Firms in different industries likely vary in terms of their patenting propensity and intensity (Wojan et al., 2015). I control for this heterogeneity across firms by including two-digit NAICS industries associated with the respondent firms in my sample. The industries included in the 2014 NSBC are: mining, quarrying, and oil and gas extraction (NAICS 21); food, beverage, textile, and animal products manufacturing (31); wood products, paper, chemical, petroleum, plastics and rubber, and nonmetallic mineral products manufacturing (32); metal, machinery, computer and electronic products, transportation equipment, furniture and related products, and miscellaneous manufacturing (33); wholesale trade (42); transportation (48), information (51), finance and insurance (52); professional, scientific, and technical services (54), management of companies and enterprises (55); and arts, entertainment, and recreation (71).

County-level controls. To control for regional heterogeneity and the business environment in which the firms operate, I include university R&D per capita in own county of firm location, university R&D in neighboring counties located within 100-mile radius (variable constructed as a spatial lag of university R&D), percentage of population with bachelor or higher degree of education, number of high-tech establishments as a percentage of total establishments, variety of high-technology industries, foreign-born population as a percentage of total population, share of employment in professional, scientific, and technical services sector to total civilian employment, unemployment rate, and total taxes per capita. With the exception of the last two terms (unemployment and taxes which I anticipate to be negatively correlated with patenting), I expect these parameter to be positively associated with patenting.

Finally, I include state-level fixed effects to control for the heterogeneity among states. I use California as the reference state as it is well known for innovation centers such as Silicon Valley (Mann & Shideler, 2015). Since I construct separate models for rural and urban firms, I examine the state fixed-effects in terms of which states may provide a relative advantage or disadvantage to firms compared to California. The state fixed-effects are discussed more at the end of the results section 5.3.

3.4. Methods

I operationalize firm innovation by using the number of patent applications that firms reported filing between 2011 and 2013 as the dependent variable and are guided by the traditional literature on modeling patents counts (e.g., see Allison & Waterman, 2002; Hall, et al., 1986). As the number of patent applications is a count variable taking on only non-negative integer values, analyses using linear regression models are not appropriate. The violation of the assumptions of linear regression regarding homoscedasticity and normal distribution of residuals, which is atypical to count dependent variable like ours, is likely to lead to biased and inconsistent coefficient estimates (Greene, 2003). The count models such as Poisson and negative binomial are more appropriate for analyzing count data such as the number of patent applications filed in a given year (Allison & Waterman, 2002; Greene, 2003; Hall, et al., 1986).

Figure 1 shows that the distribution of patent applications data in my combined sample[25] of rural and urban firms is clearly right-skewed. Thus, I turn to Poisson and negative binomial process distribution in terms of constructing my regression models. However, based on my preliminary modeling evidence, specifically the likelihood ratio tests between the initial Poisson

[25] The frequency distribution of total number of patent applications is similar for urban and rural sub-samples (not reported)

and negative binomial models (discussed more in the results section), indicate that patent applications data in my sample are over-dispersed. In the presence of such over-dispersion of the

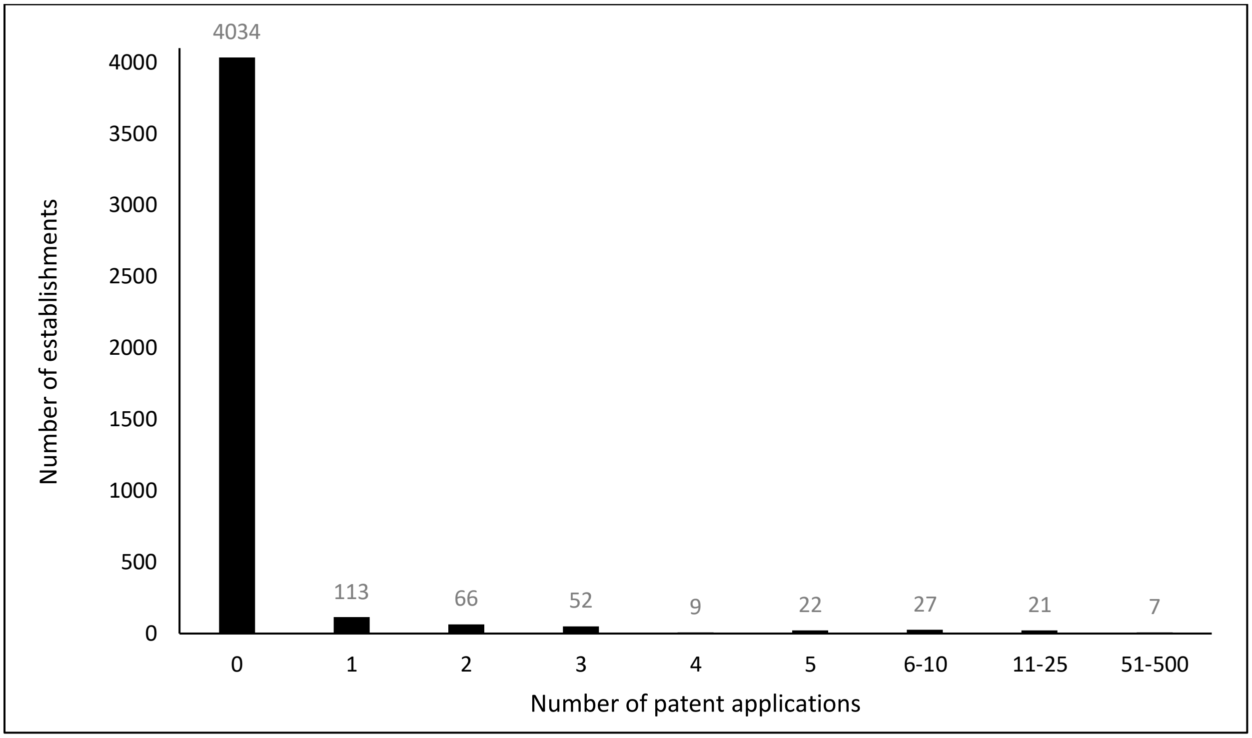

Figure 3.1 Frequency distribution of firm-level total patent applications during 2011-13 (pooled sample)

dependent variable, the Poisson regression model is inappropriate as the over-dispersion likely causes spurious significance of the coefficient estimates due to underestimation of standard errors (Cameron & Trivedi, 1986). On the other hand, the negative binomial models allow over-dispersion (variance > mean) through its separate parameterization. Therefore, I settle on the negative binomial model which has the following form expressed in terms of its log-likelihood function (Hilbe, 2011):

$$\mathcal{L} = \sum_{i=1}^{n} \left\{ y_i ln\left(\frac{\alpha\exp(x_i'\beta)}{1+\alpha\exp(x_i'\beta)}\right) - \frac{1}{\alpha}\ln(1+\alpha\exp(x_i'\beta)) + ln\Gamma\left(y_i + \frac{1}{\alpha}\right) - ln\Gamma(y_i+1) - ln\Gamma\left(\frac{1}{\alpha}\right) \right\}$$

where y_i represents the outcome variable for firm i, measured with reported patent applications it filed during 2011-13; x_i represents the vector of explanatory variables, including firm-level variables, industry controls, county-level controls, and state indicators; and α and β represent overdispersion parameter and the vector of other model parameters to be estimated respectively.

3.5. Results

Results and discussion are presented as follows. First, I include a brief discussion of the summary statistics of the model data. Second, I discuss the regression diagnostics and model selection. Third, I present the significant finding and their potential implications.

3.5.1. Summary Statistics

I report the descriptive statistics in table 2 separately for the sample of firms located in urban and rural counties and those for the combined sample. The Spearman's rank correlation coefficients for the combined sample are presented in table 3. The combined sample of 4,351 firms are across 1,562 US counties, and 25% (1,117 firms) are located in 422 urban counties and remaining 75% in 1,140 rural counties.

Firms located in urban areas had higher values for the average number of patent applications (1.24) compared to those in rural area (0.27). Overall however, 93% (4,034 out of 4,351 firms, see figure 1) of the firms reported zero patent applications during the period 2011-13, and the average number of patent applications for my combined sample is 0.52. Firm age is negatively correlated with patenting, and urban areas, on average, host younger firms compared to rural areas. All other variables that are positively correlated with patenting, except direct innovation and green tech, have higher average values or frequencies for urban regions compared to rural regions. The observations from the descriptive statistics indicate that rural firms innovate less frequently than urban firms. I also discuss selected variables' summary statistics in the context of the parameter results in section 3.5.3.

3.5.2. Regression Model Diagnostics and Interpretation of Results

As discussed in the methods section, I first estimated Poisson models separately for rural and urban firms and the combined sample. Most coefficient estimates were statistically significant at the 5% and 1% levels (Poisson results not shown). I then estimated negative binomial models, which allowed incorporation of the over-dispersion of the patents data. The results reported in table 4 for *alpha* (the over-dispersion parameters) provides a test of appropriateness of the Poisson models. The statistically significant *alpha*s in all three columns of the coefficient estimates demonstrated that the null hypothesis of zero dispersion is rejected at 1% significance level, thus suggesting the statistically significant coefficients in the Poisson regression models were likely due to underestimated standard errors arising from the over-dispersed patent data.

Additionally, I estimated four specifications of the negative binomial regression model: (i) no county-level controls or state-level fixed effects, (ii) county-level controls only, (iii) state-level fixed effects only, and (iv) both county-level controls and state-level fixed effects. While I do not report the results from the first three specifications, (i)-(iii), the model chosen is based on the AIC and BIC selection criteria which identified the county-level controls and state-level fixed effects as the better specification of the four.

Table 3.2 Summary Statistics

Variables *	Combined Mean (SD)	Urban Mean (SD)	Rural Mean (SD)	Variables	Combined Mean (SD)	Urban Mean (SD)	Rural Mean (SD)
Firm-level Variables				*Industrial (2-digit NAICS)*			
Patent apps. (counts)	0.52	1.24	0.27	21	2%	1%	2%
	(8.41)	(16.12)	(2.28)	31	5%	3%	6%
Firm size (# employees)	55.04	63.57	52.09	32	9%	6%	9%
	(275.55)	(398.94)	(217.23)	33	18%	17%	19%
Firm age (# years)	32.86	26.46	35.07	42	17%	21%	16%
	(28.09)	(23.92)	(29.07)	48	6%	3%	7%
Percent man. and prof. (% points)	23.57	28.55	21.85	51	8%	5%	9%
	(21.00)	(25.22)	(19.03)	52	4%	2%	4%
Academic information				54	25%	34%	22%
not valuable	13%	16%	12%	55	3%	5%	3%
somewhat valuable	52%	48%	53%	71	3%	3%	3%
very valuable	35%	36%	35%	*County-level variables*			
Bachelor's degree (1=yes; 0=no)	56%	67%	52%	Univ. R&D peer cap. ($)	94.05	187.07	59.62
Difficulty hiring (1=yes; 0=no)	26%	22%	27%		(776.48)	(512.11)	(851.46)
High-tech (NSF def.)	20%	31%	17%	SPLAG univ. R&D peer cap. ($)	336.3	283.64	355.8
Final innovative output (1=yes; 0=no)	71%	70%	72%		(994.36)	(740.47)	(1072.85)
Other IP activities (1=yes; 0=no)	33%	46%	29%	Percent bach. degree pop. (% points)	9.15	11.89	8.13
Abandoned innovation (1=yes; 0=no)	26%	30%	25%		(3.63)	(3.72)	(3.02)
R&D activity (1=yes; 0=no)	60%	66%	58%	High-tech variety	15.99	28.57	11.33
Angel/venture funding (1=yes; 0=no)	2%	2%	1%		(9.94)	(8.58)	(5.28)
Rejected loan (1=yes; 0=no)	5%	5%	5%	Percent foreign born pop (% points)	4.71	8.57	3.28
Green tech (1=yes; 0=no)	33%	31%	34%		(5.45)	(7.32)	(3.66)
Internet sales (1=yes; 0=no)	48%	50%	48%	Unemployment rate (% points)	7.28	7.15	7.32
Export products (1=yes; 0=no)	27%	35%	25%		(2.46)	(2.04)	(2.59)
				Tax per capita ($)	1441.25	1686.57	1350.44
Number of Observations (N)†	4,351	1,117	3,234		(955.72)	(91354)	(737.17)

* Variables defined in table 1; † Number of counties in combined, metro, and non-metro samples are 1562, 422, and 1140 respectively

Table 3.3 Spearman's Rank Correlation Coefficients

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Patent applications																	
2. Academic information	0.01																
3. Bachelor's degree	0.08*	-0.08*															
4. Difficulty hiring	-0.02	0.04*	-0.09*														
5. High-tech industry	0.1*	-0.07*	0.11*	-0.03*													
6. Firm size	0.2*	0.03	-0.03	-0.01	0.06*												
7. Firm age	-0.05*	0.02	-0.02	-0.01	-0.12*	0.15*											
8. Percent man. & profs.	0.1	-0.11*	0.18*	-0.09*	0.18*	-0.27*	-0.04*										
9. Final innovative output	0.14*	0.9*	-0.08*	0.04*	0.001	0.18*	-0.03	-0.11*									
10. Other IP activity	0.33*	-0.02	0.09*	-0.02	0.001	0.21*	-0.11*	0.04*	0.2*								
11. Abandoned innovation	0.13*	0.001	0.03*	0.06*	0.06*	0.11*	-0.05*	-0.02	0.12*	0.2*							
12. R&D activity	0.21*	0.01	0.07*	0.01	0.13*	0.21*	-0.06*	0.01	0.3*	0.35*	0.17*						
13. Angel/venture funding	0.12*	-0.01	0.02*	-0.01	0.03	0.09*	-0.06*	0.01	0.02	0.08*	0.06*	0.09*					
14. Rejected for loan	0.02	0.01	-0.03*	0.05*	-0.01	-0.04*	-0.1*	-0.02	0.001	0.04*	0.05*	0.02	0.08*				
15. Green tech	0.11*	0.08*	0.01*	0.04*	0.11*	0.13*	0.02	-0.07*	0.13*	0.09*	0.06*	0.17*	0.02	0.001			
16. Internet sales	0.06*	0.01*	-0.03*	0.01	-0.04*	0.07*	0.01	-0.03	0.11*	0.1*	0.08*	0.11*	0.02	0.02	0.02		
17. Export products	0.30*	0.01	0.01	-0.01	0.06*	0.24*	-0.05*	-0.04*	0.22*	0.28*	0.09*	0.28*	0.05*	-0.02	0.1*	0.14*	
18. Rural	-0.06*	0.01	-0.13*	0.06*	-0.15*	-0.02	0.14**	-0.1*	0.01	-0.16*	-0.05**	-0.07*	-0.02	-0.01	0.03	-0.01	-0.1*

* significant at 5% or lower levels

Regression results are presented in table 4 (below) for the combined, urban, and rural models. The beta coefficients and their statistical significance are shown in the first column under each group category and the incidence rate ratio minus 1 (IRR-1) are shown in the second[26]. I report the IRR-1 as this is a straightforward method to interpret results (for recent examples in the literature using this method see Howell, 2017; Murray & Stern, 2007; Rothaermel & Hess, 2007). For example, under the combined model, the *percentage of management and professional employees* (*percent man. and prof.*) parameter is interpreted as a one-unit increase (in this case a 1% increase) in the share of management and professional employees is expected to increase patenting activity by 1%.

3.5.3. Rural and Urban Firm Innovation

To address my broader research question, I used a likelihood ratio test to empirically check for a difference in the innovation creation (patenting) behavior of rural and urban firms that were otherwise identical in their characteristics. For this test, the null hypothesis is framed as rural and urban firms represented by common innovation behavior parameters. Thus, I consider if the use of the combined model for analysis is appropriate, or are two models more appropriate, one for rural and one for urban. The test statistics for the first test is derived as two times the difference between the log-likelihood of model 1 and the sum of the log-likelihoods of model 2 and 3 (-1410.0 -487.1 + -858.1 = -64.8, 2*|-64.8| = 129.6)[27]. The critical chi-square value with 76 degrees of freedom and at the 99% confidence level (107.58) is less than the test statistic. Thus, I reject the null hypotheses in favor of using individual urban and rural models in place of the

[26] These are constructed as IRR-1 = exp(β)-1 (Hilbe, 2011).
[27] See Brooks and Lusk (2010) for the inferential approach using likelihood ratio test.

combined model. In other words, my test reveals that there are some differences between ruraland urban firms in terms of potential influencers of patenting activities.

Of the statistically significant firm-level parameters, participation in other forms of IP protection has the largest magnitude in difference between rural and urban firms; however, both are positive and support the findings from Lin et al., (2013). The IRR-1 reveals that for urban firms, use of other forms of IP protection is correlated with a three-fold increase in patent activity compared to rural firms. In other words, other forms of IP protection appear much more important for urban firms in terms of their innovation creation. It may be that closer proximity to or higher density of other innovative firms contributes to this result. It could also be that since innovative urban firms compete in broader markets more frequently compared to innovative rural firms (revealed by the internet sales and export products summary statistics and parameters), increased participation in other forms IP protection is necessary. These results may also be influenced by some of the other differences in characteristics in which urban firms have an advantage, for example, access to angel and venture funding, participation in green tech (urban firms show a greater magnitude impact on patenting activity although a higher portion of rural firms participate in this), and using R&D activity. It may be that private equity investors insist on more protections for the innovations developed and greater R&D investment motivate broader IP protection.

Another interpretation of these results is that there is a connectedness between different innovation related activities, and that urban firms are able to better capitalize from the combined effort of these activities relative to rural firms. Thus, my results support prior studies that demonstrate that urban firms have higher levels of innovative activity compared to rural firms.

Table 3.4 Negative Binomial Regression Results

Variables (DV: Number pat. Apps)	Combined		Metro		Non-metro	
Firm-level	Beta	IRR-1	Beta	IRR-1	Beta	IRR-1
Rural	0.077	0.08				
Academic information						
somewhat valuable	-0.612**	-0.46	-0.167	-0.15	-0.251	-0.22
very valuable	-0.699**	-0.50	-0.588	-0.44	-0.201	-0.18
Bachelors Degree	0.465**	0.59	0.219	0.24	0.550***	0.73
Difficulty hiring	-0.233	-0.21	-0.285	-0.25	-0.140	-0.13
High-tech (NSF Def.)	0.314	0.37	0.830**	1.29	0.042	0.04
Firm size	0.655***	0.93	0.823***	1.29	0.576***	0.78
Firm age	-0.025	-0.02	-0.200	1.28	-0.036	-0.04
Percent man. & prof.	0.013***	0.01	0.014*	0.01	0.011*	0.01
Final innovative output	0.255	0.29	0.462	0.59	0.465	0.59
Other IP activity	2.200***	8.03	3.109***	21.40	2.092***	7.10
Abandoned innovation	0.452***	0.57	0.066	0.07	0.473***	0.60
R&D activity	1.299***	2.67	1.593**	3.92	1.460***	3.31
Angel/venture funding	0.983**	1.67	1.736*	4.67	0.319	0.38
Rejected for loan	-0.117	-0.11	-0.758	-0.53	0.042	0.04
Green tech	0.195	0.22	0.696*	1.01	0.308*	0.36
Internet sales	0.393**	0.48	0.965***	1.62	0.302	0.35
Export products	1.152***	2.16	1.354***	2.87	1.056***	1.87
County-level Variables						
Univ R&D per cap.	0.083**	0.09	0.049	0.05	0.121**	o.13
SPLAG Univ. R&D per cap.	0.046	0.05	0.126	0.13	0.000	0.00
Per. pop. bach. Degree	-0.005	0.00	0.039	0.04	-0.048	-0.05
High-tech variety	0.016	0.02	0.009	0.01	0.025	0.03
Per. foreign born	0.021	0.02	-0.018	-0.02	0.018	0.02
Unemp. Rate	0.035	0.04	-0.142	-0.13	0.082	0.09
Tax per capita	0.810**	1.25	0.895	1.45	0.596	0.81
Constant	-18.112***		-17.309**		-17.860***	-0.05
ln(alpha)	1.689***		1.745***		1.221***	0.03
Industry-level Fixed Effects	Yes		Yes		Yes	
State-level Fixed Effects	Yes		Yes		Yes	
Number of obs.	4,351		1,117		3,234	
Log-likelihood	-1410		-487.1		-858.1	
Model DF	81		79		76	
AIC	2985.78		1136.17		1872.22	
BIC	3515.17		1542.66		2346.58	

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

The results in this study also provide additional insight on this phenomenon, showing more detail on some of the nuances of how and/or why this may be occurring. One surprising result is that the urban-rural innovation gap is not seen in some of the county-level business environment control variables, such as high-tech variety or educated labor force in the urban areas. For example, using county-level regional model, Aryal et al., (2018) show these factors are positive and statistically significant for urban and rural firms but to different degrees.

There are two firm-level results that also provide additional insight. First, the bachelor's degree parameter (one or more jobs at the firms requiring a bachelor's or higher degree) is relevant (and statistically significant) to rural firms in the context of my modeling of patenting activity, but not to urban firms. In terms of innovative activity, the difference in the human capital needs between rural and urban firms may show that there is greater variability among rural firms. In short, innovative rural firms need educated employees but other rural firms rely less on an educated workforce; whereas, for urban firms both patenting and non-patenting firms depend on an educated labor force. The summary statistics showed that 67% of urban firms had at least one (or more) positions that required at least a bachelor's degree (it was 52% for rural firms). Thus, for urban firms, the employee standard appears higher which may imply that the distinction between different urban firms is more about bachelor's degree field and less about whether or not the employee has the degree (the opposite appears relevant for rural firms).

Second, the "abandoned innovation" parameter is also positive and statistically significant for rural firms but not for urban firms. This implies that rural firms that expend effort, even if not successful, are more likely to create new innovations. From the summary statistics, only 25% of rural firms reported that they had innovation activities that were abandoned (compared to 30% of urban firms). This result may also reveal something about the level of risk rural firms are willing or able to manage. In short, rural firms that are more willing to take on risk—as evident from starting and abandoning innovations—innovate more frequently.

For the county-level controls, only university R&D per capita is statistically significant for rural firms. This result is interesting considering that the coefficient estimate of the variable “academic information” is not statistically significant (and negative). Howells, et al. (2012) reported that firms place a low value on the impacts of university technology transfer and partnerships, yet firms were shown to greatly benefit from these relationships. Given the results in my study, this may be especially true for rural firms.

Finally, table 5 shows the statistically significant (10% or lower level) state-level fixed-effect parameter estimates for the rural firms model. While not the main focus this research, the results provide an interesting contrast to firms operating in urban areas. In both the urban and rural models, California is the reference state. Only Kentucky in the urban model was statistically different than California (and negative). However, 16 states in the rural model were different than California and all were positive. California and Massachusetts historically are the leaders in innovation creation (Mann & Shideler, 2015), but much of the literature focusing on firm- or regional-level innovation creation is relative to urban firms. The results of table 5 suggest that when it comes to rural firm innovation creation, many other states may be ahead of the traditional leaders. Interestingly, several of the tops states in this table are small in terms of population (Wyoming, Montana, and Vermont). But larger states, such as New York and Texas also appear in this list. It may be that the location of some of these firms in rural areas benefit more substantially from urban spillovers (in the case of rural areas adjacent to urban centers). On the other hand, some state’s policies may also be better geared to serve innovative rural firms relative to California.

Table 3.5 Rural Innovative firms - Statistically Significant State Fixed Effects (Ref. State=CA)

State	Coefficient Estimate	IRR-1
Wyoming	4.21	66.3
Nevada	3.64	36.9
Vermont	3.39	28.5
Montana	3.29	25.8
Alabama	3.18	23.1
Kansas	2.98	18.7
Missouri	2.82	15.8
New York	2.75	14.7
Texas	2.74	14.5
Iowa	2.73	14.3
Ohio	2.68	13.5
Colorado	2.67	13.4
Kentucky	2.57	12.0
Mississippi	2.52	11.4
Tennessee	2.45	10.6
Minnesota	2.41	10.2

3.6. Summary and Conclusion

Much of the innovation creation literature is focused on urban firms or areas, or relies heavily on data based on these (NSF, 2016). Less studied are rural firms and areas in this context. The goal of this paper is to empirically test if and how much rural and urban firms differ in terms of behaviors and characteristic that may influence innovation creation. To accomplish this goal, I use the 2014 NSBC and combine it with regional secondary data that reflects the business and innovative environments in which these firms operate. My overarching finding is that urban firms are able to better capitalize on firm characteristics and behaviors that may influence innovation creation relative to rural firms. This finding is revealed as most of the parameters that are statistically significant for urban firms are also statistically significant for rural firms, but the magnitudes are higher for urban firms. While my main finding supports prior studies that show rural firms lag behind urban firms, my study also provides a few other insights as to how and why this is occurring.

First, my results suggest that innovation creation within rural firms is influenced more by university R&D than for urban firms. At the same time, information from universities (for example from extension services) may not necessarily be perceived by firms as impactful with respect to innovation creation. This finding supports Howells et al.'s (2012) counterintuitive results—with specific applications to rural firms—that while firms may not perceive value from universities, they do benefit in economic terms from their interactions with universities. Second, rural firms that are willing to try, but fail, in terms of innovation creation have a slight advantage over other rural firms less willing to take on the risk. This result is shown by the "abandoned innovation" parameter (from the 2014 NSBC question asking firms if any innovation project had been abandoned during 2011-2013 period). The implication is that rural firms that are more risk averse may also be less likely to innovate. Third, workers with at least a bachelor's degree appear to be more important for rural firms regarding innovation creation than for urban firms. However, I do not suggest here that an educated labor force is not important for urban firms in this regard. My summary statistics show that 2 out of 3 urban firms require a 4-year degree for at least some positions compared to about half of rural firms. Instead, it is likely that for rural firms, having qualified workers capable of innovation creation is a higher barrier relative to urban firms. Fourth, there are several factors that suggest urban firms are more competitive than rural firms, for example, due to their proximity to other innovative firms or based on the degree/intensity of accessing broader markets (such as via exports and ecommerce). Thus, for urban firms mixed IP protection strategies appear much more important compared to rural firms. Combined, these findings suggest potential opportunities for policies directed toward rural firms that can: (1) help mitigate the risk in innovation creation; (2) proved university support in terms of R&D, for example, access to intermediate R&D outputs such as licensing technologies; (3) provide qualified labor/assistance in terms of innovation creation or development; and (4) help rural firms access broader markets. One example may be improving access to public or private equity for R&D, such as through the Small Business Innovation Research program, or access to other kinds of programs designed to fund early stage R&D. Such improved access could occur with the aid of university-based training or research partnerships, and may include improved access to university developed technologies.

Fifth, the states that typically lead innovation creation among urban firms and areas are not necessarily the same for rural areas. Although the evidence presented to support this notion is only suggestive (state-level fixed effects parameters), it provides a topic for further research. For example, Wyoming, Vermont, and Montana appear in the top four of these rural leader states and all three are ranked near the bottom with respect to population, and Wyoming and Montana are lowest in population density among the 48 contiguous states. Thus, it may be state-level policies that impact innovation creation in these states cater to rural firms. An analysis of these policies relative to those for the leading states for urban firm innovation creation could provide important insights for other states wishing to improve rural firm innovation creation.

REFERENCES

REFERENCES

Acs, Z. (2006). How is entrepreneurship good for economic growth? *Innovations*, *1*(1), 97-107.

Acs, Z. J. & Audretsch, D. B. (2005). Entrepreneurship, innovation, and technological change", *Foundations and Trends in Entrepreneurship*, *1*(4), 149-195.

Acs, Z. J., Anselin, L., & Varga, A. (2002). Patents and innovation counts as measures of regional production of new knowledge. *Research policy*, *31*(7), 1069-1085.

Aghion, P. & Howitt, P. (1990). *A model of growth through creative destruction* (No. w3223). National Bureau of Economic Research.

Allison, P. D. & Waterman, R. P. (2002), "Fixed-effects negative binomial regression models", *Sociological methodology*, *32*(1), 247-265.

Aryal, G., Mann, J., Loveridge, S., & Joshi, S. (2018). Drivers of Differences in Inventiveness Across Urban and Rural Areas. *In Review*.

Arrow, K. J. (1972). Economic welfare and the allocation of resources for invention. In *Readings in Industrial Economics* (pp. 219-236). Palgrave, London.

Audretsch, D.B. & Feldman, M.P. (1996). R&D spillovers and the geography of innovation and production. *The American Economic Review*, *8*(3), 630-640.

Audretsch, D.B. & Feldman, M.P. (2004). Knowledge spillovers and the geography of innovation. *Handbook of Regional and Urban Economics*, *4*, 2713-2739.

Brooks, K. &Lusk, J. L. (2010). Stated and revealed preferences for organic and cloned milk: combining choice experiment and scanner data. *American Journal of Agricultural Economics*, *92*(4), 1229-1241.

Cameron, A. C. & Trivedi, P. K. (1986). Econometric models based on count data. Comparisons and applications of some estimators and tests. *Journal of applied econometrics*, *1*(1), 29-53.

Carlino, G. A., Chatterjee, S., & Hunt, R. M. (2007). Urban density and the rate of invention. *Journal of Urban Economics*, *61*(3), 389-419.

Chandler, A. D. (1977)., *The visible hand: The managerial revolution in American business*. Belknap Press, Cambridge, MA.

Chatterji, A., Glaeser, E., & Kerr, W. (2014). Clusters of entrepreneurship and innovation. I*nnovation Policy and the Economy*, *14*(1), 129-166.

Christensen, C. M. (2013). *The innovator's dilemma: when new technologies cause great firms to fail*. Harvard Business Review Press.

Cohen, W. M. & Klepper, S. (1991). Firm size versus diversity in the achievement of technological advance. In Acs, J.D. and Audretsch, D.B. (Ed.), *Innovation and Technological Change: An International Comparison*, University of Michigan Press, Ann Arbor, pp.183-203.

Cohen, W. M. & Klepper, S. (1992). The tradeoff between firm size and diversity in the pursuit of technological progress. *Small Business Economics*, *4*(1), 1-14.

Cohen, W.M. & Levin, R.C. (1989). Empirical studies of innovation and market structure. *Handbook of Industrial Organization*, *2*, 1059-1107.

Czarnitzki, D., Kraft, K., & Thorwarth, S. (2009). The knowledge production of 'R'and 'D'. *Economics Letters*, *105*(1), 141-143.

Dabson, B. (2007). Rural-urban interdependence: Why metropolitan and rural America need each other. Available at http://www.rupri.org/Forms/Dabson_Brookings.pdf. Accessed 06/07/2017.

Dabson, B. (2011). Rural regional innovation: a response to metropolitan-framed place-based thinking in the United States. *Australasian Journal of Regional Studies*, *17*(1), 7.

Fortunato, M. W. P. (2014). Supporting rural entrepreneurship: a review of conceptual developments from research to practice. *Community development*, *45*(4), 387-408.

Galbraith, J. (1956). *American Capitalism: The Concept of Countervailing Power*. Routledge.

Glaeser, E.L., Kallal, H.D., Scheinkman, J.A., & Shleifer, A. (1992). Growth in Cities. *Journal of Political Economy*, *100*, 1126–1153

Greene, W. H. (2003). *Econometric analysis*. Prentice Hall, New Jersey.

Griliches, Z. (1979). Issues in assessing the contribution of R&D to productivity growth. *Bell Journal of Economics*, *10*(Spring), 92-116.

Grilliches, Z. (1990). Patent statistics as economic indicators: A survey. *Journal of Economic Literature*, *28*(4), 1661-1707.

Hall, B. H., Griliches, Z., & Hausman, J. A. (1986). Patents and R and D: Is there a lag? *International economic review*, 265-283.

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. *RAND Journal of economics*, 16-38.

Henderson, J. (2002). Building the rural economy with high-growth entrepreneurs. *Economic Review-Federal Reserve Bank of Kansas City*, *87*(3), 45.

Henderson, J.V. (2003). Marshall's scale economies. *Journal of Urban Economics, 53*(1), 1-28.

Hilbe, J. M. (2011). *Negative binomial regression*. Cambridge University Press.

Howells, J., Ramlogan, R., & Cheng, S. L. (2012). Innovation and university collaboration: paradox and complexity within the knowledge economy. *Cambridge Journal of Economics*, *36*(3), 703-721.

Huang, C., Arundel, A., & Hollanders, H. (2010). *How firms innovate: R&D, non-R&D, and technology adoption* (working paper no. 2027).

Kleinknecht, A. (1987). Measuring R & D in small firms: how much are we missing?. *The Journal of Industrial Economics*, 253-256.

Kleinknecht, A. & Verspagen, B. (1989). R&D and market structure: the impact of measurement and aggregation problems. *Small Business Economics, 1*(4), 297-301.

Knickel, K., G. Brunori, S. R.,m & Proost, J. (2009). Towards a better conceptual framework for innovation processes in agriculture and rural development: from linear models to systemic approaches. *The Journal of Agricultural Education and Extension*, 15(2),131-146.

Kuznets, S. (1962). Inventive Activity: Problems of Definition and Measurement. In Nelson, R.R. (Ed.), *The Rate and Direction of Inventive Activity*, National Bureau of Economic Research Conference Report, Princeton, NJ, 19-43.

Lin, E. S., Hsiao, Y. C., & Lin, H. L. (2013). Complementarities of R&D strategies on innovation performance: evidence from Taiwanese manufacturing firms. *Technological and Economic Development of Economy*, *19*(sup1), S134-S156.

Lyons, T., Miller, S., & Mann, J. (2018). A new role for land grant universities in the rural innovation ecosystem? *Journal of Regional Analysis and Policy*, forthcoming.

Mann, J. & Shideler, D. (2015), "Measuring Schumpeterian activity using a composite indicator", *Journal of Entrepreneurship and Public Policy, 4*, (1), 57-84.

Mansfield, E. (1984). Comment on using linked patent and R&D data to measure interindustry technology flows. In Griliches, Z. (Ed.), *R&D, patents, and productivity,* University of Chicago Press, Chicago, IL, 462-464.

Markley, D. (2001). Financing the new rural economy. *Exploring Policy Options for a New Rural America*, *69*, 80.

Mowery, D.C., Nelson, R. R., & Martin, B.R. (2010). Technology policy and global warming: why new policy models are needed (or why putting new wine in old bottles won't work). *Research Policy, 39*(8), 1011-1023.

Nagaoka, S., Motohashi, K., & Goto, A. (2010). Patent statistics as an innovation indicator. *Handbook of the Economics of Innovation*, *2*, 1083-1127.

Orlando, M. J., & Verba, M. (2005). Do only big cities innovate? technological maturity and the location of innovation. *Economic Review-Federal Reserve Bank of Kansas City*, *90*(2), 31.

Pakes, A. & Griliches, Z. (1980). Patents and R&D at the firm level: a first report. *Economics Letters*, 5, 377-381.

Parker, J., Mann, J., & Loveridge, S. (2017). Rural V. Urban: A National Survey on Determinants of Business Innovation Activities. Paper presented at the Midwest Economics Association (MEA), 31 March-2 April, Cincinnati, Ohio.

Porter, M. E., Ketels, C. H., Miller, K., & Bryden, R. (2004). Competitiveness in rural US regions: learning and research agenda. *US Economic Development Administration (EDA)*, Washington, DC.

Ratner, S. & Markley, D. (2014). Rural wealth creation as a sustainable economic development strategy: introduction to the special issue. *Community Development*, *45*(5), 435-442.

Renski, H. & Wallace, R. (2012). Entrepreneurship in rural America. *Financing Economic Development in the 21st Century, 245*.

Rosenthal, S.S. & Strange, W.C. (2004). Evidence on the nature and sources of agglomeration economies. *Handbook of Regional and Urban Economics*, *4*, 2119-2171.

Scherer, F.M. (1984). *Innovation and Growth: Schumpeterian Perspectives*. MIT Press, Cambridge, MA.

Schumpeter, J. A. (1942). *Capitalism, socialism, and democracy*. Harper and Row, New York.

Stauber, K. N. (2001). Why invest in rural America--and how? A critical public policy question for the 21st century. *Economic Review-Federal Reserve Bank of Kansas City*, *86*(2), 57.

Trajtenberg, M. (1987). *Patents, citations, and innovations: tracing the* links (working paper no. 2457). National Bureau of Economic Research, Cambridge, MA.

Wojan, T. R., Dotzel, K. R., & Low, S. A. (2015). Decomposing regional patenting rates: how the composition factor confounds the rate factor. *Regional Studies, Regional Science*, *2*(1), 535-551.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press, Cambridge, MA.