**Rural-Urban Divide in County Level Patent Applications**

Giri Aryal[a,b]
John Mann[a]
Scott Loveridge[a]
Satish Joshi[a]

[a]Department of Agricultural, Food, and Resource Economics
Michigan State University
446 W. Circle Drive, Room 87
East Lansing, MI 48824

[b]Corresponding Author
email: aryalgir@msu.edu

***Selected Paper Prepared for Presentation at the 2017 AAEA Annual Meeting, Chicago, Illinois***
***July 30-August 1, 2017***

# Rural-Urban Divide in County Level Patent Applications

## 1. Introduction

Innovation is central to economic competitiveness. A large body of literature has identified and analyzed economic and non-economic factors driving innovation, and variations in innovation-related outputs across time and regions. Population densities, critical mass of educated employees, research and development (R&D) expenditures by universities and private industries, innovation and communication infrastructure, and network externalities are all drivers of regional innovation (Artz et al., 2016; Guimaraes, 2015; Lee et al., 2010; Adelaja et al., 2009; Carlino et al., 2007; Barkley et al., 2006; Andersson et al., 2005; Audretsch & Feldman, 2004; Acs et al., 2002; Cooke et al., 2002; Anselin et al., 1997). Most innovation studies focus on big cities and metropolitan regions, and their roles in economic growth as they draw on the idea that innovation is primarily an urban phenomenon (Florida et al., 2016; Carlino et al., 2007; Acs et al., 2002; Feldman and Audretsch, 1999; Glaeser et al., 1992, 1995, 2010). This idea is based on the theoretical argument that urban regions are conducive to innovation due to their tendency to generate higher network externalities and knowledge spillovers, i.e., the agglomeration effect (Carlino et al., 2001, 2007).  Less studied are drivers of innovation in nonmetropolitan and rural areas. However, most empirical regional studies comparing rural innovation with that in urban areas conclude that that rural America lags in its innovation performance when compared to its metropolitan counterpart (Wojan et al., 2015; Zheng and Ejermo, 2015; Orlando and Verba, 2005; Porter et al., 2004).

At the same time, rural economies in the US continue to change with mixed results, for example from industrialization of agriculture, improved communication infrastructure, suburbanization and associated spillovers, immigration, and public policies promoting rural innovation and entrepreneurship (Woods, 2013; Kilkenny, 2010; Barkley et al., 2006; Porter et al., 2004; Atkinson, 2004). Increased globalization provides extended market opportunities to all regions, whether rural or urban, and reduces rural market constraints arising from low population densities (Woods, 2013). However, rural regions face globalization challenges which are illustrated by the observation of Munchin et al. (2003, p.3).

"… for myriad reasons, both economic and demographic in nature, rural areas have been seriously challenged by this shift, one that relocates the source of competitive advantage to now being primarily a function of knowledge; it is today commonplace to refer to industries competing in a knowledge-based economy".

Innovation-led competitiveness fuels the economic growth in today's knowledge-based economies and regions emerge as units of innovation organization (Florida et al., 2016). Further, innovation is driven by a host of regional characteristics. While urban regions benefit from positive information/network externalities due the proximity of like firms and other resources (Carlino et al., 2007; Jaffe et al., 1993; Lucas, 1988), negative externalities due to congestion and scarcity of land accrue as more people migrate to these areas to exploit the positive externalities (Atkinson, 2004 and Porter et al., 2004). On the other hand, rural regions may provide opportunities to reduce some of these negative urban externalities by offering large open spaces and to reinforce positive externalities, such as rural areas adjacent to urban centers supplying labor. Thus, urban and rural regions are inter-dependent sharing positive and negative externalities to innovation (Dabson, 2007, 2011). Several empirical studies conducted at state and metropolitan levels suggest that innovation occurs less frequently in rural areas than in urban areas (Packalan, 2015; Crescenzi and Rodriguez-Pose, 2013; Carlino et al., 2007; Ohuallachain, 2005, 1999; Akai and Sakata, 2002; Acs et al., 2002; Anselin et al., 1997; Feldman and florida, 1994; Jaffe, 1989). Barkley et al. (2006) conducted a county level analysis of the nonmetro area of the US Midwest and found that innovation occurs more frequently in the counties that are proximate to the urban areas.

In this study, we empirically explore the urban-rural gap in innovative activity and the factors creating the gap. We define innovative activity as the utility patent applications per capita at county level in the US. We incorporate a wide range of factors affecting regional innovation potential, including human capital, R&D capital, industrial specialization, communication infrastructure, and other economic and demographic factors. Recognition of the count nature of the patents data has been limited to firm and industry level studies in the innovation literature. To take this into account at regional level, we use negative binomial regression model. Using a panel data on 2,847 US counties between 2009 and 2013, we test the potential gap in rural-urban patenting based on a random effects negative-binomial regression model with only the main effects of the independent variables. To test the factors that contribute to the gap, we interact the independent variables found to affect the county-level patenting with the three-category rural-urban location indicator.

Our research builds on the theoretical concept of a "regional knowledge production function" developed by Griliches (1990) for the analysis of effects of R&D expenditures on firm-level innovation and adapted by Jaffe (1989) for the analysis of the knowledge spillover across innovating regions. Econometrically, we analyze patent application counts using the panel regression techniques developed by Hausman et al. (1984) and apply the random effects negative binomial modeling to the county level of observation, as fixed effects Poisson and negative binomial models were not suitable for our sample. Our analysis supports prior studies that found urban counties have higher patent frequencies than the rural area, both the metro-adjacent rural counties (non-metro counties adjacent to metro area) and remote rural counties (non-metro counties not adjacent to metro area). We find that rural counties patent less frequently than the metro counties (remote rural counties 75% less and metro-adjacent counties 84%[1] less) but the difference in patenting rates between metro-adjacent rural and remote rural counties is not statistically significant. Our other findings show that variables such as university expenditures, share of high-tech establishments, high-tech variety, share of information industry employment, and ethnic diversity are associated with the increasing patenting gap between the urban and rural (both metro-adjacent and remote) counties. However, higher level of high-tech industry specialization and share of agricultural employment[2] are associated with the reduced gap in patenting rates between urban and remote rural counties. These results suggest that policies focused on development of the clusters of high-tech industries and promotion of agricultural firms and industries to exploit the unutilized business potential in remote rural areas are likely to lead to increased innovation activities in these areas and to more equitable regional economic growth.

What follows is a brief review of the literature on innovation and patents as its measure, rural-urban differences in innovation, and the regional characteristics that affect innovation. We develop our conceptual and empirical models, and describe the data and estimation methods in section three. In section four, we present and discuss our findings. Section six concludes with our discussion of potential policy implications and areas for further study.

---

[1] The difference in patenting frequencies between metro-adjacent rural counties and metro counties is statistically significant only at 10% level. But the difference between remote rural and metro counties is significant at 5% level.
[2] The role of higher share of agricultural employment in reducing the gap between the urban and remote rural counties is statistically significant for the counties where agriculture industry provides less than 16% of the total jobs.

## 2. Review of Patenting and Regional Innovation and Related Factors

Innovation is an undisputed source of economic growth and regional development as new ideas, products, and processes drive regional prosperity (Acs et al., 2002; Feldman and Florida, 2005). Intellectual debates arise from the conceptual frameworks treating the role of innovation in growth. Neoclassical frameworks beginning with Solow's (1957) work attribute an exogenous role to the innovation. However, the endogenous growth framework following Arrow's (1962) work endogenize innovation and explains that growth sustains itself due to the continuous pursuits of economic agents, such as firms and industries, to create something new either in the form of improvement in existing quality or the development of completely new products, output markets, new sources of production, and new methods of production and organization (Cameron, 1996; Mann and Shideler, 2015). In the empirical literature, researchers have been facing the unresolved problem regarding the measurement of innovation (Acs et al., 2002; Cameron, 1996). However, patents are the most frequently used measures of innovation because the data are readily available and they continue to provide a reliable measure of regional innovation relative to other measures (Acs et al., 2002; Griliches, 1990).

Patent statistics have been used in the literature to understand their association both with the economic growth indicators and different innovation inputs. In one of the earliest empirical applications of the patent statistics, Scherer (1965a, 1965b) used a cross section of 448 US companies on Fortune's list during 1955-1959 to explain the profit growth of the firms in terms of their inventive output, number of annual patents issued to them, and vice versa. After controlling their size, industrial structure, and R&D spending, he found that the firms' patenting increased their profits and vice versa. Up until the 1990s, most studies that used patents were focused on firm or industrial level observations (Sincera; 1997, Griliches, 1990; Bound et al., 1984, Hall et al., 1984; Hausman et al., 1984; Pakes and Griliches, 1980). However, Florida (2016) suggests that regions (and cities) are appropriate units of analysis for studying innovation. He points out that firms and industry level analyses may overlook inputs such as knowledge spillovers and other potential benefits from agglomeration economies that can only been considered at the region level view.

At the regional level of analysis, most of the studies on innovation are conducted at the state level of observation (Ohuallachain, 2005; Akai and Sakata, 2002; Feldman and florida, 1994; Jaffe; 1989), metropolitan (Packalan, 2015; Crescenzi and Rodriguez-Pose, 2013; Carlino et al., 2007; Acs et al., 2002; Ohuallachain, 1999; Anselin et al., 1997; Jaffe, 1989), and labor market area (Niebuhr, 2010; Andersson et al., 2005) levels.  Based on the conceptual idea of knowledge production function that relates patent output with R&D and human capital as its knowledge inputs, these studies also include various other regional factors that enhance the productivity of the knowledge inputs through knowledge spillover effects. The common findings of these studies suggest that R&D and human capital investments are concentrated in urban areas due to the presence of knowledge externalities generated in the presence of denser populations, entrepreneurs, and businesses.

Innovation studies conducted with higher level of regional aggregation such as state and metropolitan areas as their units of analysis "inevitably obscure the spatial (innovation) processes that occur within a region or across its regional boundaries" (Feldman and Florida, 1994, p. 216). Thus, the top performing members of a regional unit potentially dominate the innovation performance of other members. In fact, the US counties, whether be metropolitan or nonmetropolitan, are heterogeneous in terms of their innovation and related geographical characteristics (Castle et al., 2011, 1995; McGranahan et al., 2011; Henderson and Abraham, 2004; Porter et al., 2004). County level analyses have been conducted to examine the association of patenting (Monchuk and Miranowski, 2010; Adelaja, 2009; Feser and Isserman, 2006) and innovative entrepreneurship (Low et al., 2013; McGranahan et al., 2011; Henderson and Weiler, 2010; Henderson and executive, 2007) with regional economic growth. Findings show positive association between the level of innovative activities, measured by patenting and entrepreneurship, and county regional growth during two decades after 1990. Some of these studies examining the role of the proximity factor on the county level growth have found that the spillover of growth benefits arising from entrepreneurship and innovation are stronger in counties that have denser population and are (more) proximate to metro counties than in distant counties (Stephens et al., 2013; Henderson and Weiler, 2010; Monchuk and Miranowski, 2010; Henderson and Executive, 2007; Feser and Isserman, 2006). However, other analyses at the county level found negative (Young et al., 2014) or no significant association between patenting and county level growth (Stephens et al., 2013).

Only a very few other researchers have studied county level patenting activity considering both urban and rural patent outputs and various regional characteristics including skilled and educated workforce, R&D expenditure, technological infrastructures, entrepreneurial environment, and socio-economic environment that that determine the differential outputs. Using Knowledge Production Framework (KPF), Zheng and Slaper (2016) estimate that industry and university R&D, human capital and venture capital inputs are positively associated with patent outputs across US counties between 2009 and 2011. Barkley et al. (2006) summarize patents data to illustrate that rural (non-metro) counties in the southern USA witnessed less innovative activities than the urban (metro) counties during 1990-1999. They also report that proximity of a non-metro county to a metro county did not result in significantly increasing patenting. However, after controlling several county and regional characteristics including education and skills of the labor force, structure and diversity of local economy, high-tech industry employment, patenting in the neighboring counties, and the proportion of large and small firms in corresponding LMA of a county, they found significantly positive association of proximity.

Henderson and Abraham (2004) identified higher population density as the key factor essential to supporting knowledge-based activities in rural America. Using the 1990-2000 data on 3053 US counties, they found that rural counties lagged the urban counties in terms of concentration of high knowledge occupations that contributed to a widening of the innovation gap. Among the 2,246 rural counties, high knowledge occupations were concentrated more in rural places having larger towns or higher population densities. They did not find significant correlation between the rural counties' proximity to the metropolitan area and their concentration of high-knowledge occupations, indicating that remoteness is a less formidable challenge in the development of knowledge-based rural economies.

Monchuk (2003) conducted a US county level analysis in the US Midwest and found that the sum of a county's patent applications during the period was positively associated with its average percent of population with college graduates, average per capita personal income, and proximity to metro regions. While these relationships held for sub-samples with intervals of five-year periods, prior period patenting had a positive association with current period patenting. Other literature also shows that American counties that are urbanized or proximate to the urban counties fare better in terms of benefits from economic growth (Stephens et al., 2013; McGranahan et al., 2011; Monchuk and Miranowski, 2010; Higgins et al., 2009; Henderson and Executive, 2007; Henderson and Weiler, 2007; Feser and Isserman, 2006; Henderson and Abraham, 2004; Monchuk, 2003;).

Our research contributes to the literature in fours ways. First, it is a county level analysis in the contiguous US, and not isolated to a specific region in the US. Second, it examines the rural-urban divide in terms of patenting and the regional factors that differentially affect the rural and urban patenting. Third, the examination considers recent data, during the period 2009-13, and examines if there exists rural-urban divide in patenting by controlling several local regional characteristics related to entrepreneurship, innovation, and growth. Fourth, this study applies count models used to analyze firm level innovation to the county level of analysis.

## 3. Research Design

The concept of KPF (Knowledge Production Function) that was introduced by Griliches (1979) provides useful framework for understanding innovation as the product of various resource inputs that are allocated toward generation of economic knowledge. Incumbent firms engage in their pursuit of new economic knowledge that is crucial for innovation. R&D expenditures serves as one of the key inputs that generates new economic knowledge. Patents serve as the output of the knowledge production function and a reasonable proxy for innovation (Griliches, 1990 and Czarnitzki et al., 2009). In his pioneering work, Jaffe (1989) applied the KPF to study regional innovation and knowledge spillovers by shifting the units of observation from firms to the geographic entities. The Griliches-Jaffe KPF has been widely applied in a large number of empirical studies on regional innovation and knowledge spillovers and is popular in the regional literature as regional knowledge production function, RKPF (see Charlot et al., 2015. Although the large body of literature has focused on the role of various R&D capital inputs in regional innovation and knowledge spillovers, several studies have extended the RKPF to include the region-specific factors such as human capital and socio-economic and demographic factors that influence the regional innovative output (Charlot et al., 2015, Bluesa et al., 2010; Ponds et al., 2010, and Varga, 2000). The extended model of RKPF that takes into the factors such as human capital and several regional conditions that give rise to regional differences in innovation can be represented as

$$K = f(\boldsymbol{RDK},\ HK,\ \boldsymbol{z}) \tag{1}$$

where new economic knowledge, *K*, is the product of regional innovation process generated by the R&D capital inputs, *RDK*, and human capital inputs, *HK*, that are conditioned by several region-specific characteristics, **z**.

We test the rural-urban differences in innovation under the extended RKPF framework by estimating the average number of annual utility patents produced by rural and urban US counties after controlling the innovative inputs in (1), county level regional characteristics, and the state level fixed effects. To identify the potential factors that explain the regional differences in innovation, we interact the independent variables found to affect the county-level patenting with the three-category rural-urban regional indicator.

### 3.1 Data

This study uses longitudinal data on 2,847 counties located in 48 contiguous states of the USA excluding District of Columbia between 2009 and 2013. The data come from secondary sources including, United States Patent and Trademark Office (USPTO), Community Business Patterns (CBP), American Community Survey (ACS) and Current Population Survey (CPS) under the US Census Bureau, US Bureau of Economic Analysis (BEA), Economic Research Service (ERS) and Census of Agriculture under the US Department of Agriculture (USDA), Small Business Innovation Research (SBIR), and National Science Foundation (NSF).

Table 1 provides the definition of the variables and their data sources. The number of the domestic utility patent applications at county level for the post-recession period between 2009 and 2013 serves as the dependent variable and as our measure of the rate of innovation in the US counties. We aggregated the patent applications originating from residential zip codes of the primary inventors to derive the county level patent applications. The details on the various sources of USPTO data, HUD USPS zip-to-county crosswalk and the aggregation process are included in appendix A2.

**Table 1: Definition of Variables and Data Source**

| Variables | Definition [a] | Source | Year |
|---|---|---|---|
| Patent Applications Per Capita | Number of annual utility domestic patent applications per 10k population in the US | USPTO | 2009-13 |
| University R&D | University R&D expenditures | NSF | 2006-13 |
| Spatially Lagged University R&D | University R&D expenditures in the neighboring counties within 100-mile radius | | |
| Business R&D | State level R&D expenditures by private businesses | NSF | 2009-13 |
| SBIR awards | Amount of innovation research awards to small businesses | SBIR | 2009-13 |
| Percent Bachelor Plus (Human Capital) | Share of bachelor or higher degree holders in total population 25 years and over | Census Bureau (ACS) | 2009-13 |
| Ethnic Diversity | Herfindahl-Hirschman Index of racial mix in the population | Census Bureau (ACS) | 2009-13 |
| Personal Income Per Capita | Per capita personal income | BEA | 2009-13 |
| Percent Foreign-born Population | Share of “naturalized US citizens” and “not US citizens” in total population | Census Bureau (ACS) | 2009-13 |
| Percent Agricultural Employment | Share of civilian population 16 years and over employed in agriculture, forestry, fishing and hunting, and mining industries | Census Bureau (ACS) | 2009-13 |
| Percent Manufacturing Employment | Share of civilian population 16 years and over employed in ‘manufacturing’ industries | Census Bureau (ACS) | 2009-13 |
| Percent Information Employment | Share of civilian population 16 years and over employed in ‘information’ industries | Census Bureau (ACS) | 2009-13 |

| | | | |
|---|---|---|---|
| Percent Financial Employment | Share of civilian population 16 years and over employed in ‘financial’ industries | Census Bureau (ACS) | 2009-13 |
| Percent Professional Employment | Share of civilian population 16 years and over employed in ‘professional’ industries | Census Bureau (ACS) | 2009-13 |
| Establishment Size | Average number of employees per business establishment | Census Bureau (CBP) | 2009-13 |
| Industrial Specialization | Herfindahl-Hirschman Index of the mix of business establishments by two and three-digit NAICS | Census Bureau (CBP) | 2009-13 |
| Percent High-tech Establishments | Share of high-tech in total business establishments | Census Bureau (CBP) | 2009-13 |
| High-tech Variety | Number of three-digit NAICs for high-tech industry categories | Census Bureau (CBP) | 2009-13 |
| Percent Net Business Formation | Net change in the number of business establishments as percent of previous year | Census Bureau (CBP) | 2009-13 |
| Percent net Job Creation | Net change in number of jobs as percent of previous year | Census Bureau (CBP) | 2009-13 |
| High-speed internet access | Number of households out of 1,000 households having high-speed internet | Census Bureau (CPS) | 2009-13 |
| County classification | Metro, metro-adjacent rural, and remote rural classification | USDA (ERS) | 2013 |

[a]Detailed discussion about the variables used in the analysis is provided in the empirical model section.

The university R&D expenditure data from NSF are available at city level. we 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 present college(s) or university(ies) do not spend in R&D activities, we assume in this study that the county has zero university R&D. To account for the fact that expenditure on a university research project is likely to be a continuous process, we follow Hausman et al. (1984) and control past annual university R&D expenditures by using three-time lags of the variable. For example, the R&D in 2010, 2011, and 2012 correspond to the three lags of the observation for 2013 and the R&D in 2006, 2007, and 2008 correspond to the lags of the observation for 2009. Additionally, we account the spillover effects of the university

R&D from the counties hosting the university/colleges to their neighboring counties by deriving a spatially lagged university R&D variable based on a distance decay function within 100 miles from the county centroids. More detail on the derivation of the spatial lags is included in the appendix (A2). We use the NSF data on the state level R&D expenditures by the private businesses because the data are not available at county level.

The data on the households with high-speed internet access come from Current Population Survey of the Census Bureau in five ordinal scales with intervals of 200 for a block of 1,000 households. For example, a county is assigned with a category 1 if it has less than 200 households in 1,000 are connected with high-speed internet, 2 if 200 to 400, and so on with 5 if more than 800. In our regression analysis, we categorize the counties into two categories – the counties with more than 800 out of one thousand households having high-speed internet and the remaining counties, where we term the former category as "high internet access".

We retrieved the firms-related data including the share and variety of high-tech industries from US Census Bureau's Community Business Patterns (CBP). The number of total business establishments were derived by summing it at the three-digit level of industry codes across the 2012 North American Industry Classification System (NAICS), 2012. We 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 the share foreign-born population, higher education, and civilian employment in agricultural, manufacturing, information, financial and professional service industries[3] come from the five-

[3] ACS surveys 295,000 households randomly each year and no household is repeated 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

year estimates for 2013 of American Community Survey (ACS). We combine the population of "naturalized US citizen" and "not a US citizen" to derive the foreign-born population variable.

To derive the industrial specialization, we follow Carlino et al. (2007) and derive the Herfindahl-Hirschman Index of the mix of business establishments corresponding to two and three-digit level NAICS industries using the CBP data as follows:

$$Industrial\ Specialization = \sum_{n=1}^{N_i} s_{int}^2$$

where $s_{int}$ is the share of establishments under two and three-digit level NAICS industry *n* in region *i* in year *t*. Higher value of the index represents higher ethnic diversity. The higher index value represents the presence of more specialized industries and lower value represents more diversified industries.

The ethnic diversity variable was derived as Herfindahl-Hirschman Index following Rupasingha et al. (2002) and using ACS population mix data as:

$$Ethnic\ Diversity = 1 - \sum_{n=1}^{N_i} s_{int}^2$$

where $s_{int}$ is the share of population with race *n* in total population of region *i* in year *t*; $N_i = \{Black, White, Asian\ and\ Pacific\ Islander, American\ Indian, and\ other\}$. Higher value of the index represents higher ethnic diversity.

In our attempt to proxy the entrepreneurship with start-up businesses, we derived the percent change in the number of establishments and employments from previous year using CBP data. Our measure essentially represents the net change (number of new establishments less number of deaths of existing establishments) because we do not have the data at county level on the business death rates.

We classify counties by combining the codes in Rural-Urban Continuum Codes (RUCC), 2013 developed by ERS unit of USDA[4]. Codes 1, 2, and 3 combined form the metropolitan counties, 4, 6, and 8 the non-metro counties adjacent to metro areas, and 5, 7, and 9 the non-metro counties not adjacent to metro areas[5]. For the analysis in this study, we consider metro counties as "urban", non-metro counties adjacent to metro area(s) as "metro-adjacent rural", and non-metro counties not adjacent to metro areas as "remote rural".

### 3.2 Empirical Model and Econometric Estimation

We denote the number of annual patent grants to a county *i* in year *t* by $P_{it}$ . To account for the count nature of the patents data, we first start with an assumption that $y_{it} \sim P[\mu_{it} = \alpha_i \lambda_{it}]$. That is, patent applications by each county is drawn from a Poisson distribution with parameter $\mu_{it}$, which varies across both counties and years. The probability that $P_{it} = c$ is given by

$$f(\mathrm{p};\mu) = \Pr(P_{it} = c) = \frac{\mu_{it}^{c}\, e^{\mu_{it}}}{c!} \;\; ; \;\; \mathrm{i} = 1 \text{ to } 2847, \;\; \mathrm{t} = 2009, \ldots, 2013$$

$$\text{and } c = 1, 2, 3, \ldots \qquad (2)$$

[4] We accessed 2013 RUCC from https://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx

[5] 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

We specify our Poisson regression models with multiplicative county specific term $\alpha_i$ representing county-level heterogeneity, say, in terms of propensity to patent and exponential mean function $\lambda_{it} = \exp(\boldsymbol{X'}_{it}\boldsymbol{\beta})$. As a result, our Poisson regression specification becomes

$$\log(\mu_{it}) = \alpha_i + \boldsymbol{X'}_{it}\boldsymbol{\beta} + u_{it} \tag{3}$$

where $\boldsymbol{X_{it}}$ represents the vector of time-varying predictor variables, $\alpha_i$ represents unobserved county-specific effects, and $u_{it}$ is the disturbance term.

The Poisson log-likelihood are obtained by transforming the probability function (1) "for which the parameters are estimated to make the given data most likely" (Hilbe, 2011 p.80) as

$$\mathcal{L}(\mu; p) = \sum_{i=1}^{n}\sum_{t=1}^{n_i}[P_{it}(\boldsymbol{X'}_{it}\boldsymbol{\beta} + \alpha_i) - \exp(\boldsymbol{X'}_{it}\boldsymbol{\beta} + \alpha_i - \ln\Gamma(P_{it} + 1)\Gamma(P_{it} + 1)] \tag{4}$$

where $\ln\Gamma(\ )$ is the log-gamma function.

There are two different approaches that differ in terms of how they address the individual heterogeneity in maximizing the log-likelihood function (3) and estimating the parameters. Fixed effects (FE) estimators assume separate county-specific unobserved parameters, $\alpha_i$'s, for each county and these parameters can be correlated with the observed predictor variables $\boldsymbol{X_{it}}$. Standard (unconditional) fixed effects estimators using the maximum likelihood estimation method are likely to be inconsistent in regression models with short panels (Cameron and Trivedi, 1998). A very widely used strategy following Hausman et al. (1984) avoids the inconsistency due to the correlation by eliminating $\alpha_i$ from the equation (4) by means of conditioning on the sum of patents $\sum_{t=1}^{T} P_{it}$. Thus, they are known as conditional FE Poisson estimators. On the other hand, random effect (RE) models assume that $\alpha_i$'s are iid and not correlated with the observed predictors. The RE models incorporate $\alpha_i$'s into the error term and estimate the resulting model

$$\log(\mu_{it}) = \boldsymbol{X'}_{it}\boldsymbol{\beta} + v_{it} \qquad \text{where} \quad v_{it} = (\alpha_i + u_{it}) \tag{5}$$

Conditional FE Poisson models are estimated by maximizing the conditional log-likelihood function in which the individual effects are eliminated (Hausman et al., 1984, Hilbe, 2011)

$$\mathcal{L}(\beta; P) = \sum_{i=1}^{n} \begin{bmatrix} ln\Gamma\left(\sum_{t=i}^{n_i} P_{it} + 1\right) - \\ \sum_{t=i}^{n_i} ln\Gamma(P_{it} + 1) + \sum_{t=i}^{n_i}\left(P_{it}\ (\boldsymbol{X'}_{it}\boldsymbol{\beta}) - ln\ \sum_{l}^{n_i} \exp(\boldsymbol{X'}_{it}\boldsymbol{\beta})\right) \end{bmatrix} \tag{6}$$

However, the log-likelihood of the RE model 5 takes a different form based on the assumption about the distribution of the individual-specific effects. We assume that $\alpha_i$'s are drawn from a gamma distribution with each county panel independent of the others. The log-likelihood of the RE model becomes (see Hausman et al., 1984 and Hilbe, 2011 for derivation)

$$\mathcal{L}(\beta; P) = \sum_{i=1}^{n} \begin{bmatrix} ln\Gamma\left(\theta + \sum_{t=i}^{n_i} P_{it}\right) - ln\Gamma(\theta) - ln\Gamma(P_{it} + 1) + \theta \ln(u_i) + \left(\sum_{t=1}^{n_i} P_{it}\right) \\ \ln(1 - u_i) - \left(\sum_{t=i}^{n_i} P_{it}\right) \ln\left(\sum_{t=i}^{n_i} \exp(\boldsymbol{X'}_{it}\boldsymbol{\beta})\right) + \sum_{t=i}^{n_i} P_{it}\ (\boldsymbol{X'}_{it}\boldsymbol{\beta}) \end{bmatrix} \tag{7}$$

where $\theta$ is the variance of the gamma distribution of $\alpha_i$ and $u_i = \left({\theta}\big/{\left(\theta + \sum_{t=i}^{n_i}\left(exp(\boldsymbol{X'}_{it}\boldsymbol{\beta})\right)\right)}\right)$.

Poisson RE model 5 is estimated by maximizing the log-likelihood function (7).

In Poisson models, the conditional mean and variance of each observation of patent count are restricted to be equal. That is, $(P_{it}) = var(P_{it}) = \lambda_{it}$. When this restriction is violated in the counting process, the estimated results do not become useful for making inferences. In fact, in most of the cases count data are over dispersed due to the positive correlation among the outcomes of the response variable, the excess variation among the outcomes and the assumptions about the data distribution (Hilbe, 2011). The over dispersion can be addressed by including additional term in the Poisson models that allows variance to exceed the mean. This is achieved by using the models based on Negative Binomial (NB) distributional assumption about the data. The log-likelihoods (6) and (7) for the negative binomial models include the dispersion parameter. Hausman et al. (1984) developed the fixed-effect version of the negative binomial model by allowing the dispersion parameter (variance to mean ratio) in the RE Poisson model to grow with the mean and by incorporating conditional fixed effects $\alpha_i$ along with its possible correlation with the independent variables $\boldsymbol{X_{it}}$ **.**

The *conditional FE and RE* negative binomial models differ from the models 3 and 5 by including an additional parameter that allows the variance to exceed the mean. Additionally, unlike the $\alpha_i$'s in RE Poisson model 5 following gamma distribution, $\alpha_i$'s in RE negative binomial models are assumed to follow beta distribution. In both *conditional FE* and *RE* negative binomial models the dispersion parameter is assumed to be constant within the same county over the observation period. Its variation across the counties is assumed to be random across counties in the RE negative binomial but it can take any value in *FE* because it is conditioned out of likelihood function as shown in Poisson model. The FE models are estimated using conditional likelihood method and the RE models using maximum likelihood method. Regression analysis was performed using Stata. Unlike the conditional FE Poisson regressions, conditional FE negative binomial regressions provide the estimates of the effects of observed time-invariant predictors. So, these models do not serve as "true fixed-effects" models like Poisson models (Allison and Waterman, 2002). They suggest estimating the unconditional (with dummy variables for the counties) fixed effects negative binomial models. There doesn't exist any theoretical proof of whether the parameter estimates suffer from incidental parameter bias and therefore inconsistency. With the help of simulation of a two-time period for 100 individuals and

500 samples, they find little evidence of such problem. However, our attempts to run such a regression failed due to computational challenges arising from our large dataset.

In our empirical analysis, application of the fixed effects negative binomial model would leave out 41% of the observations corresponding to the counties with zero patents in every year during 2009-13. Therefore, we estimate of the following empirical extension of model 5

$$\log(\mathrm{Pat}_{it}) = \boldsymbol{x}'_{it}\boldsymbol{\beta} + \boldsymbol{z}'_{it}\boldsymbol{\gamma} + v_{it} \qquad \text{where} \quad v_{it} = (\alpha_i + u_{it}) \tag{8}$$

where $\mathrm{Pat}_{it}$ , annual utility patent applications, represents the knowledge output K in regional knowledge production model 1; $\boldsymbol{x}_{it}$ represents the vector of innovation inputs – county level university R&D expenditures and their spatial and time lags, state level private business R&D, and human capital; and $\boldsymbol{z}_i$ represents the vector of regional characteristics and the interaction of some relevant characteristics with regions. We hypothesize that the regional patenting depends on the levels of university R&D, business innovation awards, population density, industrial diversity (both total businesses and high-tech businesses), and sectoral employments (population employed in agriculture and manufacturing). Therefore, we form interactions of these variables with the regions. The pool of educated population measures human capital (*HK*) in this study. Following Monchuk (2003), we use the share of population 25 years and over with bachelor's or higher degree in county population as a measure of human skills and innovative capabilities

**Regional Characteristics/Controls***:* We control the regional characteristics at the county level. We follow Monchuk and Miranowski (2010), Barkley et al. (2006), Andersson et al. (2005), Rupasingha et al. (2002), and Monchuk (2003) for our choice of variables to measure the regional characteristics. The county level controls include firm characteristics, entrepreneurial/innovative environment, industrial characteristics, and economic and demographic characteristics. The average size of the business establishments (total employment divided by the total number of establishments) represent the firm characteristics at the county level. The range of the number of households with access to the high-speed internet and the share of civilian population 16 years and older employed in information sector represent the presence of communication infrastructure for innovation. The percent change in number of establishments and employments and the shares of employment in financial and professional service industries constitute the entrepreneurship environment.

The industrial specialization, share of high-tech establishments, varieties of high-tech industries at three-digit level of NAICs, share of civilian population16 years and older employed in agricultural, forestry, fishing and hunting, and mining industries), and share of civilian population16 years and older employed in manufacturing industries constitute the industrial characteristics at county level. Following Carlino et al. (2007), we derived the industrial specialization variable by calculating the Herfindahl-Hirschman Index (HHI) as the sum of the squares of the industry employment shares in counties, at two and three-digit levels in our study. The higher index value represents the presence of more specialized industries and lower value represents more diversified industries. The similar derivation of HHI for measuring the specialization of high-tech industries would not allow us to distinguish between the counties with zero employment and those with a completely specialized industry, as several counties in our sample have zero high-tech employment. We 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 35, 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. Per capita personal income reflects the regional economic prosperity. Ethnic diversity and share of foreign-born population constitute demographic characteristics.

## 4. Empirical Results

### 4.1 Summary Statistics

The definitions of the variables summarized in the table 2 come from table 1. The first group of independent variables enter the equations 9 and 10 in their natural logarithmic transformation but the second group in their level form. A casual look at the patent applications data would reveal that metro (urban) counties dominate the rural counties remarkably both in terms of overall period average per capita patent applications (table 1) and annual per capita averages (figure 1) the adjacent rural counties have higher average patenting rates than the remote rural counties. One common observation among three types of counties is that there exists a large deviation from average values of per capita patenting and most of the independent variables within each group.

**Table 2: Summary Statistics**

| Variables | Combined | | Metro | | Adj. Rural | | Remote Rural | |
|---|---|---|---|---|---|---|---|---|
| | Mean | S.D. | Mean | S.D. | Mean | S.D. | Mean | S.D. |
| *Dependent Variable* | | | | | | | | |
| Patent Applications per 10k Population | 2.23 | 28.59 | 5.14 | 45.63 | 0.55 | 5.04 | 0.23 | 1.60 |
| *Independent Variables (I)* | | | | | | | | |
| University R&D ($1M) | 21.75 | 151.67 | 53.83 | 239.90 | 1.58 | 15.47 | 1.64 | 17.42 |
| Spatially Lagged University R&D ($1M) | 22.97 | 105.41 | 48.03 | 164.39 | 10.79 | 27.79 | 2.97 | 9.92 |
| SBIR Awards ($1M) | 0.66 | 5.24 | 1.65 | 8.31 | 0.02 | 0.16 | 0.05 | 0.76 |
| Establishment Size | 11.49 | 4.21 | 13.28 | 4.42 | 11.03 | 3.40 | 9.55 | 3.78 |
| Per Capita Personal Income ($1k) | 35.58 | 9.40 | 37.83 | 9.65 | 32.73 | 6.74 | 35.91 | 10.80 |

| *Independent Variables (II)* | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Percent Bachelor Plus (pct. points) | 8.48 | 3.48 | 9.86 | 3.69 | 7.11 | 2.53 | 8.21 | 3.46 |
| Ethnic Diversity (Herfindahl_Hirschman Index) | 0.27 | 0.18 | 0.32 | 0.18 | 0.25 | 0.18 | 0.23 | 0.17 |
| Industry Specialization (Herfindahl-Hirschman Index) | 0.12 | 0.06 | 0.11 | 0.04 | 0.12 | 0.05 | 0.12 | 0.08 |
| Percent High-tech Establishments (pct. points) | 5.27 | 2.65 | 6.42 | 2.87 | 4.55 | 2.10 | 4.55 | 2.38 |
| High-tech Variety | 14.20 | 9.56 | 21.20 | 10.56 | 10.94 | 5.27 | 8.45 | 5.17 |
| Percent Net Business Formation (pct. points) | -0.82 | 3.50 | -0.71 | 2.50 | -1.05 | 3.71 | -0.69 | 4.33 |
| Percent Net Job Creation (pct. points) | -0.62 | 8.13 | 0.63 | 5.46 | -0.97 | 7.87 | -0.21 | 10.98 |
| Percent Foreign-born Population (pct. points) | 4.27 | 5.13 | 5.92 | 6.29 | 3.20 | 3.64 | 3.25 | 4.21 |
| Percent Agricultural Employment (pct. points) | 6.24 | 6.48 | 2.76 | 3.22 | 6.52 | 5.31 | 10.69 | 8.11 |
| Percent Manufacturing Employment (pct. points) | 12.87 | 7.03 | 12.36 | 5.70 | 14.88 | 7.47 | 11.16 | 7.57 |
| Percent Information Employment (pct. points) | 1.55 | 0.85 | 1.82 | 0.79 | 1.33 | 0.72 | 1.43 | 0.95 |
| Percent Financial Employment (pct. points) | 4.83 | 1.97 | 5.93 | 2.19 | 4.13 | 1.31 | 4.15 | 1.59 |
| Percent Professional Employment (pct. points) | 6.54 | 3.06 | 8.56 | 3.01 | 5.61 | 2.48 | 4.86 | 2.04 |
| Number of Observations | 14,235 | | 5,490 | | 4,770 | | 3,975 | |

Note: [a] The private business R&D expenditures are averaged at state level

Figure 2 shows that the distribution of the patents per capita is highly skewed to the right. Roughly two-third of the county-year observations (9,097 observations out of total 14,235 county-year observations) have zero patent applications and 0.2% of the observations (28 observations) have more than 100 patents per 10k population.

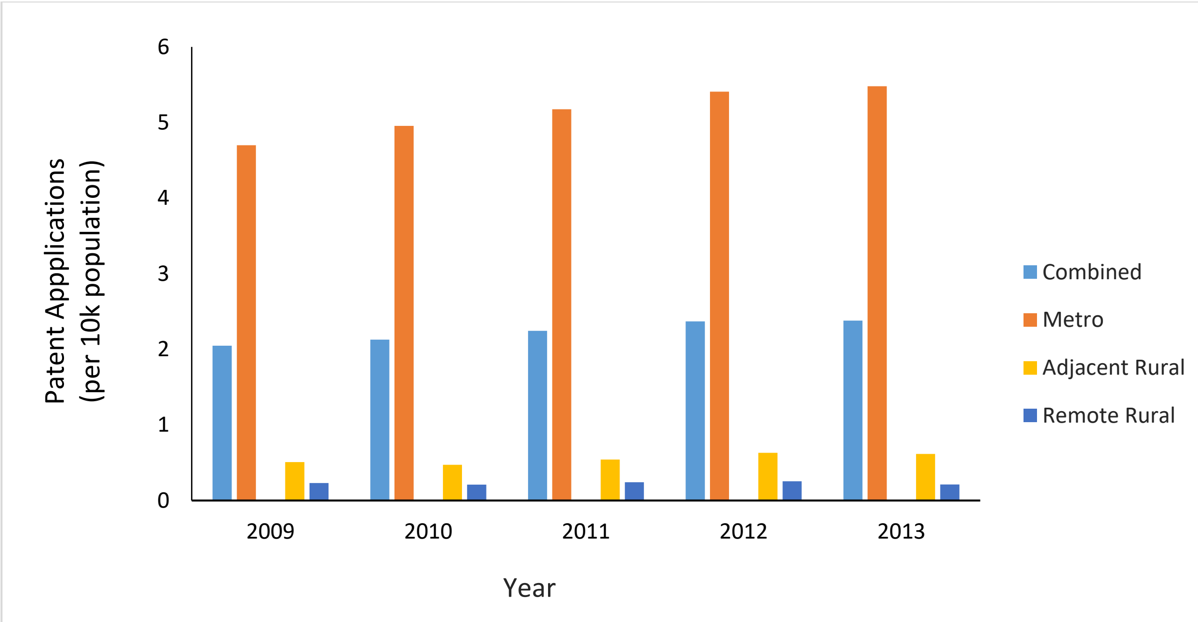


**Figure 1: Regional Trend of Patenting Intensity (2009-2013)**

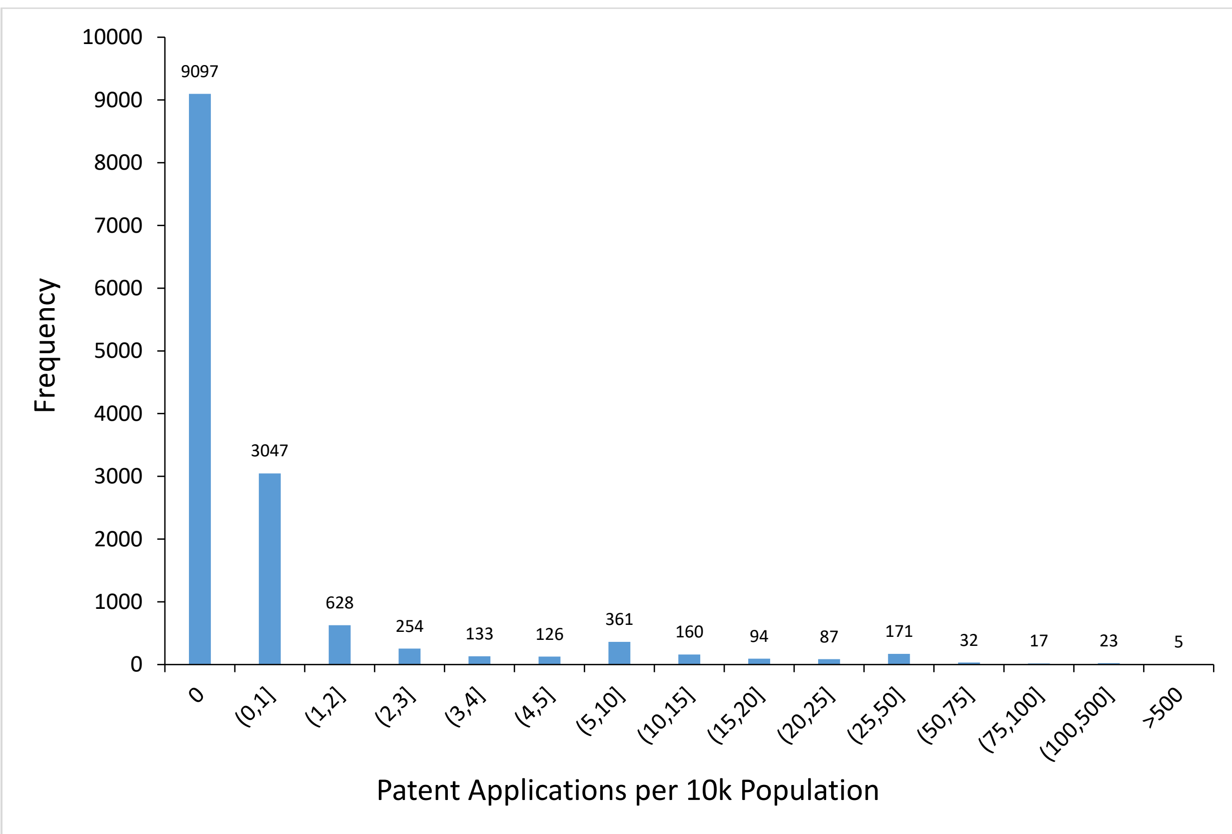


**Figure 2: Frequency Distribution of County-Year Patent Applications per 1k Population (2009-13)**

Table 1 also shows that metro counties have higher average values for university R&D expenditure, share of population with college education, amount of SBIR awards, share of high-tech industries, varieties of high-tech firms, firm size, and foreign-born share of population than the rural counties. Summary statistics and graphs suggest that counties are heterogeneous both in terms of patenting intensity and innovation inputs within urban and rural regions.

### 4.2 Regression Results and Discussion

Initially, we conducted our empirical analysis on the results from estimation of empirical model (8) using both Poisson and negative binomial regressions. we found significant over-dispersion of patents in our sample. This implies that the conditional variance of the patents exceeds its conditional mean. A Poisson model data underestimates the standard errors of the coefficient estimates in Poisson regressions. It induces more chances of rejection of the null hypotheses when, in fact, they are true (Hilbe, 2011). Therefore, negative binomial model is an appropriate choice for our analysis. To model the county heterogeneity effects, we ran both the conditional fixed and random effect models. The conditional maximum likelihood method for estimation of conditional fixed effects negative binomial model that was proposed by Hausman et al. (1984) drop 41% of the counties that didn't patent at all during our study period, that is when $\sum_{t=1}^{T} Pat_i$=0. Fixed-effect models should essentially control the unobserved stable covariates. Further, no "true" conditional fixed effect for negative binomial model exists in the literature (Allison and Waterman, 2002). Thus, we have chosen random effects negative binomial models for our analysis in this study.

### 4.2.1 Urban-Rural Difference in Innovation

Using the random effect model 8, we first ran the main effects model (no interactions) to examine the difference in patenting among urban and two rural county types. The results from the model with metro counties as the base category and the rural counties as the comparison categories (table 3) suggest that patenting is likely to occur 84% $\{(e^{-0.169} - 1) * 100\}$ less frequently in metro-adjacent counties and 75% $\{(e^{-0.286} - 1) * 100\}$ less frequently in remote rural counties than the metro counties (see Hilbe, 2011 and Abramovsky, 2007 for interpretation). However, the coefficient for the metro-adjacent category is statistically significant only at 10% level. The results of the similar model with metro-adjacent counties as the base category and metro and remote rural counties as the comparison categories indicate no statistically different patenting activities between the metro-adjacent and remote rural counties. These results from the model without interaction effects support the notion in the literature that innovation activities take place more in urban areas than the rural. The time lags of the University R&D and its spatial lag were tested for their potential effect on county level patenting but were dropped as these variables were not statistically significant and contaminated the effects of the contemporaneous levels of the variables. Similarly, private business R&D expenditures were also not found to be statistically significant and did not add to the model fit based on the AIC criterion; therefore, it was dropped from our analysis.

From the model with only the main effects, we found that university R&D expenditures and its spatial lag, human capital (percent population with college or higher degree), average establishment size, per capita personal income, ethnic diversity, percent high-tech establishments, and high-tech variety, and shares of employment in information and professional industries were positively associated with county level patenting while we found that share of employment in agriculture industry had negative association. However, our main interest lies in investigating the factors that are associated with the gap in patenting among urban and rural counties. So, rather than undertaking extended discussion of these results, we present our results from further investigation of a model with interaction effects in the next subsection.

**Table 3: Random Effects Negative Binomial Regression Results with Only Main Effects on County Level Innovation**

| *Dep. Variable:* Annual Domestic Utility Patent Applications per 10k Population | Coefficient Estimates | S.E. | z | P>z |
|---|---|---|---|---|
| *Independent Variables (I)* [a] | | | | |
| University R&D | 0.017 | 0.005 | 3.43 | 0.001 |
| Spatially lagged University R&D | 0.022 | 0.010 | 2.19 | 0.028 |
| SBIR Awards | 0.004 | 0.003 | 1.51 | 0.130 |
| Per Capita Personal Income | 1.032 | 0.220 | 4.68 | 0.000 |
| Establishment Size | 0.424 | 0.122 | 3.48 | 0.001 |
| *Independent Variables (II)* [b] | | | | |
| Metro-Adjacent Rural (1=yes; 0=metro) | -0.169 | 0.100 | -1.69 | 0.092 |
| Remote Rural (1=yes; 0=metro) | -0.286 | 0.124 | -2.30 | 0.021 |
| Percent Bachelor Plus | 0.067 | 0.012 | 5.44 | 0.000 |
| Ethnic Diversity | 0.812 | 0.330 | 2.46 | 0.014 |
| Percent Foreign-born Population | 0.009 | 0.009 | 0.97 | 0.330 |
| Industrial Specialization | 5.686 | 3.470 | 1.64 | 0.101 |
| Percent High-tech Establishments | 0.068 | 0.014 | 5.00 | 0.000 |
| High-tech Variety | 0.038 | 0.006 | 6.82 | 0.000 |
| Percent Agricultural Employment | -0.065 | 0.010 | -6.30 | 0.000 |
| Percent Manufacturing Employment | 0.006 | 0.008 | 0.73 | 0.464 |
| Percent Information Employment | 0.065 | 0.032 | 2.03 | 0.042 |
| Percent Financial Employment | -0.016 | 0.017 | -0.98 | 0.328 |
| Percent Professional Employment | 0.076 | 0.015 | 5.17 | 0.000 |
| High Internet Access (1=yes; 0=no) | -0.042 | 0.027 | -1.56 | 0.118 |
| Percent Net Business Formation | -0.002 | 0.006 | -0.29 | 0.775 |
| Percent Net Job Creation | 0.001 | 0.003 | 0.32 | 0.752 |
| Constant | -12.499 | 2.260 | -5.53 | 0.000 |
| Year Fixed Effects | | Yes | | |
| State Fixed Effects | | Yes | | |
| # Observations | | 14,235 | | |

### 4.2.2 Factors Explaining the Urban-Rural Difference in Innovation

We interacted the urban and rural county category with the variables of our interest based on the literature and our results from the main effects model- university R&D, SBIR awards, human capital, ethnic diversity, foreign-born population, share of high-tech establishments, high-tech industry variety, and shares of employment in agricultural, information, and professional industries. We present these results in table 4, which essentially contains the coefficient estimates from the negative binomial regression of a single model with all counties in the sample. In this model metro counties serve as the base category and metro-adjacent and remote rural as the comparison categories. Modeling interaction terms, we are fundamentally hypothesizing that the patenting among urban and rural regions depends on the levels of the interacting variables. Column 1 of table 4 contains the coefficients estimates representing the main effects of the independent variables; column 2 and column 3 consist of the coefficient estimates for the terms of interaction of corresponding independent variables with the metro-adjacent and remote rural county categories respectively.

**Table 4: Random Effects Negative Binomial Regression Results with Interaction Effects on County Level Innovation**

| *Dependent Variable:* Annual Domestic Utility Patent Applications in the US | Coefficient Estimates | | |
|---|---|---|---|
| | Baseline Model[a] | Metro-adj Rural | Remote Rural |
| *Independent Variables (I)* [b] | | | |
| University R&D | 0.015** | -0.005 | -0.009 |
| Spatially lagged University R&D | 0.022 | 0.005 | -0.005 |
| SBIR Awards | 0.004 | | |
| Per Capita Personal Income | 1.343*** | 0.136 | -2.071*** |
| Establishment Size | 0.509*** | 0.212 | -0.429* |
| *Independent Variables II* [c] | | | |
| Percent College or Higher Education | 0.069*** | -0.034 | 0.061 |

| | | | |
|---|---|---|---|
| Ethnic Diversity | 1.405*** | -1.363*** | -1.256** |
| Percent Foreign-born Population | 0.002 | | |
| Industry Concentration | 3.571 | | |
| Percent High-tech Establishments | 0.066*** | -0.010 | -0.039 |
| High-tech Variety | 0.033*** | 0.001 | 0.023 |
| Percent Agricultural Employment | -0.072*** | 0.030 | 0.004 |
| Percent Manufacturing Employment | 0.005 | | |
| Percent Information Employment | 0.104* | -0.032 | -0.149* |
| Percent Financial Employment | -0.017 | | |
| Percent Professional Employment | 0.069*** | 0.015 | -0.074 |
| High Internet Access (1=yes; 0=no) | -0.033 | | |
| Constant | -15.659*** | 1.696 | 22.733*** |
| Year fixed effects | | Yes | |
| State fixed effects | | Yes | |
| # Observations | | 14,235 | |

[a] *The Estimates for metro counties as the base category in the model with both main and interaction effects*
[b] *Variables in natural log;* [c] *Level Variables;* *** $p<0.01$, ** $p<0.05$, * $p<0.1$

The state fixed effects and year fixed effects that we control in our regression are not presented in table 4 but reported in Appendix table A1. The time lags of the university R&D and its spatial lag have not been controlled in our analysis due to the similar reasons as in main effects model.

The coefficient estimates in table 4 corresponding to the variables in logarithmic form can be interpreted as the elasticities because we have the log-link function in our count model. The other coefficients are interpreted as the change in the log of the odds ratio of the patents that is associated with the log of the odds ratios of the corresponding variables (see Hilbe, 2011). Alternatively, following Abramovsky (2007), these coefficients can be converted to incident rate ratios (IRR) and interpreted as percent change. However, the interpretation of the coefficient estimates for the interacting variables is less straightforward. Even with a linear model (OLS estimation), the results for the interacting variables cannot be directly interpreted as slopes like in usual models with only main effects because the fundamental hypothesis behind interaction is that the slope of one variable depends on the value (s) of the variable(s) it interacts with. In the non-linear model like Poisson or negative binomial, the reported estimation results for the interaction terms (both the coefficients and standard errors) from the available estimation methods do not represent the actual interaction effects and their statistical significance (see Ai & Norton, 2003; Hilbe, 2011). Alternatively, it is not unlikely that the actual interaction effects have different sign and statistical significance than as reported in table 4. We derive the actual interaction effects and standard errors and p-values following Hilbe (2011, p. 520-29) and using representative range of values of the interacting variables within our sample. In this way, we get different interaction effects (both magnitudes and significance) for different levels of the interacting variables. For, example by interacting university R&D variables with metro-adjacent rural and remote rural county categories we are hypothesizing that the regional patenting intensity depends on the level of the university R&D variable. Thus, we get different gaps between the patenting rates of urban (metro county as the base category) and the rural counties. Instead of reporting all these calculated interaction effects, we present them with the help of graphs.

We analyze the results in table 4 by calculating the marginal effects of the interacting variables, at their representative values in our sample, in rural and urban areas and graphing the difference in the marginal effects between the urban (base category) and the rural areas. We got these results by using Stata's *margins* and *marginsplot* commands.

**University R&D:** Figure 3 shows the difference in the marginal effects of university R&D expenditures and its spatial lag (university expenditures in counties neighboring within 100-mile radius) between urban (metro as comparison category) and the two types of rural counties. It shows that the rural-urban difference in patenting intensity increases with the increasing levels of the variables. For example, one percent increase in university R&D expenditures is associated with a range of 1.5-2.3% less increase in the patent applications per 10k population in the remote rural areas than in the urban areas. Similar interpretation follows for the spatially lagged university R&D variable and the comparison between the urban and the metro-adjacent rural areas. The difference in the marginal effects between urban and remote rural areas are statistically significant in the entire range of the values of the variables while the difference between the urban and metro-adjacent rural regions is significant only at the lower range of the values. As this result shows that university R&D expenditure is likely to yield to higher patenting activities in urban areas than in rural areas, the university decisions to locate in urban areas is likely to widen the rural-urban gap rate of innovation. Alternatively, as the university R&D is associated with positive patenting intensity in the main effect model, the gap is likely to reduce due to larger presence of universities in rural areas.

Our similar analysis of the results of the regression model with metro-adjacent rural counties as the base category show that there is not statistically significant differential marginal effect of university R&D and its spatial between metro-adjacent rural and remote rural regions.

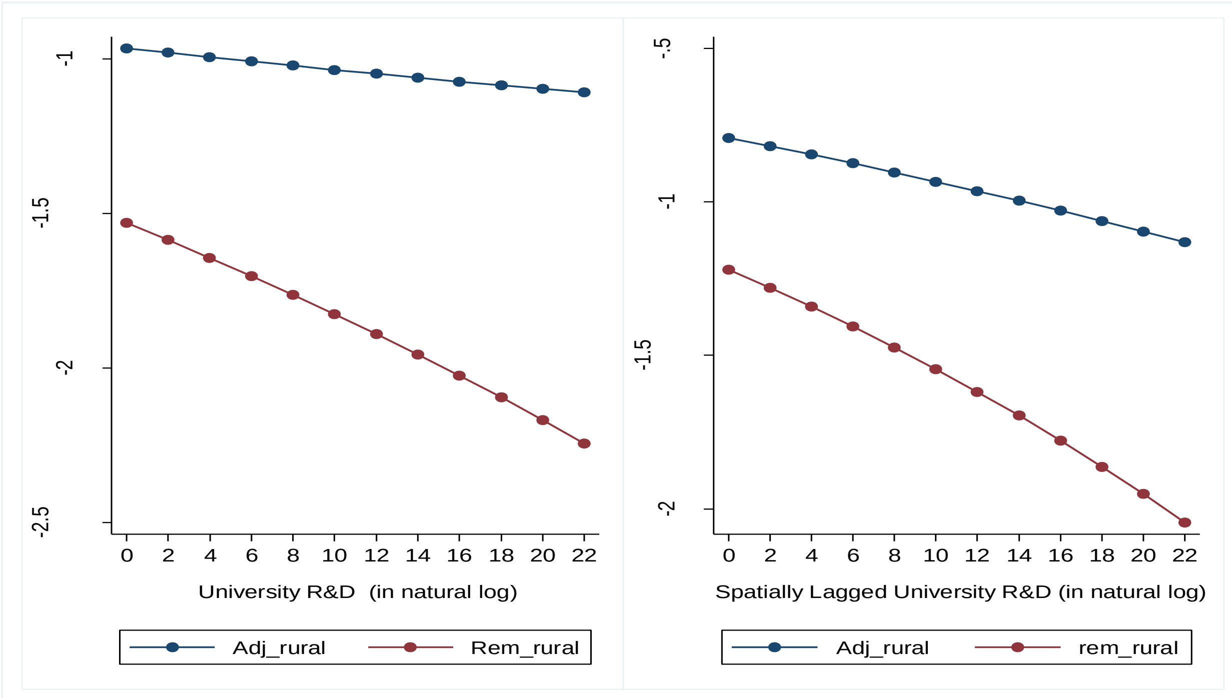


**Figure 3: Difference in Marginal Effects of University R&D on Rural-Urban Innovation**

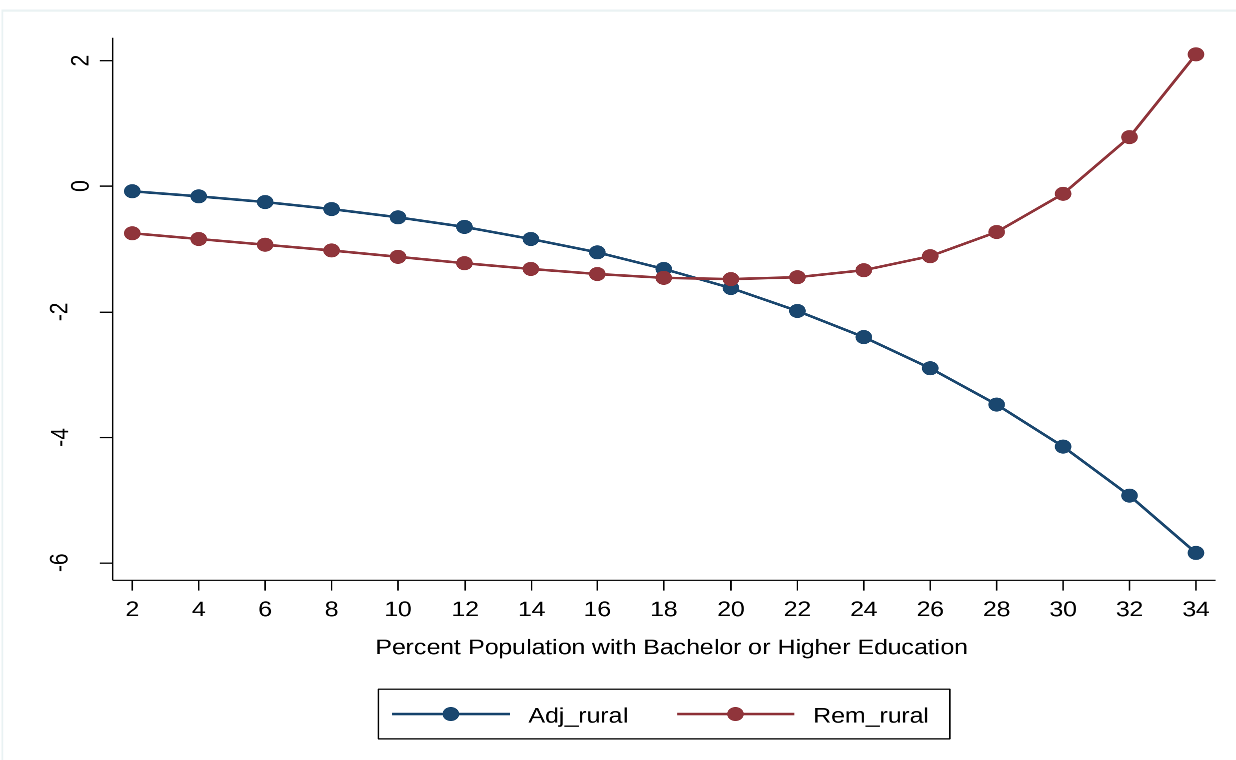


**Figure 4: Difference in Marginal Effects of Human Capital on Rural-Urban Innovation**

**Human Capital**: It appears from figure 4 that the difference in marginal effect of the stock of human capital between metro counties and the metro-adjacent rural counties continuously increases with the higher share of highly educated population but the difference between the urban and remote rural counties increases until the upper range of values of the variable and decreases at the upper range values. The differences between the urban and metro-adjacent rural counties are statistically significant at the middle and upper range of values while the differences between the urban and remote rural counties are statistically significant at the range of values at the first half. In summary, within the range of values that are associated with the statistical significance, the increased share of human capital is likely to be more effective in generating patenting activities in urban areas. Thus, concentration of population with higher education in urban areas will potential widen the rural-urban innovation gap.

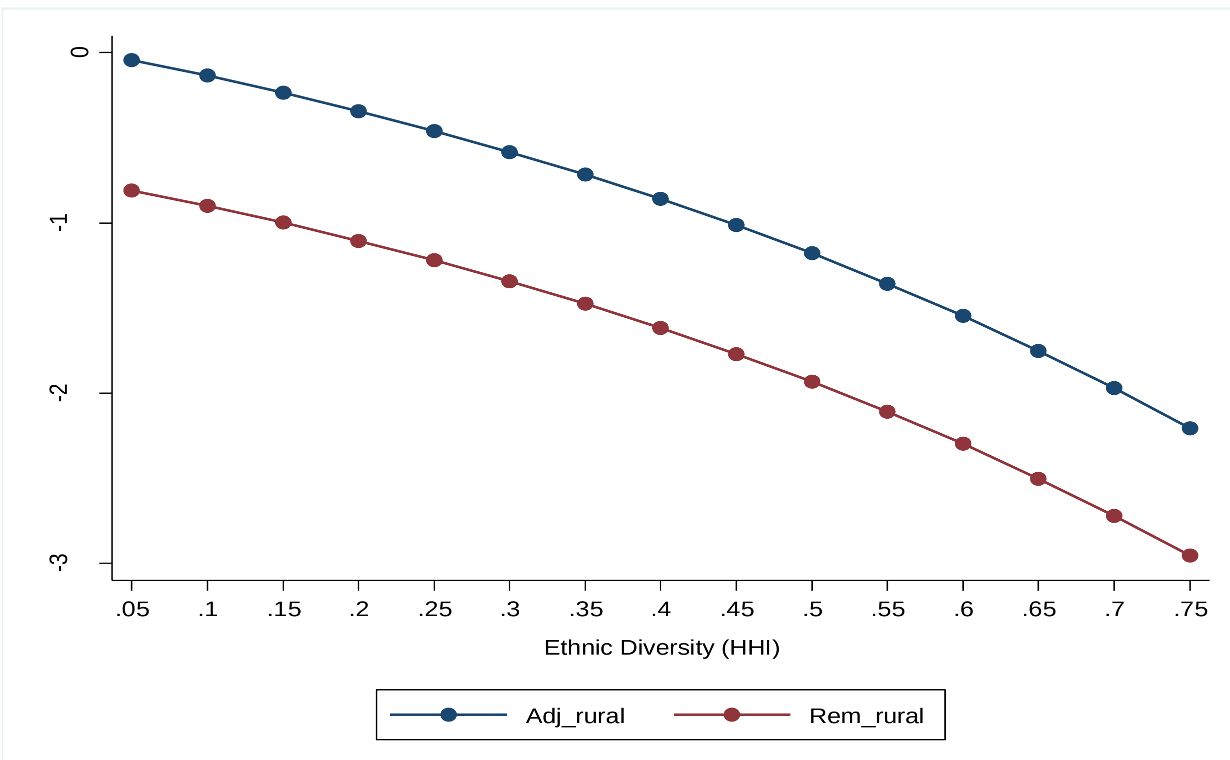


**Figure 5: Difference in Marginal Effects of Ethnic Diversity on Rural-Urban Innovation**

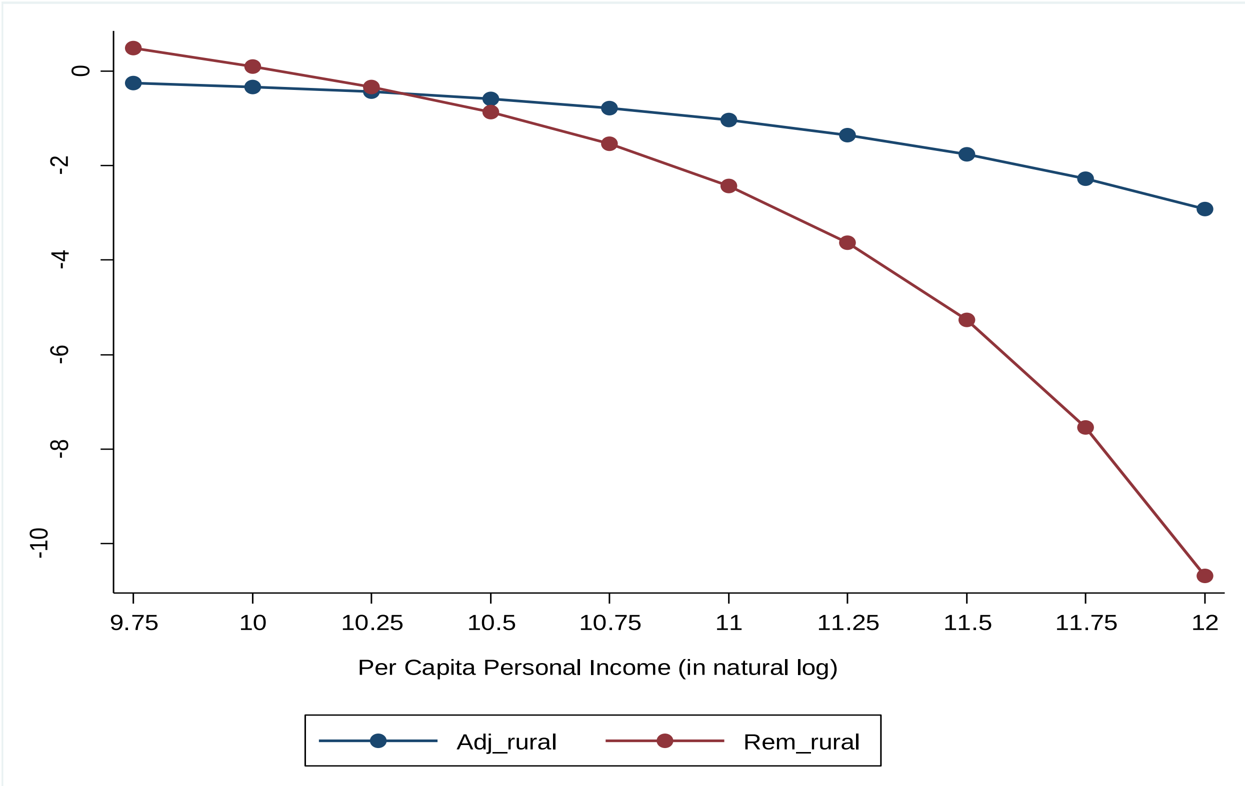


**Figure 6: Difference in Marginal Effects of PCPI on Rural-Urban Innovation**

**Ethnic Diversity:** Increasing level of ethnic diversity is associated with the widening rural-urban innovation gap (figure 5). The difference in the marginal effects of the ethnic diversity between urban and remote rural regions is statistically significant at the entire range of values of the variable but the difference between urban and metro-adjacent rural regions is significant only at the range of values at the upper end. Although these results suggest that the innovation activities increase faster in urban areas with higher ethnic diversity, the positive association of ethnic diversity with patenting in the main effect model suggest that the rural areas with higher ethnic diversity are likely to grow in terms of innovation activities.

**Personal Income Per Capita:** The increasing gap between the urban and metro-adjacent rural counties due to the marginal effects of per capital capita personal income (figure 6) is statistically significant only at within a small range (10.25-10.75) but the results for the difference between the urban and remote rural counties are significant at the values higher than 10.50. Thus, increase in income inequality between urban and remote rural regions is likely to further widen the innovation gap between these regions.

**High-tech Establishments**: Figure 7 shows that the rural-urban patenting gap increases with the incremental share of high-tech establishments and diversity of high-tech industries. According to our reasoning in section 3 that the variety of high-tech industries after controlling the share of high-tech establishments the total number of establishments represents the diversity of high-tech industries, we focus on the results for the high-tech variety.  The differences in marginal effects of the high-tech variety between rural (both types) and urban counties are statistically significant at the almost the entire range of values for the variable. This result imply that the diversity of high-tech industries is likely to produce higher marginal effects on patenting in urban areas than in rural regions. Alternatively, it suggests that specialization of high-tech industries is likely to reduce the rural-urban innovation gap.

**Agriculture Industry**: We do not find the statistically differential marginal effects of agricultural employment between urban and metro-adjacent rural counties (Figure 8). However, our finding on the difference between the urban and remote rural counties is significant at lower range of values (below 15%).  The reducing difference in the marginal effects between the urban and remote rural counties suggests that the latter regions, where agriculture contributes to smaller share of total employment, are likely to reduce their innovation gap with the former regions if there exist unutilized business opportunities in agriculture industry.

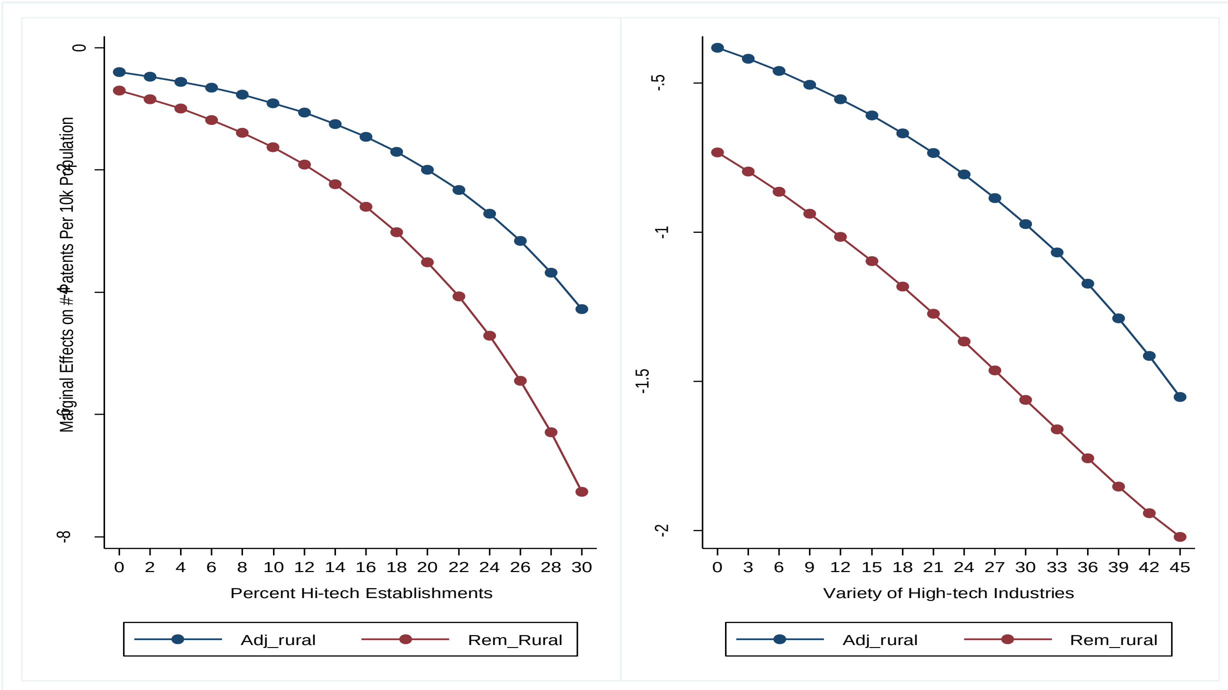


**Figure 7: Difference in Marginal Effects of High-tech Establishment Share and the Diversity of High-tech Industries on Rural-Urban Innovation**

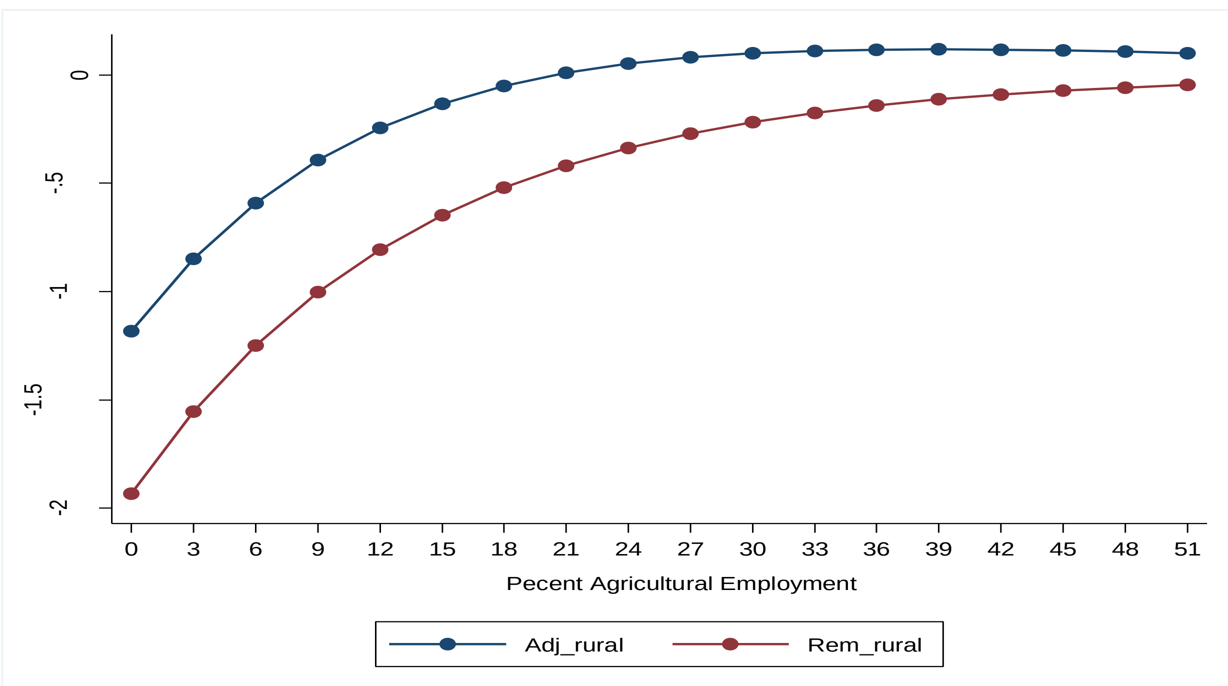


**Figure 8: Difference in Marginal Effects of Agricultural Employment Share on Rural-Urban Innovation**

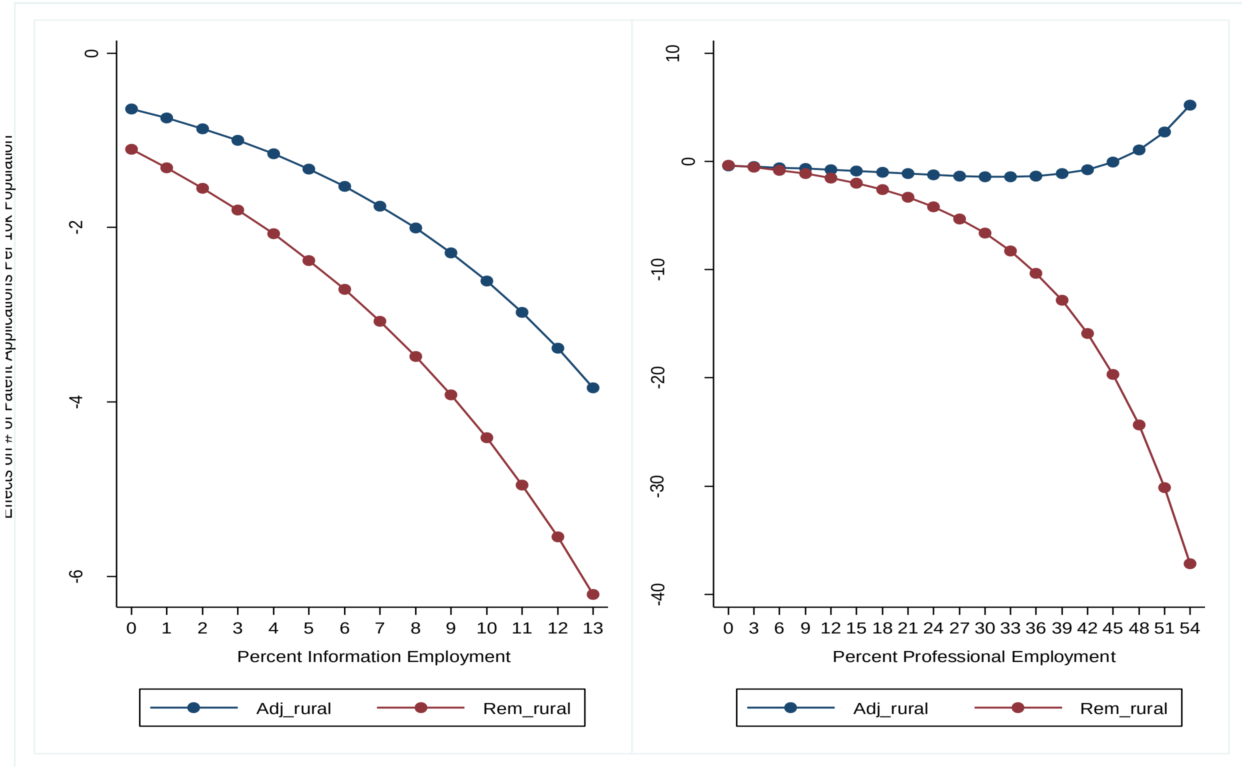


**Figure 9: Difference in Marginal Effects of Information and Professional Industry Employment Shares on Rural-Urban Innovation**

**Information and Professional Industries:** The results in figure 9 show that the incremental entrepreneurship opportunities provided by the larger presence of information and professional services, measured as the shares of employment in information and professional industries, are likely to generate more innovation activities in urban areas than in remote rural areas (the results for the metro-adjacent rural areas are not statistically significant). However, such opportunities are associated with increased innovation activities wherever they are available (table 3).

## 5. Summary and Conclusion

We examined the innovative activities in the US at its county level by using panel data for post-recessionary period between 2009 and 2013. Innovative activities in 2847 counties within 48 contiguous states excluding District of Columbia were measured by patenting intensity (utility patent applications per 10,000 population).

We use the concept of regional knowledge production function, that is based on the seminal works by Griliches (1979) and Jaffe (1989). Using this concept, the innovation inputs such as university R&D expenditures, SBIR innovation awards, and human capital measure and several county characteristics were regressed on the patenting intensity. The conditional fixed effects Poisson and negative binomial regression techniques developed by Hausman et al. (1984) we applied for our econometric analysis. we found that the patents data were over-dispersed. So, the Poisson regression analysis was not appropriate. Conditional fixed effects regression of the negative binomial model was also not appropriate as we would lose 41% of the data and the resulting analysis also would not entirely control the unobserved county level propensity to patent arising from stable county characteristics. Therefore, our analysis is based on the random effects negative binomial regression with an assumption that the propensity to patent among counties follows gamma distribution.

Our findings from the random effects negative binomial model with only the main effect terms show that there exists difference in patenting intensity between rural and urban counties (metro-adjacent rural counties patent 84% less frequently, at 10% significance level, and remote rural counties 75% less frequently, at 5% significance level, than the metro counties). We find the patenting intensity in average US county to be positively associated with university R&D, human capital, ethnic diversity, per capita personal income, average establishment size, diversity of high-tech industries, and shares of employments in information and professional industries but negatively associated with the share of employments in agricultural industries.

From the results of random effects negative binomial regression with interaction of the county level variables, which were associated with patenting intensity in the main effects model, with county categories, we analyzed the variables that are associated with the gap in innovation between urban and rural counties. We found that university R&D, human capital, ethnic diversity, per capita personal income, shares of employments in information and professional industries are more effective in generating patenting activities in urban areas than the rural (especially remote rural) areas. However, our findings of positive association of these variables with patenting intensity in an average county in the main effects model suggest that the rural-urban gap is subject to depend on the policy focus. A policy that intends to reduce the rural-urban innovation gap, hence the inequality in rural-urban economic growth, should consider directing resources to strengthen these variables in rural areas despite their larger effectiveness on urban innovation activities.   We also find that incremental specialization of high-tech industries and share of agricultural employment (in remote rural counties with unutilized competitive advantage in agricultural industry) reduce the rural-urban innovation gap

This study has limitations that come inherently from the nature of the available secondary data. We dropped the counties that had missing values for some key variables such as per capita personal income; almost a half of the counties in Virginia.  The missing values occur due to the failure in reporting by some counties.  If the non-reporting is systemically related to the rural areas, then our sample might be misrepresenting the population of rural counties.  We also do not have available at the county level the data on the R&D expenditures by firms and the more appropriate measure of entrepreneurship such as number of business start-ups, which the innovation literature finds to have significant role in innovation. Our data for the net formation of new business establishments and creation of new jobs does not account for the death rates of the businesses.

## Appendix

### A1: Aggregation of the Patents Data at County Level

The United States Patents and Trademark Office (USPTO)[6] provides separate reports on the publicly available patent applications and the correspondence address(s) of the inventor(s). These addresses do not provide the county of the inventors' residence but the zip codes. Some patent applications might not be publicly available due to such reasons as arising from the requests by the inventors for not disclosing their claimed invention until they are provided patent protection (for detail see Graham et al., 2015). we matched the patent application numbers until 2015 from the application data with those in the correspondence address data to derive the aggregate patent applications originating at zip code level. Aggregation of zip code level patents to county level was not obvious as some zip codes are associated with more than one county. we used zip-to-county crosswalk data from the HUD USPS[7] for matching zip codes to county(ies) published for the second quarter of 2011, a middle point of our study period. The crosswalk data contain 29, 854 zip codes uniquely associated with their counties. Remaining 9,242 zip codes have one-to-many associations with counties, ranging up to one-to-six.

The USPTO provides the county level information on the patents for the granted patents from 2000 to 2013[8]. we matched the application numbers of the granted patents with the application number of the patent applications available at the zip code level and identified unique relationship with counties for additional 4,268 zip codes. Matching the patent award numbers from the combined application and correspondence data at zip code level with those from

[6] https://www.uspto.gov/learning-and-resources/electronic-data-products/patent-examination-research-dataset-public-pair
[7] https://www.huduser.gov/portal/datasets/usps_crosswalk.html
[8] https://bulkdata.uspto.gov/data2/patent/ptmtdvd/

granted patents data at county level, we identified 116 additional unique zip-to-county relationships. Next, we assigned 520 zips that had one or less patent applications during 2009-13 randomly to one of the counties of their multiple association. Finally, we had 4,338 counties out of total 39,906 zip codes (roughly 11%) with multiple county associations.

we applied two strategies to derive our patents data – (we) dropping the 11% of the zip codes and using only those remaining with unique associations and (ii) distributing the number of patent applications originating from these counties equally to the multiple counties of their association. The first strategy would miss 19,502 utility patent applications (roughly 4.5% of the total utility patent applications in all the years during 2009-13). we applied the data sets obtained from both the strategies on our regression analysis. we did not find any change in the signs of the coefficients using the two datasets but in the statistical significance in some instances and a very little change in magnitudes of a few coefficients. Following USPTO practice of our strategy[9] (ii), we decided to use it for matching the zip-codes to counties and subsequently aggregating the zip-code level patenting to county level.

**A2: Derivation of the Spatial Lags of the University R&D Expenditure**

Let URD denote the current year university R&D in the US counties in 2009. Then our spatially lagged university R&D variable can be expressed as **W**URD, where W is the n×n spatial weight matrix and URD is the n×1 column vector. With ‘n’ as the total number of counties in our sample,

[9] https://www.uspto.gov/web/offices/ac/ido/oeip/taf/data/county.pdf

$$\mathbf{W} = \begin{bmatrix} 0 & w_{1,2} & w_{1,3} & \cdots & w_{1,n} \\ w_{2,1} & 0 & w_{2,3} & \cdots & w_{2,n} \\ w_{3,1} & w_{3,2} & 0 & \cdots & w_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{n,1} & w_{n,2} & w_{n,3} & \cdots & 0 \end{bmatrix} \text{ and } \quad \text{URD} = \begin{pmatrix} URD_1 \\ URD_2 \\ URD_3 \\ \vdots \\ URD_n \end{pmatrix}$$

The resulting n×1 vector of spatially lagged university R&D variable can be expressed as

$$\mathbf{W}\text{URD} = \begin{pmatrix} \sum_{j=1}^{n} w_{1,j} URD_j \\ \sum_{j=1}^{n} w_{2,j} URD_j \\ \sum_{j=1}^{n} w_{3,j} URD_j \\ \vdots \\ \sum_{j=1}^{n} w_{n,j} URD_j \end{pmatrix}$$

whose $j^{th}$ element represents the weighted average of the university R&D expenditures in the neighboring counties of county j. The weights were assigned based on the spatial weight matrix that we calculated by using the distance decay concept and limiting the university R&D spillovers from one county to another county to a geographical distance of one hundred miles (d=100 miles). Specifically,

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

we use a power function of distance decay with the parameter $\delta = 4$ to generate a spatial weight matrix and derive the spatially lagged university R&D expenditure variables.

**Table A1: Random Effects Negative Binomial Regression Results with Interaction Effects for County Level Innovation (Table 4 Continued…..)**

| Dep. Variable: Annual Domestic Patent Applications Per 10k Population | Coefficient Estimates | | |
|---|---|---|---|
| | Baseline | Metro-adjacent Rural | Remote Rural |
| *Time Dummies*[d] | | | |
| t10 | 0.012 | | |

| | |
|---|---|
| t11 | -0.019 |
| t12 | -0.049 |
| t13 | -0.046 |
| *State Dummies*[e] | |
| AL | -1.608*** |
| AR | -1.394*** |
| AZ | -0.059 |
| CO | -0.167 |
| CT | -0.165 |
| DE | -0.202 |
| FL | -0.488 |
| GA | -1.052*** |
| IA | -0.735** |
| ID | -0.353 |
| IL | -0.621** |
| IN | 0.486 |
| KS | -0.224 |
| KY | -0.996*** |
| LA | -1.516*** |
| MA | -0.515 |
| MD | -1.011** |
| ME | 0.272 |
| MI | 0.168 |
| MN | -0.466 |
| MO | -0.858*** |
| MS | -2.102*** |
| MT | -0.140 |
| NC | -0.496 |
| ND | -0.451 |
| NE | -1.609*** |
| NH | 0.627 |
| NJ | -0.621 |
| NM | -1.253*** |
| NV | 0.779 |
| NY | 0.249 |
| OH | -0.334 |
| OK | -0.681* |
| OR | 0.044 |
| PA | -0.641* |
| RI | -0.962 |
| SC | -1.026*** |

| | |
|---|---|
| SD | -0.987** |
| TN | -0.818** |
| TX | -0.773*** |
| UT | 0.882** |
| VA | -0.231 |
| VT | -0.984* |
| WA | 0.268 |
| WI | -0.254 |
| WV | -0.258 |
| WY | 0.090 |

[d]Base year is 2009; [e]Base state is California (CA); *** p<0.01; ** p<0.05, * p<0.1

Note: The 2nd and 3rd columns are blank because state and time dummies don't interact with county categories