Geographic Distribution of Renewable Energy Sector Industries: An Analysis Using Recent Developments in Industry Concentration Measurement

D. Lane Register (1), D.M. Lambert (2,*), B.C. English (3), K.L. Jensen (3), R.J. Menard (4), M.D. Wilcox (4)

Paper prepared for presentation at the
Agricultural & Applied Economics Association's 2012
Annual Meeting, Seattle, Washington, August 12-14, 2012

Department of Agricultural & Resource Economics, The University of Tennessee, 302 Morgan Hall, 2621 Morgan Circle, Knoxville, TN 37996-4518. (1) Graduate Research Assistant; (2) Associate Professor, corresponding author (dmlambert@utk.edu); (3) Professor; (4) Research Associate, (5) Assistant Professor.

Abstract

Recent developments in firm location analysis are applied to explore the concentration patterns of firms making up the green energy sectors in 2002 and 2006. A two-step procedure is applied in this analysis. First, Guimarães, Figueiredo, and Woodward's spatial adaption of Ellison and Glaeser's industry concentration index are applied to estimate the degree to which firms making up the so-called green energy sectors tend to exhibit concentration. In the second stage, the spatial distribution of concentration is analyzed using a statistical framework, also suggested by Guimarães, Figueiredo, and Woodward. Preliminary results suggest that green energy subsectors exhibit significant global concentration, but localized concentration appears to be random.

Key words: global, local, industry concentration measures, green energy sectors

Introduction

The development of renewable energy sectors is expected to become an important economic driver of local and regional economies. Renewable energy production targets have been declared by federal and state agencies to encourage the expansion of so-called "green energy" production. For example, the 2005 Energy Policy Act required that 7.5 billion gallons of fuel come from renewable sources. In 2007, the Energy Independence and Security Act (EISA) mandated that 36 billion gallons per year of biofuels be produced by 2022, with 58% derived from renewable fuels (other than grain ethanol) with lifecycle greenhouse gas emissions meeting a 50% reduction over baseline emissions. Other efforts complementing EISA objectives include solar, wind, and methane production, for which recent federal funds also support.

To meet these goals, local planners and policy makers will require information about which attributes communities can leverage to attract, retain, and expand the businesses comprising specific renewable energy sectors. From the perspective of investors, the decision to select a location will be partly influenced by the concentration of business support networks and the potential for costs savings arising from network externalities emerging from up- and downstream linkages. Information about which regions exhibit comparative advantage with respect to the natural resources bases, skilled labor, and business support services will be important information for potential investors as they seek least-cost sites. Other important decision making criteria includes information about feedstock production, materials extraction, transportation, and transformation, and the production and distribution of energy products. Local comparative advantage will also be determined by skilled work force availability, access to capital, and infrastructure. The effects of spatial competition for limited energy production resources on job, income, and business establishment growth will vary, depending on the competiveness of a county and its ability to leverage local resources and connections to wider regional economies.

This preliminary research applies recent developments in the measurement and analysis of industry concentration and firm location theory, applying global and local geographic concentration indices to determine if the firms making up a particular renewable energy sector are clustered across space. The renewable energy sectors analyzed in this research are involved in the extraction, production, and distribution of renewable energy intermediate and final products (e.g., electricity; ethanol; drop-in fuels including biobutanol; or biodiesel materials), and the financing of businesses supporting these activities. The sectors considered are biodiesel, coal co-firing, wood direct fire, ethanol production from switchgrass, wood biobutanol and ethanol, landfill gas, dairy methane, solar energy, and wind power, with each value chain comprised of between 12 and 25 firm types (Jensen et al., 2010). The analysis is at the county level (3,078) for 2002 and 2010. Detailed sector information on business establishments and employment at the NAICS 6-digit level combines IMPLAN 2002 and 2010 National County Level Datasets and the US Census Bureau's County Business Patterns data for the same years.

The analysis proceeds in two stages. First, a global concentration measure is estimated following recent advances in industry concentration theory (Guimarães, Figueiredo, and Woodward, 2011; hereafter GFW). GFW's measure extends conventional global industry concentration indices (see GFW, 2007) to account for the location of areal units in space. This application extends GFW (2011) by developing a nonparametric bootstrap procedure to estimate confidence intervals for the global industry concentration measures. The null hypothesis is that firms belonging to a specific renewable energy sector are randomly distributed across spatial units. Rejection of the null hypothesis suggests that industry concentration is significant, warranting further inspection of the geographic distribution of firms comprising the sector. Second, given significant global concentration of a sector, a more detailed profile of regional sector concentration is developed using local concentration indices in the second stage. New developments in the derivation of the traditional location quotient (LQ) as a maximum likelihood estimator (GFW, 2009) are also applied in this stage.

Data

Global and local concentration indexes are estimated using employment and business establishment data. Business establishment data is from the US Census Bureau's County Business Patterns (CBP) database. This data has the advantage of being arranged by county divisions and by six-digit North American Industry Classification System (NAICS) codes. The analysis examines green energy industry concentration for all counties in the contiguous United States (J = 3,078)[1]. However, employment data with similar detail is difficult to obtain. Due to non-disclosure rules, CBP data lacks sufficient detail to facilitate disaggregated employment pattern analyses. Researchers are often compelled to search elsewhere for employment data or conduct analysis at highly aggregated levels. Methods exist whereby missing employment data can be imputed. For example, WholeData.net, the Minnesota IMPLAN Group[2], and other companies use proprietary algorithms to estimate missing employment information, offering enhanced datasets to the public (though generally at significant cost). This study uses employment data for 2002 compiled by WholeData.net. Employment data for 2006 are from the

[1] In 2002 and 2006, the contiguous United States consisted of 3107 counties and county-equivalents. This analysis excludes Washington, D.C., and combines most independent cities in the state of Virginia with nearby counties.
[2] Note that IMPLAN datasets are arranged by an alternative sector specification to the six-digit NAICS codes. However, a key exists by which the sector codes can be matched to facilitate cross-database analysis.

Minnesota IMPLAN Group database. Both data sets allow for highly disaggregated analysis of employment patterns at the county level.

Employment data from WholeData.net are enhanced versions of CBP datasets, thus representing employment captured at a single point in time. Furthermore, the CBP excludes self-employed individuals, agricultural production, railroad, and government employees, and employees of private households. Employment data from IMPLAN, on the other hand, are annual average estimates created from multiple datasets including the CBP, the Bureau of Labor Statistics' Quarterly Census of Employment and Wages (QCEW) datasets, and the Bureau of Economic Analysis' Regional Economic Information System (REIS) database. In spite of these differences, data from both sources generally aggregate to observed national and state employment levels. The analysis proceeds by acknowledging some incongruence between the 2002 WholeData.net and 2006 IMPLAN data sets, formulating inter-year conclusions based on these caveats. With respect to the global indexes used to determine industry concentration, analysis of changes should not be too problematic because the estimated indexes represent industry national averages.

Global Measures of Industry Concentration

Ellison and Glaeser (1997) proposed that the degree to which firms in industry *k* are concentrated is measured as

$$\hat{\gamma}_k = \frac{G_k - \left(1 - \sum_{j=1}^{J} x_j^2\right) H_k}{\left(1 - \sum_{j=1}^{J} x_j^2\right)(1 - H_k)}, \tag{4}$$

where the "raw" industry concentration index is $G_k = \sum_{j=1}^{J} \left(s_{jk} - x_j\right)^2$; $H_k = \sum_{i=1}^{c_k} z_{ik}^2$ is a plant size Herfindahl index; z_{ik} is plant *i*'s share of industry employment; c_k is the number of firms in industry k; and the variable of interest, $s_{jk} = e_{jk}/e_k$, is the location's share of industry employment.

It is important to emphasize that γ_k characterizes industry concentration arising from the combination of natural advantages and external economies, but because it is an employment-based measure, the parameter also accounts for economies of scale internal to individual firms. While the Herfindahl index accounts for variation in the distribution of firm size, γ_k primarily reflects external economies realized by larger plants and discounts those realized by smaller plants. Guimarães, Figueiredo, and Woodward (2007) observed this shortcoming of the EG employment-based index and developed an alternative index based on plant-counts (the GFW index). An advantage of this modified EG index is its ability to adjust for "lumpiness" or variation in establishment sizes that could arise when employment is used to measure establishment concentration. The GFW index, therefore, moderates the influence of scale economies internal to individual firms. Furthermore, the GFW index tends to have a lower variance compared to the EG index in simulation studies (GFW 2007). The reformulated EG index based on plant counts is

$$\hat{\gamma}_{ck} = \frac{c_k G_{ck} - \left(1 - \sum_{j=1}^{j} x_j^2\right)}{(c_k - 1)\left(1 - \sum_{j=1}^{j} x_j^2\right)}, \tag{5}$$

where $G_{ck} = \sum_{j=1}^{j}\left(n_{jk} - x_j\right)^2$; $n_{jk} = c_{jk}/c_k$ is region j's share of plants belonging to renewable energy industry k; c_{jk} is renewable energy industry k's establishments in region j; and c_k is the renewable energy industry's overall number of establishments in the nation (or entire analysis region). This concentration measure differs from γ_k by emphasizing firm counts and replacing the Herfindahl index, H_k, with $1/c_k$.[3] Naturally this substitution is based on the rather strong assumption that all plants in a renewable energy value chain contribute equally to concentration.

To illustrate the statistical implication of the concentration coefficients, when $\gamma_{ck} = 0$ ($\gamma_k = 0$), establishment (employment) concentration is purely random. On the other hand, when γ_{ck} = 1 ($\gamma_k = 1$), it is expected that all establishments (employment) in a renewable energy industry would locate in a single region. In their study, Ellison and Glaeser (1997) adopted the convention that concentration estimates less than 0.02 were "not very concentrated" while those greater than 0.05 were considered "highly concentrated".

Spatially-Adjusted Global Industry Concentration Measures

Despite these recent advances in firm concentration measurement theory, shortcomings remain with respect to capturing a complete picture of the effects of external economies on firm site selection. The first problem relates to the "modifiable areal unit problem" (MAUP), which stems from the imposition of arbitrarily defined spatial boundaries on to data which may or may not have been generated as a function of the selected boundaries (e.g. zip code, city, county, state, or MSA boundaries). The MAUP may bias estimates due to this aggregation problem (Openshaw and Taylor 1979). The second complication, the "checkerboard problem", results from the inability of previous industry concentration measures to account for the proximity of areal units (White 1983; Griffith 1983) [4].

GFW formulated a spatially-adjusted EG index, which accounts for spillovers, firm size effects, and natural advantages spanning a region. The spatially adjusted EG index establishes a link between the Moran's *I* statistic and the aspatial EG index by including a scaling factor

[3] The derivation of this index's relationship to the EG index is presented in Appendices A and B of GFW (2007).
[4] GFW (2011) (Section 2) present a detailed description of this problem by comparing various concentration measures among permutations of economic activity across a hypothetical region.

estimated using a neighborhood relational matrix, $\mathbf{\Psi} = \mathbf{I} + \mathbf{W}$, where $\mathbf{I}$ is an identity matrix and $\mathbf{W}$ is a matrix identifying neighboring spatial units. In the absence of neighborhood effects, $\mathbf{\Psi}$ is simply the identity matrix and the index collapses to the standard EG index. The estimator of γ_k^s (where the superscript s denotes the spatial reformulation of γ_k), is

$$\hat{\gamma}_k^s = \frac{G_{sk} - H_E(1 - \mathbf{x}'\mathbf{\Psi}\mathbf{x})}{(1 - H_E)(1 - \mathbf{x}'\mathbf{\Psi}\mathbf{x})} \tag{6}$$

where $G_k^s = (\mathbf{s} - \mathbf{x})'\mathbf{\Psi}(\mathbf{s} - \mathbf{x})$, $\mathbf{x}$ is a $J \times 1$ vector of shares of a reference employment distribution (previously, x_j); $\mathbf{s}$ is a $J \times 1$ vector of employment shares (previously, s_{jk}); and $H_E = \boldsymbol{z}'\boldsymbol{z}$, where $\boldsymbol{z}$ is a $J \times 1$ vector of average establishment size (previously, z_{ik}). Additionally, as GFW indicate (Appendix A; 2011), the spatially adjusted EG index can be extended to their firm-count index, γ_{ck} (Eq. 5).

The spatially adjusted firm-count index, $\hat{\gamma}_{ck}^s$, takes as its starting point the reparamertization of the spatially weighted EG index to a spatially weighted plant count index (justified on the same set of assumptions proposed by GFW (2007)). As with the conventional γ_{ck} index, H_k is replaced by $1/c_k$ and the raw concentration index, G_{sk} is modified as $G_{ck}^s = (\mathbf{n} - \mathbf{x})'\mathbf{\Psi}(\mathbf{n} - \mathbf{x})$, where $\mathbf{n}$ is a $J \times 1$ vector of industry establishment shares (previously, n_{jk}). The spatially weighted plant count index follows

$$\hat{\gamma}_{ck}^s = \frac{c_k G_{csk} - (1 - \mathbf{x}'\mathbf{\Psi}\mathbf{x})}{(c_k - 1)(1 - \mathbf{x}'\mathbf{\Psi}\mathbf{x})}. \tag{7}$$

Optimal Bandwidth Selection for Determining the Geographic Extent of Concentration

The weighting procedure suggested by GFW (2011) is followed here. All spatially adjusted global concentration indexes are reported and the relevant hypotheses evaluated at the optimal bandwidth distances, D_k^* and D_{ck}^*.

Local Measures of Industry Concentration

Significant global indexes of firm concentration warrant further analysis of firm location and employment distribution at a finer spatial resolution. Given significant global industry concentration, a more detailed analysis of these patterns proceeds using location quotients. The standard location quotient (LQ) compares the proportion of employment in a particular industry in a region with the proportion of employment in that industry across all regions (typically, the nation);

$$L_{jk} = s_{jk}/x_j, \tag{8}$$

where L_{jk} is the LQ for renewable energy industry k in location j, and s_{jk} and x_j (the reference distribution) are described above. It is generally assumed that when the LQ is greater than one, industry k is concentrated in location j.

GFW (2009) extended EG's (1997) dartboard model to formulate a more flexible derivation of the LQ. Under certain assumptions, their derivation is observationally equivalent to (8). They express the LQ as an estimator of the unobservable external economies and (or) natural advantages in location *j* according to EG's dartboard model. Similar to EG (1997), GFW (2009) assume that the spatial distribution of industry activity replicates the distribution of overall economic activity, using the restrictions in (2) and (3). Location event can be expressed as probabilities, given Eq. (1). The likelihood of observing a particular distribution of firms is then constructed as the product of all location probabilities weighted by a factor of w_{jk}, where $w_{jk} = s_{jk} \times \ c_k$, such that

$$l_k = \prod_{j=1}^{J} p_{j|\eta_k}^{w_{jk}} = \prod_{j=1}^{J} \left(\frac{x_j \exp{(\eta_{jk})}}{\sum_{j=1}^{J} x_j \exp{(\eta_{jk})}} \right)^{w_{jk}}. \tag{9}$$

Maximizing l_k with respect to η_{jk} and solving the first order conditions[5], it can be shown that $\hat{\eta}_{jk} = \log L_{jk}$, where L_{jk} is the LQ for industry *k* in region *j*,

$$L_{jk} = \left(\frac{w_{jk}}{w_k}\right) / x_j \; , \tag{10}$$

and w_k is the sum across regions of all w_{jk}'s in the region. As with the global concentration indexes, employment-based location quotients are unable to differentiate localization economies from internal economies of scale. Thus, Figureido et al. (2007) suggest an alternate weighting factor, $w_{jk} = c_{jk}$, which permits the formulation of an establishment-based location quotient derived from the same model.

The advantage of linking the standard LQ with the likelihood function of (9) provides a statistical framework for testing the null hypothesis: LQ > 1. GFW (2009) construct Wald statistics to test for the presence of localization economies influencing firm or employment concentration, basing their hypotheses on the η_{jk}'s. The first statistic tests if localization economies are present in a region. Following GFW (2009), a Wald test statistic is estimated as

$$W_{jk} = \frac{J[logL_{jk}]^2}{(J-2)w_{jk}^{-1} + \overline{\overline{w_k^{-1}}}} \, , \tag{11}$$

which is asymptotically distributed as a χ^2 variate with one degree of freedom. The null hypothesis is η_{jk}= 0; location-specific effects are not associated with concentration and thus, the industry is not localized in spatial unit j. Rejection of the null hypothesis suggests that the industry exhibits concentration resulting from location-specific advantages rather than random site selection which might exhibit concentration patterns.

Methods and Procedures

[5] GFW (2009) introduce a restriction on l_k requiring that $\sum_{j=1}^{J} x_j \exp{(\eta_{jk})} = x$ to resolve the resulting indeterminate solution, since the η_{jk}'s are unobservable.

A two step approach is used to determine: (1) *which* industries tend to globally concentrate across the nation, and of those industries, (2) *where* they are concentrated. In the first step, global concentration of a renewable energy industry is analyzed using the EG and GFW indexes. In step two, industries exhibiting concentration from step one are analyzed using the employment and establishment based location quotients.

Geographic concentration indexes are subject to a wide range of assumptions and different applications. For instance, EG's (1997) index was developed to measure the concentration of industry employment, whereas GFW's (2007) plant-count index was formulated to measure establishment concentration. Furthermore, these measures were developed to express concentration in relative terms. The type concentration measured also depends on the reference distribution chosen by the researcher. To illustrate, recall that EG (1997) applied their index to analyze manufacturing concentration. Geographic concentration of a particular manufacturing industry (defined by the county's share of industry employment, s_{jk}) was estimated relative to the county's share of all *manufacturing* employment, x_j (defined by the 2-digit SIC level). Though not an arbitrary choice, it would be equally admissible to measure the concentration of a particular manufacturing industry using a subset of manufacturing industries (say the 3-digit level NAICS), or all US industries as a reference distribution. It is clear that the selection of a reference distribution is by and large the choice of the researcher, keeping in mind that evidence of concentration is only meaningful relative to the reference distribution selected. Estimates indicate the relative strength of the uneven distributions of local factors attracting firms to certain locations. In other words, global indexes larger in magnitude may suggest greater proclivity of firms being attracted to factors found in relatively few locations (e.g. ethanol firms concentrating near agricultural areas specializing in grain production).

Step 1 – Measuring Global Industry Concentration

In the first step, employment concentration is estimated using Equations (4) and (6). Establishment concentration is estimated using Equations (5) and (7). Neighborhood relational matrices for (6) and (7) are constructed using county level, population weighted centroids based on the 2000 US Census Bureau definition. Therefore, neighborhood definitions are based on the distance between population centers rather than geographic centroids that lack economic or demographic context. The variances of the glopbal concentration indices are esimttaed using a bootstrap procedure (Greene, (2000) The null-hypothesis of $\gamma_{ck}^{s} = 0$ is rejected if the lower bound of the confidence interval is greater than zero.

Step 2 – Local Measures of Industry Concentration

Local measures are constructed by evaluating Equation (10) using the employment and establishment weighting schemes. Thus, the analysis of renewable energy value chains exhibiting global employment or business establishment concentration is extended in the second step. Relevant hypotheses about the strength of external economies are evaluated using equation (11). Specifically, counties where significant instances of localization are identified suggest the presence of external economies (and potentially cost savings) for a particular industry in that location ($\eta_{jk} > 0$). This frames the hypothesis about which counties exhibit comparative advantage with respect to attracting certain firm types of a renewable energy sector. The *p*-values of this test are mapped using ESRI's ArcMap software for the contiguous US to identify counties exhibiting statistically significant concentration.

Results and Discussion

Summary information about the number of businesses and employees in each renewable energy value chain appears in Table 1. It is important to emphasize that overlap in value chain roles among all industries is reflected in each sector's total. Therefore, the sum of all green energy *sectors'* employment and establishments is greater than the sum of all green energy employees and establishments.

Step 1 Results – Global Measures of Firm Concentration

Figure 1 (for 2002 and 2006, respectively) present the γ_k estimates for both global concentration indexes. All of the renewable energy sectors analyzed exhibited geographic concentration in 2002 and 2006. Considerable variation in establishment concentration compared to those of employment is apparent. The residential solar sector exhibited the greatest magnitude in terms of employment concentration in both 2002 and 2006. By comparison, the biodiesel value chain appears at the opposite end of the employment concentration spectrum.

Investigation into the nature of observed residential solar and biodiesel sector firm location patterns are discussed in detail to motivate the utility of these findings. Residential solar production is composed of two concentrated industries (Table 2). Substantially higher concentration levels are observed for the Semiconductors and Related Device Manufacturing industry ($\hat{\gamma}_k^s = 0.032$ in 2002 and 0.031 in 2006). This result is expected, given the oft-noted tendency for industries dependent on skilled labor to concentrate (EG 1997). Therefore, it appears that location decisions for firms engaged in the solar panel manufacturing step of the value chain would likely result in clustering around specialized labor sheds, similar firms, and (or) sunny locations. Panel installation firms in the Commercial Machinery Repair and

Maintenance industry are likely more footloose, encouraging site selection across a range of locations where populations demand residential solar panels.

Figure 2 plots the spatial and aspatial global concentration ratios across candidate bandwidths, suggesting that the influence of external economies/natural advantages begins to weaken beyond the distance at which the ratio reaches a maximum. In areas where firms concentrate, optimal bandwidths circumscribe the counties within 20 and 33 miles of neighboring population centers. Several implications arise from this finding. First, note that evidence of firm concentration is expected; if the transport costs of solar panels are relatively low, then location decisions based on, perhaps, access demand markets, may not be too important a source of cost savings. Instead, concentration of downstream suppliers may promote cost savings from increasing returns to scale to assemblers because of improved access to specialized labor. However, while bandwidth estimates may reflect access to labor sheds or potential interaction with other firms in the value chain, questions remain about the extent to which increasing returns from concentration may be offset by changing policy and market realities for solar energy. For example, diminished support for renewable energy subsidies and tax credits, increased foreign competition primarily from heavily subsidized Chinese firms, and expiration of US incentives at the end of 2011 (e.g. the Nonbusiness Energy Property Tax Credit) increase risk to investors and communities. While new tariffs have been imposed on cheaper Chinese solar panels (Gordon 2012), the future of a US role in this sector is unclear. Nevertheless, low transport costs and a demand for skilled labor suggest that location decisions of all firms engaged in the residential solar energy sector may be, on average, primarily driven by access to specialized labor pools.

The observed employment and firm concentration patterns of the residential solar value chain contrast the biodiesel value chain. Costly transport of biodiesel feedstock suggests that a relatively greater dispersion of firms may be expected. Extensive availability of restaurant wastes leads firms engaged in biodiesel production using yellow grease to locate in any number of locations across the US. Firms producing biodiesel using agricultural products (e.g. soybeans or canola) may also be relatively footloose, given the distribution of production potential for these crops. Lower global concentration indexes for biodiesel (compared to those of the residential solar sector) reinforce the expectations that this sector is more evenly distributed across the nation than firms supporting the residential solar value chain.

Figure 2 shows the change in the ratio of the spatial to aspatial global employment concentration estimates for the biodiesel sector in 2002. Biodiesel bandwidth estimates based on the neighborhood relational matrices ranged from 29 to 76 miles, suggesting a relatively large area of influence. The maximum bandwidth was 76 miles. Access to external economies appears to provide increasing returns to biodiesel firms across a wider distance than observed for the residential solar production sector. Note that the maximum bandwidth estimate, 76 miles (the distance at which the ratio of the spatial estimate to the aspatial estimate is greatest), was derived from employment concentration. This may reflect the demand for skilled labor juxtaposed with access to agricultural areas, perhaps implying spillover gains due to larger downstream firms across county borders. Biodiesel from agricultural production may, therefore, benefit from labor sheds that span several counties, while minimizing transport costs by locating in closer proximity to oilseed production areas. Biodiesel production near cities may be more likely to use yellow grease, and potentially, has greater access to skilled labor within influence areas. Concentration of firms in the biodiesel sector, therefore, may be primarily driven by labor availability with feedstock transported from a nearby agricultural periphery (biodiesel from oilseeds), or obtained from within the labor-rich region (biodiesel from yellow grease). This scenario appears to be in line with the disbursed nature of overall firm location activity in the biodiesel sector, perhaps characterized instead by *pockets* of concentration.

Step 2 Results – Local Measures of Firm Concentration

Local measures supplement the global concentration indices by identifying where renewable energy concentration is observed. Figures 3, 4, 5, and 6 identify local patterns (observed in 2002) in the establishment and employment location quotients for biodiesel and residential solar energy value chains. Note that establishment concentration patterns suggest more counties exhibit firm concentration in these two sectors than employment concentration. These patterns appear to be the result of comparing establishment distributions with an employment based reference distribution (x_j).

The patterns of local employment concentration of residential solar energy firms corroborate the implications of the global concentration indexes regarding firm site selection near locations exhibiting external economies (e.g. Silicon Valley or the Southwest US). The magnitude of employment concentration observed in the Appalachian Region is also remarkable – particularly in Virginia, West Virginia, and Southeastern Pennsylvania. When compared with the distribution of solar radiation (Figure 7; NREL 2008), it appears that concentration in the Appalachian Region is likely not because of sunlight. The prevalence of employment concentration across the Sunbelt is also notable. It remains unclear if firms that could participate in residential solar value chains may be more attracted to areas endowed with relatively more sunlight hours and radiation or if firms making up this sector perceive some advantage from locating near other factors (as in the case of Appalachian Region concentration). Inspection of the underlying local market structure reveals that Maricopa County, Arizona is home to number of solar panel manufacturers and related industries, including First Solar, another dominant global competitor. Arizona's Renewable Energy Tax Incentive program has also served as a factor in attracting solar companies to the region (e.g. Saint-Gobain Solar in 2011). Therefore, it may be assumed that Maricopa County exhibits comparative advantage with respect to attracting solar energy value chain firms, driven simultaneously by the availability of sunlight, a developing infrastructure of supporting businesses, and a history of favorable policy incentives to producers. In the case of the biodiesel sector, extensive employment concentration (Figure 6) supports the implication from the first step of the analysis that site selection of firms belonging to the biodiesel sector may be more geographically disbursed than firms oriented toward the residential solar sector. Additionally, note the general lack of establishment (Figure 3) or employment concentration across much of the Midwest. Employment and establishments both appear more likely to concentrate near densely populated areas (e.g. The Eastern Seaboard, Atlanta, Florida, Southeast Texas, Los Angeles, and Seattle, among others). As would be expected, county-level total biomass resources appear to be most abundant near the population centers of the East and West coasts, and across Midwest, likely due to the prevalence of agricultural resources distributed across the region. The lack of biodiesel value chain concentration in these areas (given the relatively high transport costs of soybeans and other oilseeds) suggests that many of the counties with a comparative advantage with respect to feedstock availability may lack other factors critical to the site selection decisions of biodiesel value chain firms. An employment location quotient of 1.22 (significant at the 1% level) indicates that Grundy County, Illinois may possess a comparative advantage with respect to attracting firms belonging to the biodiesel sector. REG Seneca, a local biodiesel refinery, therefore, may enjoy reduced feedstock and final product transportation costs due to access to an extensive local transportation hub (a shared resource arising associated with external scale economies). Costly transportation of biodiesel inputs and products suggests that counties with access to an efficient transportation network (e.g. Grundy County) may possess comparative advantage with respect to attracting firms engaged in the biodiesel value chain.

Conclusions

Renewable energy technologies have been the focus of many policy makers and investors for their role in promoting economic development and transitioning away from dependence on fossil fuels. Yet the patterns of industry concentration that could result in cost-savings to firms, and eventually to the flow of investments to counties have not been investigated. Data from 2002 to 2006 was analyzed to describe the geographic landscape related to ten renewable energy sectors. A two-step procedure was developed to analyze firm concentration using recent advances in industry concentration analysis. The utility of this approach was motivated by focusing on the biodiesel and residential solar energy sectors. It appears that site selection decisions of firms engaged in the solar energy value chain are made independent of access to solar resources, largely gravitating toward population centers, suggesting that access to factors associated with population centers (e.g. skilled labor or knowledge spillovers) may be of greater importance than proximity to natural solar resources.

The concentration patterns identified in this preliminary analysis suggest several implications for a range of stakeholders. If the objective is to determine where firms engaged in biodiesel value chains may experience relatively greater returns, policy makers and investors may observe that metropolitan areas generally possess the necessary structures to support various levels of value chain activity. On the other hand, localization economies appear to be relatively weak in rural counties that may produce a range of agricultural feedstock used in biodiesel and ethanol production or for electricity generation from direct firing of wood products. Investors are often forced to weigh the relative importance of access to raw materials versus access to localization economies.

References

Anselin, L. 1988. Spatial Econometrics: Methods and Models. Kluwer Academic Publishers, The Netherlands.

Cronon, W. 1991. Nature's Metropolis: Chicago and the Great West. W. W. Norton & Company, Inc., New York.

Ellison, G. and E. L. Glaeser. 1997. Geographic Concentration in US Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy* 105(5): 889-927.

Figueiredo, O., P. Guimarães, and D. Woodward. 2007. Localization Economies and Establishment Scale: A Dartboard Approach. FEP Working Papers 247.

Gordon, M. 2012. US Govt Sets New Tariffs on China Solar Panels. *Associated Press*, 20 March.

Greene, W.H. 2000. Econometric Analysis, 4th Edition. Prentice Hall. Upper Saddle River, NJ.

Guimarães, P., O. Figueiredo, and D. Woodward. 2011. Accounting for Neighboring Effects In Measures of Spatial Concentration. *Journal of Regional Science* 51(4):678-693.

Guimarães, P., O. Figueiredo, and D. Woodward. 2009. Dartboard Tests for the Location Quotient. *Regional Science and Urban Economics* 39:360-364.

Guimarães, P., O. Figueiredo, and D. Woodward. 2007. Measuring the Localization of Economic Activity: A Parametric Approach. *Journal of Regional Science* 47(4):753-774.

Guimarães, P., O. Figueiredo, and D. Woodward. 2006. Geographic Concentration and Establishment Scale: An Extension Using Panel Data. *Journal of Regional Science* 46(4): 733-746.

Guimarães, P., O. Figueiredo, and D. Woodward. 2004.Industrial Location Modeling: Extending the Random Utility Framework. *Journal of Regional Science* 44(1):1-20.

Jensen, K.L., D.M. Lambert, R.J. Menard, and B.C. English, and W. Xu. 2010. Projected Impacts of Green Jobs Development in the Appalachian Region. Bio-Based Energy Analysis Group, The University of Tennessee, Institute of Agriculture.

Lafourcade, M., and G. Mion. 2007. Concentration, Agglomeration, and the Size of Plants. *Regional Science and Urban Economics* 37: 46-68.

Marshall, A. 1890. Principles of Economics. Macmillan, London, England.

Minnesota IMPLAN Group. 2006. IMPLAN. Stillwater, Minnesota.

US Census Bureau. 2006. County Business Patterns (NAICS) Database. Washington, D.C.

US Department of Energy. 2010. *Renewable Energy Consumption and Electricity Preliminary Statistics 2009*. Washington DC: Energy Information Administration, August.

US Energy Information Administration. 2011. Average Retail Price of Electricity to Ultimate Customers by End-Use Sector, 1999 through 2010. http://www.eia.gov/electricity/ annual/pdf/table7.4.pdf.

Table 1 – Renewable energy sector employment and establishments - 2002 and 2006

Value Chain	Establishments		Employment	
	2002	2006	2002	2006
Biodiesel	2,061,911	2,366,005	12,944,391	17,275,533
Cofire	782,710	1,042,672	9,015,782	5,417,058
Wood Direct Fire	777,306	981,447	9,206,047	5,586,651
Ethanol - Switchgrass	1,122,755	1,028,278	10,552,367	6,674,636
Ethanol - Wood	1,904,160	2,225,963	20,513,715	24,810,396
Landfill Gas	763,581	878,564	9,714,170	3,702,717
Dairy Methane	757,450	947,920	7,558,072	4,648,078
Commercial Solar	146,060	336,135	5,142,390	6,888,514
Residential Solar	1,098	25,528	42,911	549,916
Wind Energy	1,045,790	1,396,584	13,778,125	10,234,347
Total Renewable Energy Sector	2,498,469	2,723,175	29,959,974	34,605,146

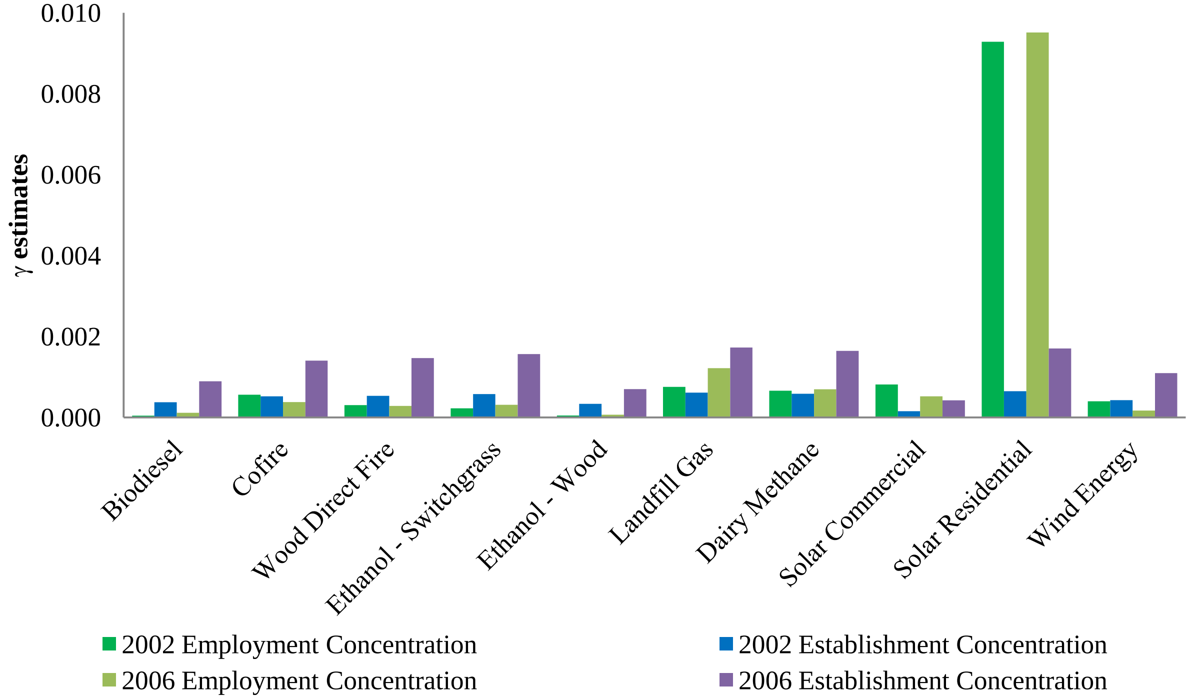
Source: 2002, 2006 US CBP, 2002 WholeData.Net, 2006 IMPLAN

Table 2 – Global industry concentration estimates for residential solar industries

Year	Industry Description	IMPLAN Code	NAICS Code	$\hat{\gamma}_{ck}^{s}$	$\hat{\gamma}_{sk}$
2006	Semiconductors & Related Device Manuf. ***	311	334413	0.03142	0.04667
	Commercial Machinery Repair & Maintenance ***	485	8113//	0.00150	0.00148
2002	Semiconductors & Related Device Manuf. ***	311	334413	0.03279	0.03869
	Commercial Machinery Repair & Maintenance ***	485	8113//	0.00121	0.00175

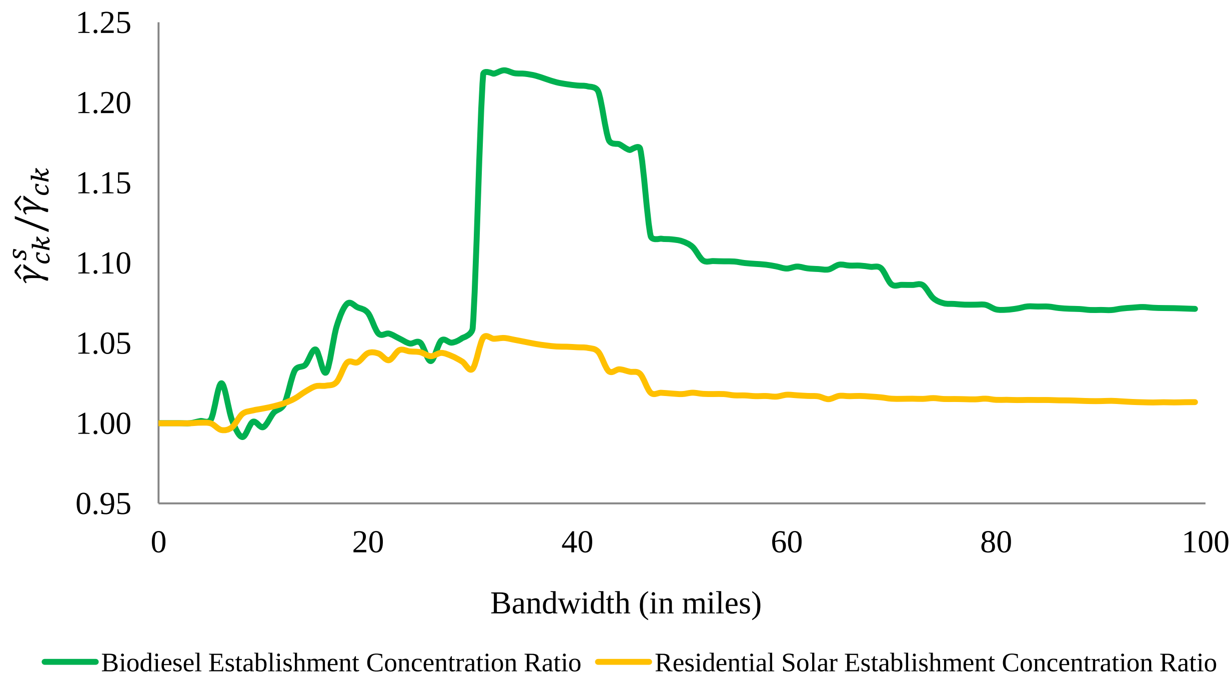
Source: 2002, 2006 US CBP, 2002 WholeData.Net, 2006 IMPLAN *** denotes significance at 1% level

Figure 1 - Global Concentration Index Comparison (2002 & 2006)



Source: 2002, 2006 US CBP, 2002 WholeData.Net, 2006 IMPLAN

Figure 2 - Spatial-Aspatial Index Ratio and Optimal Neighborhood Bandwidths



Source: 2002, 2006 US CBP, 2002 WholeData.Net, 2006 IMPLAN

Note: In this example, the optimal bandwidth for the biodiesel sector was 33 miles and for the residential solar sector was 31 miles.

Figure 3 - Residential Solar Value Chain Employment Concentration

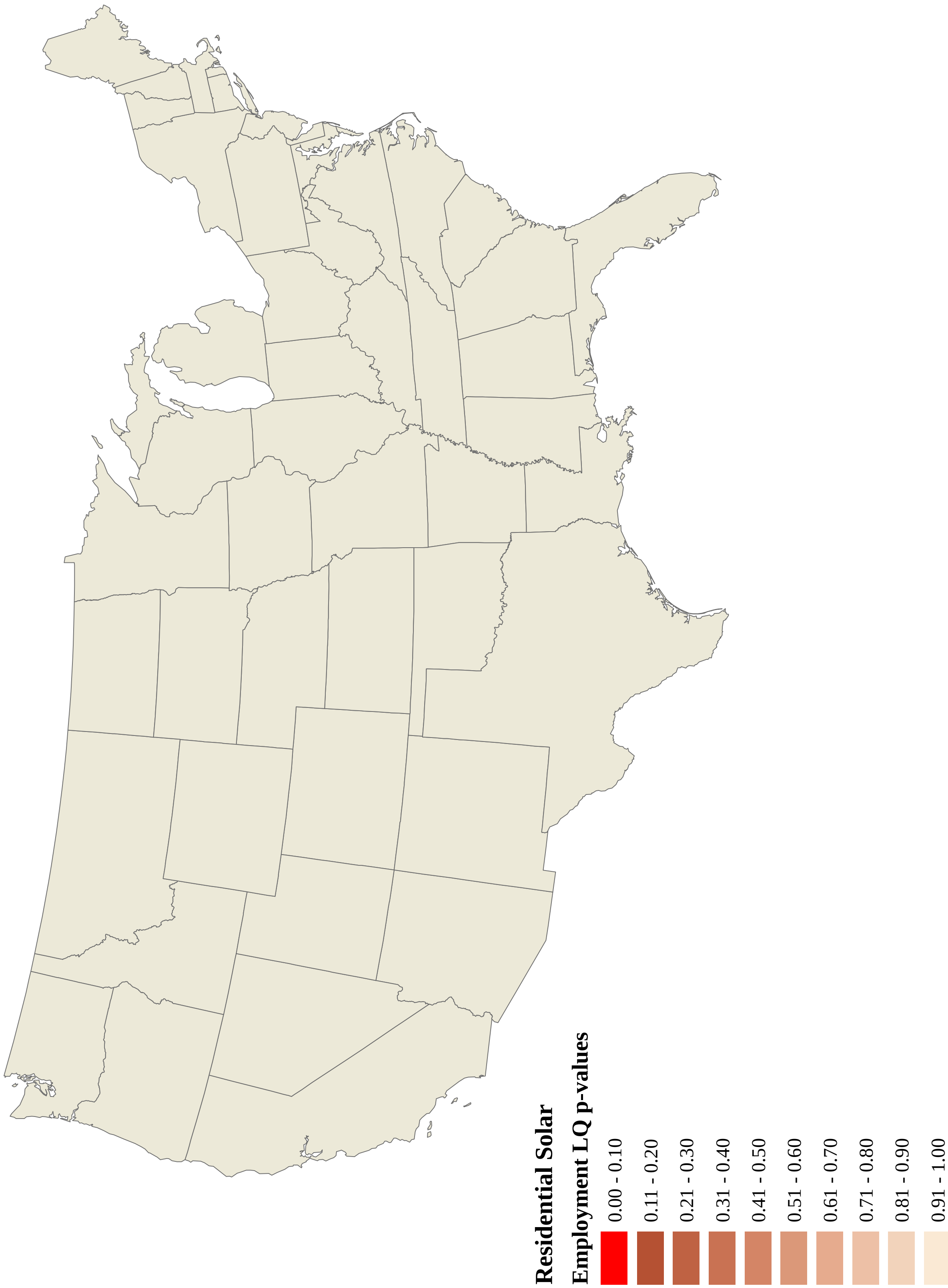


Figure 4 - Residential Solar Value Chain Establishment Concentration

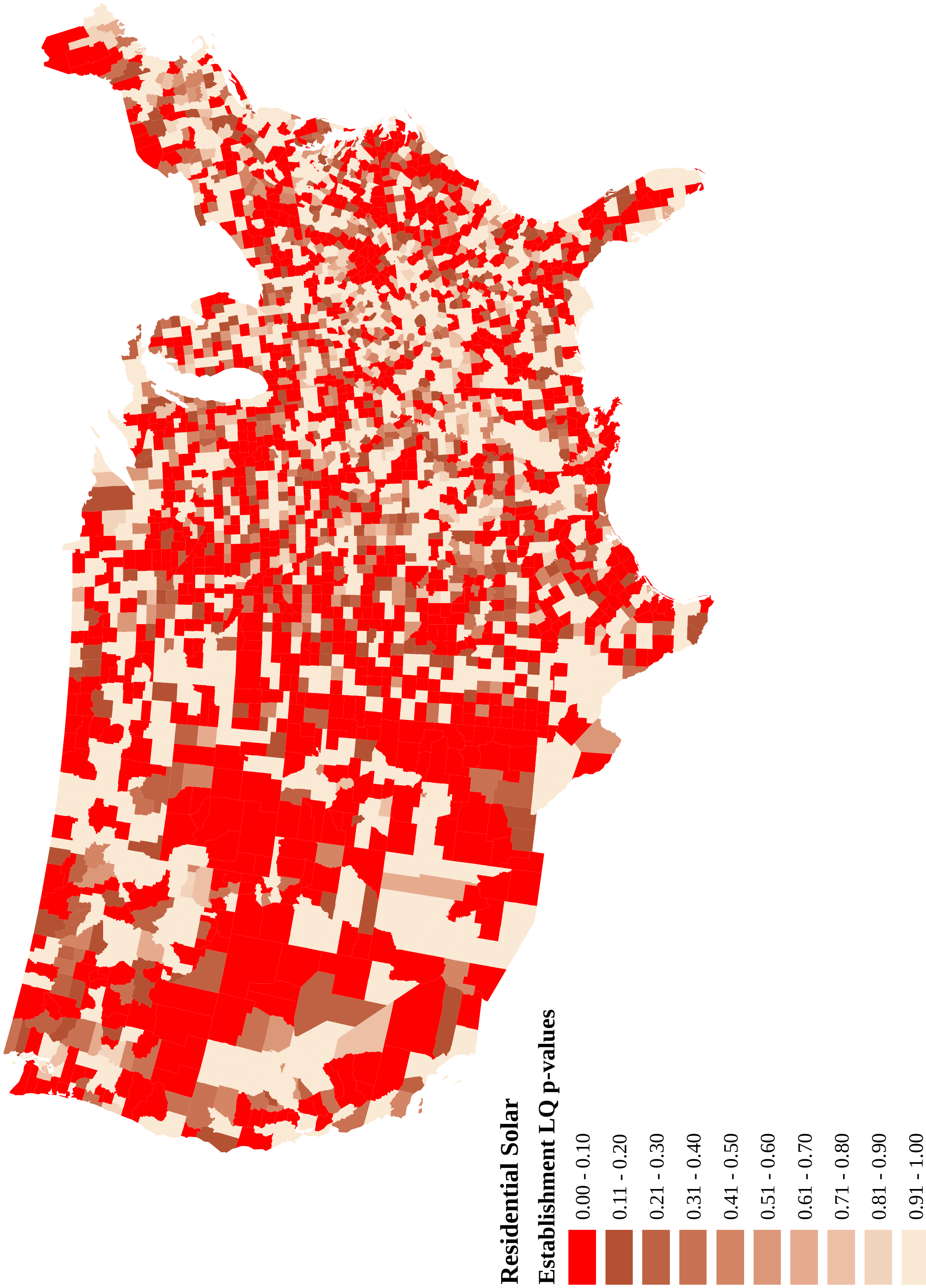


Figure 5 - Biodiesel Value Chain Employment Concentration

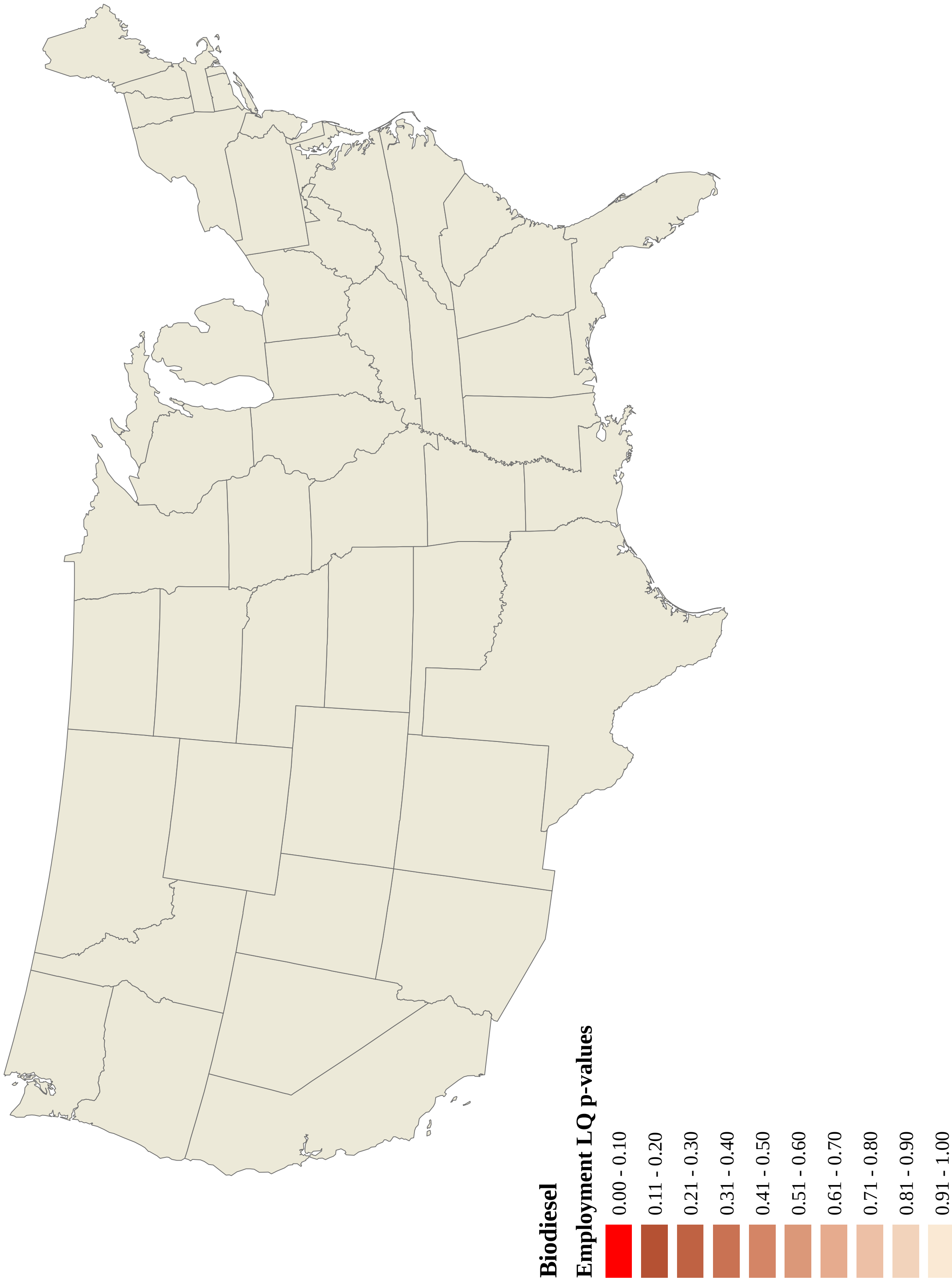


Figure 6 - Biodiesel Value Chain Establishment Concentration

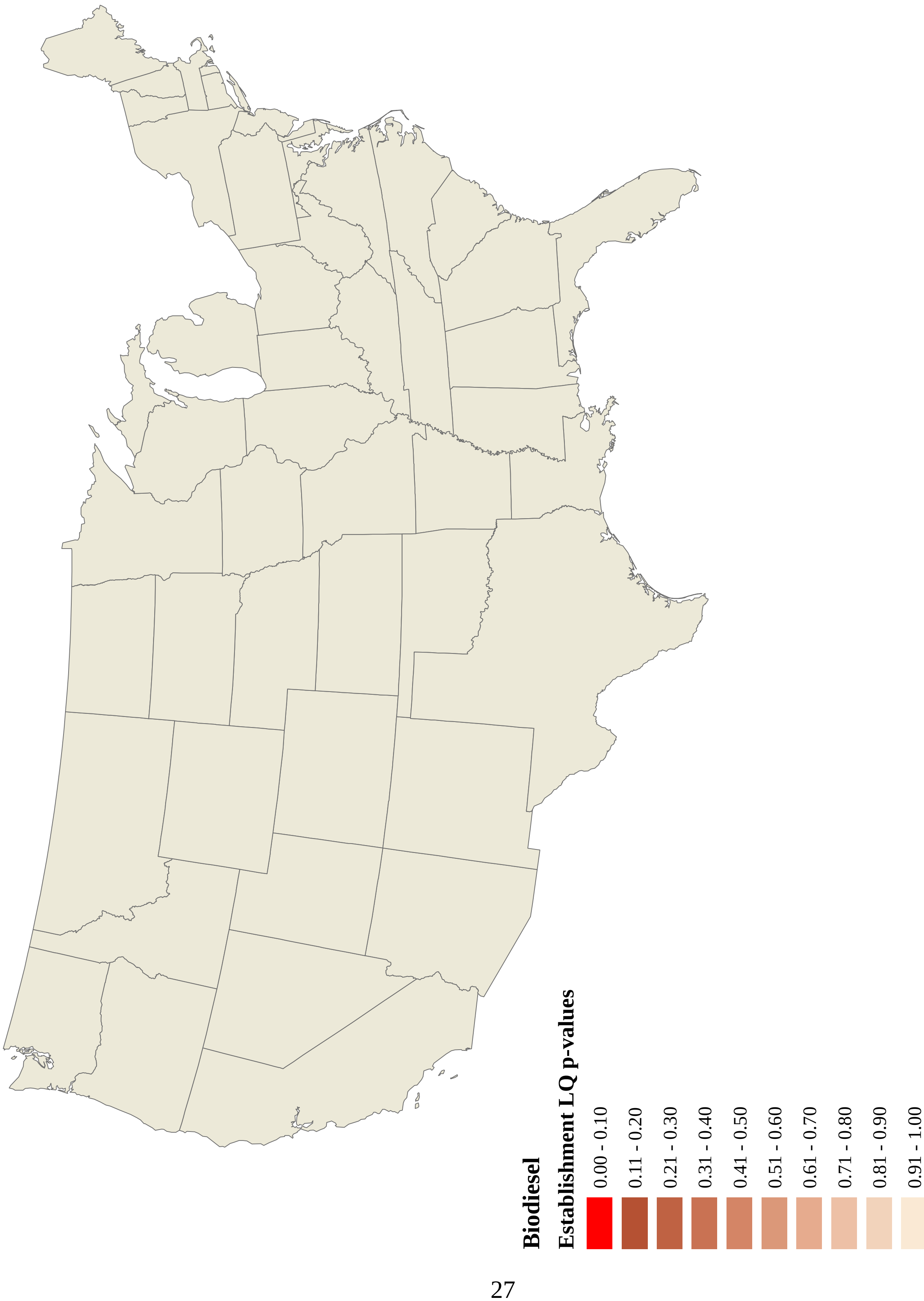


Figure 7 – US Solar Resource Map