STATSAMERICA User Guides

Proximity-Adjusted Location Quotients

Overview

The proximity-adjusted location quotient (PA-LQ) is a new measure of industry agglomeration that enhances the standard location quotient by taking into account two important aspects of industry clustering: (1) input-output (IO) linkages among related industries and (2) spatial spillovers of industry concentration across geographic units, defined here as U.S. counties. The PA-LQ also mitigates several frequently cited issues with the traditional location quotient (LQ). These include:

(1) Extremely high LQs are often found in less-populated and remote counties that typically have simple and highly specialized industry structures. Such counties emerge with high LQs but very low employment counts. This can occur because an LQ for one county doesn't take into account the industry shares of neighboring counties, despite the fact that counties in close proximity to each other are often highly integrated.
(2) The LQ's bias toward specialization without regard to size leads to statistical issues, as the distribution of LQs across counties will be highly skewed and contain many outliers that have to be dealt with in the context of regression and hypothesis testing, for example.
(3) The LQ based on national industry statistics does not usually include key information on related industries. Each industry-county pair has its own LQ, which fails to account for industries that may be closely integrated through input-output linkages, for example. This information is essential if one wishes to capture not only the concepts of specialization and spatial concentration, but also the broader concept of agglomeration.

The PA-LQ is computed from four different components, each of which captures an aspect of industry agglomeration:

(1) LQ_{ig}: The standard location quotient that measures the industry concentration of industry *i* in county *g*.
(2) ALQ_{ig}: The measure for industry concentration of related industries that connect industry *i* through IO relationships in county *g*.
(3) WLQ_{ig}: The measure for industry concentration of industry *i* in neighboring counties of county *g*.
(4) $WALQ_{ig}$: The measure for industry concentration of related industries that connect industry *i* through IO relationships in neighboring counties of county *g*.

The PA-LQ is the geometric mean of the four components, calculated as, for industry *i* in county *g*:

$$PALQ_{ig} = [LQ_{ig} * (ALQ_{ig} + \varepsilon) * (WLQ_{ig} + \varepsilon) * (WALQ_{ig} + \varepsilon)]^{1/4}$$

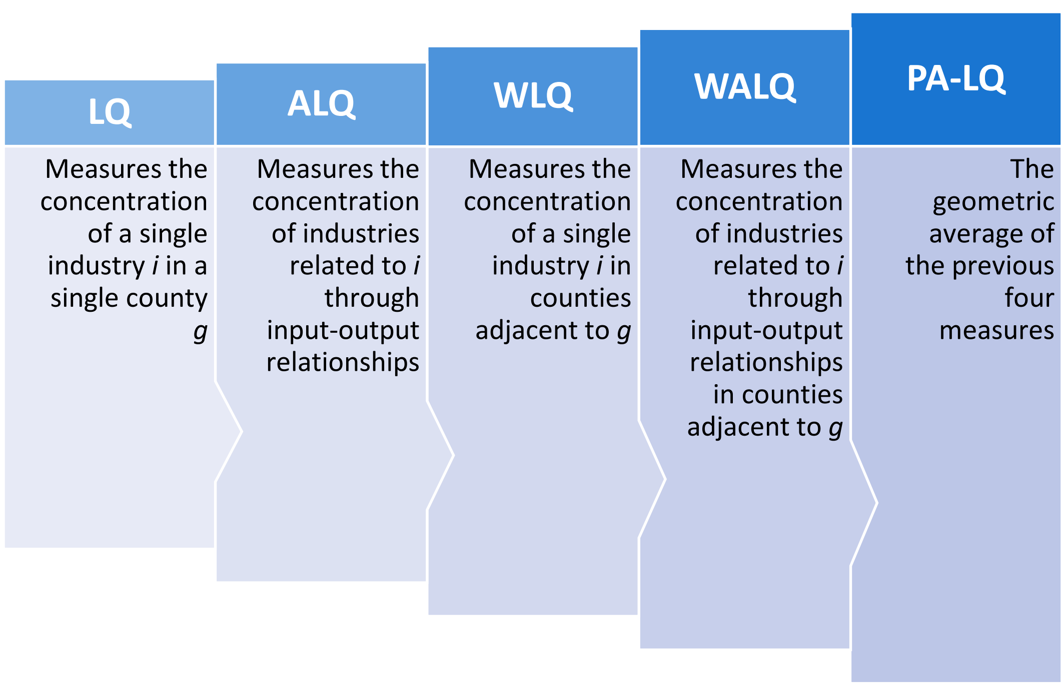
Where ε is a small positive number to make sure the ALQ, WLQ, and WALQ aren't zero, though we allow the PA-LQ to be zero if the LQ is zero to reflect the fact that not all industries are present in all counties.

A practical advantage of the PA-LQ over the traditional LQ is the attenuation of extreme values of the LQ in less-populated and remote counties, as well as the identification of large clusters in metropolitan areas like New York, Los Angeles, Chicago and others. In descriptive and applied regression analysis, the PA-LQ performs better than the LQ at capturing industry clusters and their hypothesized effects. Analyses of internal consistency of the PA-LQ also reveal the importance of input-output linkages on regional industry concentration.

How to use the data

A simple application of these data are the quantile rankings of the PA-LQ to see where selected industries in a particular county compare to the rest of the nation. Industry-county pairs whose PA-LQ falls into the 90-95th or 95-100th quantile can be considered as being highly concentrated in the county relative to the U.S. Because of the composite nature of the PA-LQ, it doesn't register values nearly as extreme for small counties with high levels of industry specialization and, thus, it can be trusted to a greater extent than the traditional LQ.

Another advantage of the PA-LQ that enhances its practicality is how it can be easily disaggregated into its four component parts. The component parts build on each other, with each one providing more information than the last:



Though the PA-LQ is the most complete measurement because it's an amalgam of the measures, each component part is informative on its own. The LQ is the most basic building block, and, despite its many flaws, it can still be informative for counties that aren't too small and have a relatively diverse industry structure. The ALQ can be used for identifying industries related through input-output linkages, which is useful to see which industries are dependent on each other. A practitioner could use the ALQ to consider, for example, how negative or positive shocks to one industry may impact other industries that are linked through supply chains.

The WLQ is concerned with regional spatial concentration and agglomeration within a single industry. A practitioner may use the WLQ to see the extent to which an industry in their county is regionally significant. This could be useful for identifying industries that could be susceptible to regional economic downturns. The WALQ combines the ALQ and WLQ to measure the concentration of industries related through both input-output and spatial linkages. It's meant to reinforce the strength of an industry's concentration in a county if the same industry and its related industries concentrate in surrounding counties. Practitioners can use this measure to get an idea of which industry clusters (which for the purposes of this measure are those industries connected through IO linkages) are regionally significant. Practitioners can use this information to consider which industries should be prioritized for business expansion, for instance.

Finally, the PA-LQ is computed as the geometric average of the four components. We use the geometric average to attenuate large values of the LQ that are likely to occur in smaller counties with simple industrial structures. Large values of any of the measures would artificially skew the PA-LQ upward if the arithmetic mean were used. A caveat with the PA-LQ is our decision to measure the industry relatedness in terms of input-output linkages, which omits other factors that can connect industries like worker skills or technology, for example. Nonetheless, our decision is based on the existing literature.

You may notice from the formula above that each of the components is given equal weight in the geometric mean calculation. This is another subjective decision but one that can easily be modified based on the user's judgment of importance. Users can simply download the data for each of the component measures, which we provide, and weight some components more heavily than others if they wish. If, for example, a practitioner is especially interested in the spatial and IO concentration of an industry and less interested in the single industry's concentration in a single county, they can weight the WALQ more heavily and discount the LQ in the calculation. Let's see how this could work using an example.

In Lake County, Florida, in the year 2015, the "insurance carriers and related activities" sector (NAICS 524) has a relatively high PA-LQ, in the 75-90th percentile. How is this calculated?

$$PALQ_{ig} = [0.349 * 0.727 * 1.035 * 0.826]^{1/4}$$

$$PALQ_{ig} = 0.682$$

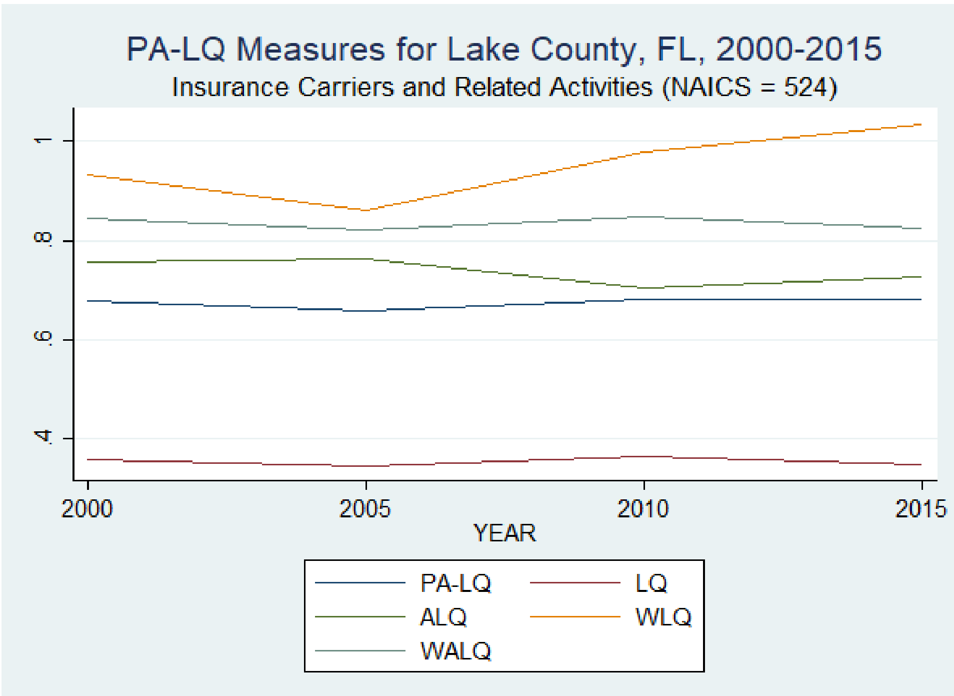
But say that a practitioner is more interested in the components that measures regional concentration in neighboring counties (i.e., Orange County [Orlando], Sumter County, Polk County, Marion County and Volusia County) and industries related through IO linkages. The practitioner can then weight the ALQ, WLQ, and WALQ more heavily. Let's say they decide to weight the ALQ, WLQ and WALQ by a factor of three. What would the calculation look like now?

$$PALQ_{ig} = [0.349^1 * 0.727^3 * 1.035^3 * 0.826^3]^{1/10}$$

$$PALQ_{ig} = 0.780$$

The modification couldn't be easier. Simply apply the weights you want as exponents and take the *n*th root, where *n* is the sum of the weights.

The time series component of the measures allows users to see how industrial structure has changed over the years. We provide the PA-LQ and the component measures during four snapshots in time, 2000, 2005, 2010, and 2015. For insurance carriers and related activities in Lake County, the measures haven't changed much over time:



Indeed, the only measure that changed considerably over this time is the WLQ (yellow line), which indicates an increase in the concentration of this industry in the counties adjacent to Lake County during the 15 years of the time series. But even this didn't move the needle substantially when it comes to the component measure.

Due to its magnitude, the gap between the traditional LQ and the other measures is worth mentioning. If one were to accept the LQ as the principal measure of industry concentration, insurance carriers and related activities wouldn't even register as structurally important to Lake County. We can interpret the LQ of about 0.35 in 2015 as meaning that Lake County's workforce is only about 35 percent as concentrated in this industry as the nation. This measure alone gives the impression that this isn't a very important sector in Lake County. By contrast, the PA-LQ in 2015 is 0.68, which indicates that, when IO linkages and the concentration of the industry in neighboring counties is factored in, the industry is much more important to the region than the traditional LQ lets on.