Identification of Knowledge and Innovation Clusters: A GIS Application of Concentration, Co-existence, and Correlation

By Hamilton Galloway and Henry Robison
Economic Modeling Specialists, Inc.

Introduction
Popular themes in recent regional economic development literature certainly include the role of human capital (Becker, 1964; Barro, 1992; Mathur, 1999) and high technology-driven economic clustering (Porter, 1998; Audretsch & Feldman, 1996; Romer, 1986; Krugman, 1991). Neither of these concepts is really new. Human capital in the development context can be viewed as roughly synonymous with Weber’s location-determining notion of “labor orientation (Weber, 1929),” while economic clustering harkens back to Marshall’s “industrial zones (Marshall, 1890).” In Weber’s case, firms chose their location by weighing the relative advantages of transportation access (transportation orientation), natural resource availability (resource orientation), and workforce characteristics (labor orientation): policy makers might spur regional development by enhancing any or all of the above. In Marshall’s case, firms that share an output market and/or use similar inputs create a positive externality and thereby realize a competitive advantage by co-locating: policy makers might spur regional development by encouraging such co-location.

Lately the nexus of economic clustering, human capital and innovation has drawn particular attention (Porter 2003; Baptista and Swann, 1998; Thompson, 1965). This is partly due a changed character of national outputs, i.e., the production of services and particularly information services in place of traditional manufactured goods, and to a parallel change in industrial processes, i.e., substitution of computers and information technologies for traditional mechanical techniques (Alcaly, 2003). Several authors have speculated on the mechanism through which innovation leads to economic growth and development (e.g., Florida, 2002; Markusen and King, 2003; Caves, 2000; Scott, 2000). The existence of universities and other higher education can certainly play a role in innovation (Goldstein, 1995), as can proximity to cultural resources (Schoales, 2006; Markusen et al., 2004). The mix of innovation factors is likely large, and the interplay and synergy among these factors is just beginning to be understood. The policy implications are many.

This paper takes a spatially descriptive look at innovation and the US economic landscape. The objective is to identify measurable characteristics of innovative clusters. We develop a theoretical outline, using a collection of appropriate data in order to portray U.S. innovative regions. The policy implications of this work are obvious: regions can spur economic development by filling gaps in their innovation milieu or studying best practices by similar economies.

Background
Innovation in the Northwest and moreover the United States, takes on many forms. Some innovation is focused around production processes while others are focused around production development. The fundamental key in all cases of innovation, however, is the
ability to enhance or develop new technology and industry. In many cases, this innovation leads to business and industry clusters. Clusters are a group of industries that are closely linked by common product markets, labor pools, similar technologies, value chains, and/or other economic ties. Clusters can take on strategic importance because activities that benefit one group member will generally have positive spillover effects on other members of the cluster. The idea behind creating clusters is to develop a strategic cohesive business or industry framework at the regional level that leverages the strengths of related innovative industries and workforce in order to create a strong comparative advantage in the region for the particular cluster. This advantage allows a region to grow and develop around the cluster, while concurrently diversifying through promoting industries to support the emerging or existing cluster. One specific example of this type of cluster is Silicon Valley. High-tech industry innovation and development led to the initial cluster and further IT support and web-development augmented the high-tech industry cluster. From this concept and framework of industry clusters came the need to define types of clusters that can be used by regions for analysis purposes. Industry clusters such as Life Sciences (Biomedical/Biotechnical) and Transportation & Logistics have been identified and defined (based largely on industry description and NAICS codes) by entities such as Pennsylvania Department of Labor and Industry and Purdue University’s Center for Regional Development.

The second wave of cluster analysis revolves more around concepts than particular industries, though industry and occupational mix play a very real and very important role in identifying and developing these concept clusters. The particular concept to be analyzed through this methodology is knowledge (Knowledge Clusters). The foundations of this concept has origins dating back to Peter Drucker’s 1959 book, *Landmarks of Tomorrow: A Report on the New Post-modern World*, where he identified the strong potential for a paradigm shift from production workers to knowledge workers. He called this new world the “knowledge economy.” At the regional level, the expansion of knowledge occupations, as a regional development initiative, has resurfaced with more vigor in recent years. Knowledge occupations focus more on the human capital of a region, arguing that occupations—especially “high knowledge” occupations such as physicists, biochemists, engineers, etc.—are the drivers of innovation and technological development. When these types of knowledge occupations become increasingly interconnected, knowledge clusters begin to form, which when leveraged by industries increase regional industry diversification and competitiveness.

The idea of knowledge can be interpreted in many fashions. For the purposes of this research project, the definition of a knowledge cluster is borrowed in part from Japan’s 2nd Science and Technology Basic Plan in which emphasis was placed on promoting the

---

1 Pennsylvania’s Department of Labor and Industry. “Pennsylvania’s Targeted Industry Clusters: Executive Summary.”
2 Purdue University’s recently released Industry Cluster definitions are largely considered the most detailed and comprehensive list of cluster identification and definitions. Furthermore, when identifying “knowledge clusters” the Industry Cluster definitions from Purdue will be utilized.
creation of "Knowledge Clusters" in different regions of Japan. The Basic Plan defined knowledge clusters as:

“a technological innovation system formed through local initiative. At the center of the clusters is a unique R&D topic for the area, and core research institutions (like Universities) with high research potential. The system also involves the participation of corporations and other groups from both inside and outside the region.

More specifically, the project forms human networks and joint research organizations to promote beneficial feedback between the "seeds" of innovative technology possessed by public research organizations and other groups forming the core, and corporate needs for practicality. This creates a chain reaction of technical innovation, which eventually results in the creation of new industries [or expansion of current industries].”

This study takes chapters from both industry and occupation cluster schools of thoughts to identify and describe innovative regions across the Northwest and U.S. in general. The proposed methodology developed for this particular segment of the project focuses on identification and description of knowledge clusters through three sets of spatial analyses, notably: concentration, co-existence, and correlation. Certain benefits are achieved by identifying the clusters deemed to be considered knowledge based. First, regions where knowledge clusters exist will have more in-depth understanding of the resources available and will be in a better position to align economic development around their existing clusters. Second, through the process of identifying and describing knowledge clusters, certain opportunities necessary to foster a stronger knowledge environment can be discovered. These discoveries can help regional economies bridge gaps, where possible, and put them on the road towards future economic development and growth that revolves around knowledge and innovation. Third, once clusters are identified, further analysis focused around how these clusters are created, and the economic returns of developing one or more specific aspects of knowledge creation over other aspects of knowledge can be defined. This will allow regions with gaps or regions seeking to develop a knowledge corridor the framework for implementing the development.

**Variables and Methodology**

The methodology for identifying knowledge and innovation clusters is a multi-tiered analysis that utilizes concentration, co-existence, and correlation across counties to identify clusters that are knowledge and innovation focused. Proxies will be used to identify knowledge and innovation on a broad and specific level. Once identified, the relative concentration (location quotient) of these proxies will be measured. When displaying results, only those proxies with relatively high location quotients (>1.2 for industry clusters, occupation clusters and educational attainment and >1.0 for patents) will be displayed. This way, counties that have a 20% or greater share of employment for a given industry or occupation compared to the U.S. share of employment for the same industry or occupation will be identified. The advantage of this method is county size or population does not dictate the results. Instead, relative proportional shares become the focus. Once concentrations are identified, the co-existence of multiple proxies will allow further narrowing of the knowledge concentrations. Finally, a correlation analysis of the

---

3 The measurement for the proxies will utilize location quotient, a measure of relative regional concentration.
concentrations will define areas where inter-county level clustering has occurred. For the correlation analysis, a new county-level innovation index is created that factors in 10 high-tech industry clusters, 10 high knowledge occupation clusters, patents granted and educational attainment at the undergraduate and graduate levels. The primary data source used to collect industry, occupation, education and patent data was Economic Modeling Specialists, Inc (EMSI). EMSI uses a sophisticated data disaggregation technique to un-suppress industry and occupation data at the 6-digit NAICS and 5-digit SOC levels. After collection and sorting the data, ArcGIS was used to spatially display and interpret results.

In order to conduct the analysis, general proxies for knowledge clusters are were defined as:

- High tech industry and high-tech industry clusters (as defined by Purdue University)
- High-Knowledge occupation clusters (as determined through a clustering algorithm)
- College and Graduate Degree Attainment
- Patents granted
- The relative density of higher education institutions

**High tech industries** tend to focus on heavy research and development in order to stay at the cutting edge of their respective industries. However, past research and data has shown that high-tech industries do not necessarily equal “knowledge,” in the sense that this term is being used. Industries in many cases utilize the comparative advantages of local areas around the U.S. for functions such as sales/marketing, production, and development. As a result, an industry (with the same NAICS code) may have a production facility in one geographic area with a relatively low level of knowledge and innovation, and a R&D facility in another area with a high level of knowledge and innovation (Barbour and Markusen, 2006). Distinguishing between these levels can be very difficult; however identifying the relative concentrations of occupations within the specific industry will enable further insight into the level of knowledge and innovation. Additionally, in order to create interpretable results, knowledge and high-tech industry groups are used. These groups are compiled according to similar functionality and value-chains. The groups used for this study are those identified by Purdue University’s Center for Regional Development. Specifically, these groups include: advanced materials, biomedical/biotechnical, chemicals, computers and electronics, electrical equipment and appliance, energy, IT/telecom, machinery manufacturing, transportation equipment manufacturing, and business/finance.

**High-knowledge occupations** provide a level of understanding of the type of work that would foster knowledge and innovation. Occupations that require a high level of training or a significant amount of education, most likely will have a higher tendency of being considered a “knowledge” occupation. Unlike the industry cluster’s focus on value-chain

---

4 A list of the industries and NAICS codes can be made available upon request.
operations, knowledge occupation clusters are formed into groups of similar job functions and knowledge components such as bio-science, or engineering and mathematics.

For this project Ward’s hierarchical agglomerative clustering algorithm was used to cluster occupations according to 33 knowledge variables identified and described by O*NET. Ward’s algorithm uses error sum of squares to minimize within cluster variance; see equation 1.1.

\[ ESS = x_i^2 - \frac{1}{n (\sum x_i)^2} \]

Where \( x_i \) is the score of the \( i^{th} \) case.

The within cluster variation is considered minimized when the sum of squared deviations from the measured knowledge variables (i.e. points) to centroids cannot be reduced any further, given a subjective restriction in the number of clusters formulated.

Occupations with a Job Zone code of 1 and 2 (those occupations that require little to no education and training) were omitted from the algorithm. ONET occupations with Job Zone codes 3, 4 and 5 were used in the algorithm. Based on the results from the algorithm, selected high-knowledge occupation groups were used in the study. These groups are: architecture, biology, management, computer science, engineering, finance, health, medical, physics and statistics.

**College and Graduate Degree Attainment** ties in with occupation. The assumption made here is that significant levels of correlation exist between knowledge, innovation and educational attainment. This is even more so for areas with higher relative concentrations of graduate degree attainment. To this end, separate location quotient (LQ) measurements of the area population 25+ years old with a 4-year college degree and graduate degree attainment are used.

**Patent data** at the county level is used as an indicator of innovation. Furthermore, where there is the existence of industries that entail knowledge and innovative occupations, a measure of patents would further support the occurrence of knowledge clusters and additionally indicate the relative innovation output of those clusters.

The patent data is collected from the U.S. Patent Office and is collected at the zip-code level. EMSI converts the zip codes into FIPS codes (county codes) and a LQ is calculated according to a patent-county employment ratio vs. patent-U.S. employment ratio.

The occurrence of **higher education institutions** provides insight into the high potential for R&D that provides “seeds” to industry development spin-offs resulting from the

---

A list of the occupation groups and SOC codes can be found in Appendix B.
funded research. A spatially allocated data set highlights the occurrence of major higher education land-grant and research institutions.

As described earlier, the identification of knowledge clusters is a three-tiered process of concentration, co-existence and correlation. Each process serves to identify a particular set of occurrences and the latter two processes serve to further isolate those occurrences and provide more insight into the level of clustering.

Concentration is simply geographically identifying all areas of knowledge indicators, starting first with high-tech industry. Individual maps of relative densities for the defined indicators will allow an analyst to see the dispersion of the indicators across space. There is an expectation of strong spatial diversity between respective high-tech industries and high-knowledge occupations. Given the number of variables used in this analysis, only selected results will be reported. Regions of interest, especially rural counties with very high LQs, will be addressed in the text.

Co-existence requires layering of the individual concentration maps in order to find where the indicators intersect. This will provide insight into level of co-existence and which variables tend to be spatially associated with each other. Furthermore, this will provide an analysis for determining the “degree” of knowledge or innovation in a region. For example, a particular area in southeast Texas may show high concentrations of high educational attainment, high-tech industry, and high-tech jobs. However, the level of patents granted may be very low. This will indicate that there may be a knowledge cluster in southeast Texas, but it would furthermore indicate that the region has high potential resources to leverage and use to market for R&D industries in order to increase innovation (in addition to economic stability) and move towards a strong knowledge cluster.

Since specific high-tech industry clusters’ R&D or innovation capacity is characterized by a limited set of knowledge occupations, or occupational groups that would most likely describe the industry and the type of innovation. For example, advanced materials industry innovation is largely developed by engineering and physics occupations. In order to determine areas where there is a higher probability that this innovation is taking place, an analyst would seek areas with high concentrations of advanced materials industries, engineering and physics occupations, high educational attainment and patents.

Correlation requires higher levels of spatial statistics and analysis. Getis-Ord Gi* statistics will be used to identify relative high and low multi-county innovation clusters. The clusters are spatial clusters and are not to be confused with industry or occupation cluster. Results from the custom innovation index will be used to conduct the spatial autocorrelation statistics. The following equation explains the process of attaining the statistics and the level of interpretation that the values will be used.
Getis Ord $G$:
\[ G_i = \frac{\sum_j w_{ij} x_j}{\sum_j x_j} \text{ where } j \neq i \]

$G_i$ is the measure of local clustering of attribute $x$ around the $i^{th}$ location, $x_j$ is the value of $x$ at $j$, and $w_{ij}$ represents the strength of the spatial relationship between units $i$ and $j$ (measured in a distance decay function). The expected value and expected variance of $G_i$ are:

\[ E(G_i) = \frac{w_i}{n-1} \]
\[ \text{Where: } w_i = \sum_j w_{ij} \text{ } j \neq i \]

\[ \text{And} \]
\[ \text{Var}(G_i) = \frac{w_i(n-1-w_i)s_i^2}{(n-1)^2(n-2)} \]

Under the first law of geography (i.e. nearer things are more related than distant things) these measurements will identify the strength and significance (z-score) of the innovation index values and how related those measurements are across space.

The knowledge index described above is combines all high-tech industry and high-knowledge occupation group LQs, patent LQs, college degree attainment LQs and graduate degree attainment LQs. All high-knowledge occupation and high-tech industry variables were weighted according to their share of the national employment for the entire group; see equation 1.5.

\[ \alpha_i = \frac{N_i}{(N_i + \cdots + N_n)} \]

Where $\alpha_i$ is the proportional weight for the $i^{th}$ occupation or industry variable and $N_i$ is the national employment for the $i^{th}$ occupation or industry variable. The index is reflective of relative county level innovation and is reported as a LQ.

Results
Since this study utilizes several data variables and mapping to identify and describe innovation and knowledge regions, the number of figures and combination of variables generated would fill a book. Given this, the results provided in this section highlight interesting findings of high LQs for specific industry, occupation and patent concentrations. In addition, specific findings of industry-occupation co-existence are described. Lastly, the innovation index created using all the indicator variables will provide further insight into inter-county regional connectedness.

Though this study evaluated all U.S. county FIPS codes, particular focus is directed towards the Western United States and Northwest region, especially when evaluating co-existence. As was quickly discovered during the course of the analysis, a national focus,
especially a visual focus, becomes convoluted very quickly as more layers of data are added. However, regional focuses can be interpreted more easily and even shed light on specific regional innovation characteristics. One such example are the strong occurrences of information technology industries, engineering occupations, computer science occupations and patents in King County, Washington; Benton and Washington Counties in Oregon; and Ada County in Idaho. When these counties are evaluated more closely the types of businesses they most likely refer to are Microsoft (and supporting tech companies) (Redmond, WA), Micron and Hewlett Packard (Boise, ID), Hewlett Packard (Corvallis, OR), and various IT companies around Portland, OR. Figure 1 displays the results. A green cross-hatched county represents where all four variables co-exist. Blue and yellow counties indicate where IT industries co-exist with strong concentrations of engineer occupations or strong concentrations of computer science occupations. In other analyses the counties that appear in blue or yellow will be green when another industry is evaluated, for example, advanced materials.

Given the level of analysis, the overall results indicate that this approach was successful in identifying both regional knowledge strengths and innovative areas, especially rural areas. Furthermore, and not surprising, a high correlation between large research and land grant educational institution location and innovation also exists (see Figure 2).

Figure 1
Concentration
As expected, relative concentrations of industries and occupations varied widely depending upon regional characteristics and strengths. Of particular interest is the strong correlation between educational attainment (especially graduate degree attainment) and the location of large research universities (see figure 2). There are multiple explanations and scenarios that can describe such a correlation. Two specific explanations focus on the urban vs. rural characteristics of the identified counties. Under the urban scenario, one would assume an increased capacity and ability to draw and retain highly educated human capital. This is arguably the result of growing or existing large industry infrastructures in the urban areas and their demand for educated human capital inputs. Rural counties with higher education research institutions, however, are characterized by a different type infrastructure, specifically an education and knowledge creation focused infrastructure. Rural counties such as Whitman Co., Washington; Latah Co., ID; Gallatin Co., MT; and Albany Co., WY have a disproportionate amount of high education attainment that primarily consists of educators. These two descriptions provide one particular insight, namely, the sources of the highly educated populace in an area. Urban areas with can expect a large highly educated human capital source from industries whereas rural counties (with research education institutions) can largely expect their highly educated human capital source from the research institution in the county.

Figure 2
When addressing relatively high levels of industry concentrations, very specific regional focuses appear. A few examples include: information technology industries in the San Francisco Bay area (i.e. Silicon Valley), transportation equipment manufacturing industries in the Detroit Metro area and the Puget Sound area (i.e. Boeing), energy industries in Texas, Oklahoma, Wyoming and West Virginia and general machinery manufacturing located throughout the Mid-West. Though each of these examples is interesting in their own right, energy industry concentrations and large disproportional regional shares of U.S. energy development are evaluated.

Figure 3 displays the high regional concentrations of energy industries in the U.S. As stated, Texas, Oklahoma, Wyoming and West Virginia have high relative concentrations of energy, though each state’s energy industry focus is vastly different. West Virginia’s energy share is predominately coal, whereas Texas and Oklahoma’s energy focus is on oil and natural gas (Texas also has a large portion of the U.S. wind energy output and Oklahoma has hydro-electric production). Wyoming has significant coal and wind energy production and Nevada has significant natural gas production around Las Vegas as well as wind and solar productions spread throughout the state. For most practical purposes, this information is widely known. Additionally, policy insight may be less applicable at this stage of the analysis, though states that make effective use of renewable energy sources may serve as an example to other states with similar natural resources and characteristics. This may serve as a platform from which economic development can occur.

---

High relative concentrations of high-knowledge occupations in many cases produced regionally specific results. Regions such as the Washington D.C. metro area, Georgia, San Francisco Bay Area and the Greater Puget sound area show pockets of statistics related occupations. In many cases, these are most likely associated with Federal and State-level government operations (i.e. Departments of Commerce, Labor, Transportation, etc.). However, the occupation cluster that generate the most spatially unique results was biology. This cluster is comprised of natural science managers, animal scientists, biologists, biochemists/biophysicists, epidemiologists, park rangers and foresters. Figure 4 illustrates the strong Western United States and North Atlantic Coast concentrations for this cluster.

---

Figure 4 illustrates the strong Western United States and North Atlantic Coast concentrations for this cluster.

---

7 A complete list of all clusters can be produced by the author, upon request.
Explanations for this spatial occurrence tend to revolve around the level of environmental interaction that takes place in the West and on coastal areas. Washington, Oregon and Idaho, for example, are abundant in natural resources such as timber and wildlife such as fish. Furthermore, endangered wildlife in these regions further creates opportunities for these occupations to locate.

**Co-Existence**

The approach, as noted when discussing Figure 1, begins to highlight the possibility of linkages between industries, occupations and innovation. The second step from this analysis would undoubtedly be a county or region-level investigation to determine what (if any) linkages are occurring and possibly how they developed. This secondary analysis is predominately outside the scope of this project, though a small level of inter-connected analysis will be conducted for specific areas.

While the combination of specific variables seems all but endless, this section will focus on logical industry-occupation cluster combinations that have the highest potential to result in innovation. For example, engineering occupations in the transportation
equipment manufacturing sector have a high probability of developing new technology in cars, trucks, trains, boats and planes. As a result, these two variables are combined. Additionally, the occurrence of high patent concentration is also evaluated. Results from this example are depicted in Figure 5.

Figure 5

What becomes very apparent from analysis of this map is the large blue corridor of transportation manufacturing in East Michigan through Alabama. This is indicative of the automotive manufacturing plants and supply-chain corridor where GM, Ford, Chrysler, Honda, Subaru, Toyota, Nissan and various alliances exist. However, only a small portion of this region appears green (i.e. where both high concentrations of engineering occupations co-exist with transportation equipment manufacturing. These regions are focused around Detroit. Other areas of green indicate other various types of transportation equipment manufacturing and engineering occupations. For example in the Puget Sound area, Boeing and various suppliers most likely dominate this sector as well as in Jefferson County, Colorado where Lockheed Martin has an aerospace facility.

Combinations such as this yield similar findings that can be traced to specific companies in counties and regions. The usefulness of this type of analysis is multi-faceted. For starters, this highlights the exact issue of regionally focused industry functions that regional scientists such as Ann Markusen describe. In one region, new transportation
innovation is created. In another region, the innovation is put on the assembly lines and a
product is produced. When dealing with policy choices, knowing the both the strengths
and characteristics such as the ones described above in a given region will allow decision
makers to properly target industries and occupations for well aligned economic
development (Thompson & Thompson, 1985; Markusen, 2004; Koo, 2005).

Correlation
This last level of analysis focuses on spatial allocation of all measured variables for this
study. In order to produce interpretable results, all variables were combined, though the
integrity of each variable was at least partially maintained through the proportional
allocation method outlined in the methodology section. The goal for this index is to track
the progression of knowledge and innovation indicators over time to further describe
regional innovation dynamics. For now, though a static approach is used. The index
highlights the counties that have high concentrated levels of various knowledge and
innovation indicators. Though the measured values are susceptible to “outlier”
concentration measurements, this is partially overcome by the proportional weighting.
As a result, educational attainment (college and graduate degree) and patents have a
higher tendency to drive the index results. High patent concentrated counties, however,
didn’t necessarily result in high innovation index values, nor vice versa. Counties that
had high concentrations of patents, but were precluded from the high innovation index
counties did not have the necessary levels of high-tech industry, high-knowledge
occupations, nor educational attainment necessary to be considered innovative. Possible
explanations for this stem from the ability for anyone to obtain a patent, so long as they
invent or discover something new and useful. Therefore, even areas without the
necessary innovation infrastructure can have high relative levels of patent generation. In
the opposite case (i.e. innovative areas without patents), this may pose an opportunity for
area decision makers to utilize their industry and human capital resources to foster an
innovation economy. Figure 6 displays the relative concentrations of innovation. The
relative concentrations of patents are super-imposed on the map to show the distinctions.
Some obvious areas of possible innovation cultivation exist in portions of the Denver
Metro Region, southern Idaho, New England and central New Mexico.
Figure 6 shows yet another type of analysis that can be conducted to broaden the understandings of regional innovation, however, the analysis does not address inter-county correlations, more specifically spatial autocorrelation of innovation. The final analysis provides an all-encompassing overview of innovation, in particular the shared inter-regional innovation characteristics.

To achieve this level of spatial interpretation, Waldo Tobler’s first law of geography (i.e. everything is related to everything else, but near things are more related than distant things) is captured through use of a spatial autocorrelation model (Tobler, 1970). A Getis-Ord “hot spot” analysis is used to identify clusters of points with values higher in magnitude than one would expect to find by random chance. The analysis produces a z score for each county that represents the statistical significance of clustering for each county. Since this analysis measures spatial similarity a fixed band threshold distance of one decimal degree (roughly 70 miles) was used. This means any innovation observations more than 70 miles away from any given county would not be included in the spatial calculation. Further more, an inverse distance decay function was included so that nearer counties would be given more weight when measuring regional innovation.

The results from this auto correlation analysis show that the statistical significant level (>95% significance) of inter-county spatial clustering is predominately focused around
metropolitan regions, most notably: Seattle, Portland, Denver, Minneapolis-St. Paul, the North Atlantic Seaboard, Denver, Chicago, Houston, Austin, San Francisco and L.A. Areas with the highest levels of non-innovation are predominately focused in the south, most notably: Georgia, Arkansas, Missouri and significant portions of Kentucky, Tennessee, and Louisiana (See Figure 7). Green color coding shows areas with statistically significant innovation clustering. Areas with yellow show signs of positive innovation spatial clustering, but did not meet the 95% statistical significance criteria. Orange areas show slight clustering of low innovation index values, though these z values are not statistically significant. Areas in red show statistically significant low innovation index value clustering. Also included in the figure are the locations of major research universities around the U.S. In many cases areas with the highest positive levels of spatial autocorrelation have one or more major research university in the region.

Figure 7

Analysis of one particular region may shed some light on what aspects and variables are driving innovation clustering. Table 1 contains all the LQ values, innovation index values and Getis-Ord z values for selected counties around the Portland region. These counties include: Washington, Clackamas, Marion, Polk, Benton and Linn. Obvious areas of innovation strength in this region, which largely contributed to the clustering results, are educational attainment, advanced material manufacturing, computer science
industries, biology occupations, engineering occupations, statistics occupations, computer science occupations and patents granted. From a policy perspective, the values in this table represent both regional strengths to be leveraged for future economic development and weaknesses to improve upon to increase the region’s economic competitiveness. For example, Linn County (home to Oregon’s State Capitol) has strengths in advanced material manufacturing, machine manufacturing, engineering occupations and even patent generation, but is significantly lacking in educational attainment, information technology industries, computer science occupations and statistic occupations, which tend to be regional strengths. Possible policy options would include increasing the availability of educational programs that focus on engineering, computer science and statistics to increase the educational output in the county and begin developing the occupations that appear to be well leveraged in the region. Further recruitment of new businesses in the above outlined industries would further create occupational draw to the county and help retain the workforce, especially those completing higher education degrees.

Table 1

<table>
<thead>
<tr>
<th>Variables</th>
<th>Washington</th>
<th>Clackamas</th>
<th>Marion</th>
<th>Polk</th>
<th>Benton</th>
<th>Linn</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graduate Degree LQ</td>
<td>1.17</td>
<td>1.05</td>
<td>0.74</td>
<td>1.01</td>
<td>2.25</td>
<td>0.48</td>
</tr>
<tr>
<td>College Degree LQ</td>
<td>1.60</td>
<td>1.37</td>
<td>1.02</td>
<td>1.28</td>
<td>2.09</td>
<td>0.81</td>
</tr>
<tr>
<td><strong>Industry</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advanced Materials LQ</td>
<td>3.71</td>
<td>1.20</td>
<td>0.43</td>
<td>0.91</td>
<td>2.20</td>
<td>1.76</td>
</tr>
<tr>
<td>Biological Sciences LQ</td>
<td>0.81</td>
<td>0.80</td>
<td>0.82</td>
<td>0.86</td>
<td>1.01</td>
<td>0.76</td>
</tr>
<tr>
<td>Computer Science LQ</td>
<td>1.27</td>
<td>2.19</td>
<td>0.53</td>
<td>1.65</td>
<td>7.77</td>
<td>0.56</td>
</tr>
<tr>
<td>Electronic Manf. LQ</td>
<td>1.11</td>
<td>0.78</td>
<td>0.41</td>
<td>0.00</td>
<td>0.00</td>
<td>0.23</td>
</tr>
<tr>
<td>Information Tech LQ</td>
<td>3.52</td>
<td>1.11</td>
<td>0.46</td>
<td>0.54</td>
<td>2.29</td>
<td>0.50</td>
</tr>
<tr>
<td>Machine Manf. LQ</td>
<td>1.89</td>
<td>0.86</td>
<td>0.58</td>
<td>0.69</td>
<td>1.19</td>
<td>1.35</td>
</tr>
<tr>
<td>Transportation Equip. LQ</td>
<td>0.19</td>
<td>0.57</td>
<td>0.14</td>
<td>1.73</td>
<td>0.03</td>
<td>0.21</td>
</tr>
<tr>
<td>Business LQ</td>
<td>0.93</td>
<td>1.06</td>
<td>0.59</td>
<td>0.45</td>
<td>0.82</td>
<td>0.39</td>
</tr>
<tr>
<td>Energy LQ</td>
<td>1.97</td>
<td>0.85</td>
<td>0.66</td>
<td>0.54</td>
<td>1.71</td>
<td>0.76</td>
</tr>
<tr>
<td>Chemical Manf. LQ</td>
<td>0.91</td>
<td>0.55</td>
<td>0.54</td>
<td>0.97</td>
<td>0.51</td>
<td>0.99</td>
</tr>
<tr>
<td><strong>Occupations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Architecture LQ</td>
<td>1.14</td>
<td>1.45</td>
<td>0.49</td>
<td>0.59</td>
<td>1.74</td>
<td>0.69</td>
</tr>
<tr>
<td>Computer Science LQ</td>
<td>1.68</td>
<td>1.16</td>
<td>0.42</td>
<td>0.39</td>
<td>2.22</td>
<td>0.40</td>
</tr>
<tr>
<td>Financial LQ</td>
<td>0.94</td>
<td>1.03</td>
<td>0.75</td>
<td>0.78</td>
<td>0.79</td>
<td>0.57</td>
</tr>
<tr>
<td>Management LQ</td>
<td>0.81</td>
<td>1.05</td>
<td>0.88</td>
<td>0.87</td>
<td>0.94</td>
<td>0.64</td>
</tr>
<tr>
<td>Physics LQ</td>
<td>0.29</td>
<td>0.14</td>
<td>0.17</td>
<td>0.10</td>
<td>1.09</td>
<td>0.40</td>
</tr>
<tr>
<td>Biology LQ</td>
<td>1.07</td>
<td>1.20</td>
<td>1.34</td>
<td>1.52</td>
<td>3.86</td>
<td>0.88</td>
</tr>
<tr>
<td>Engineering LQ</td>
<td>2.67</td>
<td>1.06</td>
<td>0.60</td>
<td>0.65</td>
<td>1.76</td>
<td>1.09</td>
</tr>
<tr>
<td>Health LQ</td>
<td>0.94</td>
<td>1.05</td>
<td>1.07</td>
<td>0.87</td>
<td>1.16</td>
<td>0.83</td>
</tr>
<tr>
<td>Medical LQ</td>
<td>1.07</td>
<td>1.03</td>
<td>0.91</td>
<td>0.57</td>
<td>1.12</td>
<td>0.76</td>
</tr>
<tr>
<td>Statistics LQ</td>
<td>1.01</td>
<td>0.75</td>
<td>2.17</td>
<td>1.92</td>
<td>1.89</td>
<td>0.42</td>
</tr>
<tr>
<td><strong>Innovation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patent LQ</td>
<td>3.92</td>
<td>7.30</td>
<td>0.35</td>
<td>0.60</td>
<td>9.53</td>
<td>1.35</td>
</tr>
<tr>
<td>Innovation Index</td>
<td>1.90</td>
<td>2.35</td>
<td>0.70</td>
<td>0.86</td>
<td>3.31</td>
<td>0.80</td>
</tr>
<tr>
<td>Getis-Ord Gi* z-value</td>
<td>3.82</td>
<td>5.74</td>
<td>6.41</td>
<td>4.46</td>
<td>5.20</td>
<td>4.95</td>
</tr>
</tbody>
</table>
Conclusion

This research endeavor sought to identify an alternative way at viewing county and regional economies, namely through education, occupation and industry characteristics and combinations thereof. Through using this approach, practitioners and decision makers can get a better scope of their regional strengths and/or weaknesses and compare their economies to areas of similar size and endowments. This approach couples visual analysis and data analysis for a comprehensive understanding of small economic areas up through very large economic regions. Overall, this research and analysis proved to be very useful in identifying innovative and knowledge regions as well as the characteristics, whether specific industries or occupations, driving innovation.

Next steps from this analysis will seek to provide a dynamic components, through which occupations, industries and educational attainment will be tracked to determine areas growing in knowledge and innovation and those areas decreasing in their knowledge and innovation output. This will allow practitioners to identify “rising stars” in the economic development playing field and possibly gain more intimate knowledge of the factors driving an area’s development. A tertiary analysis involves developing a new location quotient that measures the industry’s share of occupational employment in a region, for example: the number of engineer occupations in the bioscience industry cluster. This analysis will further allow practitioners to specifically highlight industry-occupational strengths within the region that can be leveraged for economic development. A final analysis would entail developing a system that measures county-level GDP and/or value-added and conducting a series of regression analyses, specifically geographically weighted regression (GWR) modeling to identify which components of knowledge and/or innovation strongly contribute to economic growth and development. This level of analysis would be able to identify and prioritize economic development investments and highlight the potential returns on investments given regionally specific characteristics. In other words, each county/region would have their own specifically catered regional investment strategy, highlighting the best possible economic development options.

As regional science and studies of economic development continues to grow and improve, there will continuously be a need to develop new ways to improve research and understanding. This research adds a positive contribution to the growing literature of economic development.

References


