Appendix

A.1. Methodology

This edition of *At the Heart of Texas* features new industry cluster definitions, namely those developed by the Cluster Mapping Project at the Harvard Business School.[1] The 67 clusters (51 traded and 16 nontraded) provide a comprehensive view of interconnected upstream and downstream industries. The cluster mapping project definitions capture real interindustry linkages rather than arbitrary groupings, producing clusters that better reflect how regional economies function.

While clusters based on this definition are defined by their North American Industrial Classification System identifier (or NAICS code), they do not necessarily correspond to a specific broad NAICS sector. Rather, the clusters are made up of interrelated subsectors or industries (from the three-digit level down to the six-digit level) that are part of different NAICS supersectors (two-digit level). In some instances, individual NAICS industries may be found in multiple clusters, and not all existing industries are included in a cluster.

We altered the cluster definitions by condensing the original 67 clusters and subclusters into 29 broader clusters to create a simpler and more straightforward view of the industry mix in Texas and its metro areas. For example, our agribusiness cluster includes agricultural inputs and services, food processing and manufacturing, fishing and fishing products, and livestock processing clusters from the Cluster Mapping Project.

The Cluster Mapping Project analysis focuses primarily on traded clusters, or industries that are export oriented. Thus, some important industries such as government and social assistance were omitted. We included the retail trade, social assistance (NAICS 624) and government sectors. Government broadly includes federal, state and local government workers, with the exception of those employed in education, health services, social assistance, utilities, and aerospace and defense. For these clusters, government jobs are counted together with private sector employment. More information on the industries included in each cluster is available at www.dallasfed.org/research/heart/tables.

For purposes of our metropolitan area analysis, we used Census Bureau definitions of metropolitan statistical areas (MSAs) for Amarillo, Austin, Beaumont–Port Arthur, El Paso, Houston, Lubbock, McAllen and San Antonio. For Dallas and Fort Worth, we used the Census Bureau's definitions of metropolitan divisions. For both Midland–Odessa and Tyler–Longview we combined two MSAs into one.

The analysis uses data from the Quarterly Census of Employment and Wages (QCEW), which contains employment, wages and establishment information by industry down to the six-digit NAICS level. QCEW data for Texas and its metros were retrieved from the Texas Workforce Commission, while data for the U.S. came from the Bureau of Labor Statistics.

QCEW data are suppressed at some levels of detail when the number of establishments does not reach a certain threshold, risking the confidentiality of individual respondents. The data are only available quarterly, so annual employment data were calculated by taking the average of quarterly employment, and annual total wages were calculated by summing quarterly wages. Discrepancies may exist in the wage data because some industries may be unsuppressed in one quarter and suppressed in another, leaving annual wage data incomplete. In instances where wage data for a particular NAICS code were available in some quarters and missing in others within the same calendar year, observed quarterly wage data were applied to or substituted for missing quarters. Additionally, because of suppression issues, employment in some industries and years may be understated.

The employment and wage data were aggregated into clusters using NAICS codes as per the Cluster Mapping Project definitions with Dallas Fed modifications. For each cluster, the component industry annual employment and wage data were summed and excluded industries subtracted, where applicable. Average wages for each cluster were calculated by dividing total cluster wages by total employment.

Location quotients (LQs) are the ratio of cluster employment in each metro divided by total metro employment to cluster employment in the U.S. divided by total U.S. employment.[2] An LQ greater than 1, therefore, means the cluster's share of employment in the metro is greater than its share of U.S. employment, indicating the cluster is more concentrated in the metro than in the U.S. overall.

Demographic data are from the Census Bureau's American Community Survey. For median household income, the percent receiving SNAP (Supplemental Nutritional Assistance Program) or food stamp benefits,

and the percent receiving cash public assistance, we utilize the five-year ACS data. For all other demographic data, the one-year ACS estimates are used. In our analysis, we typically compare 2023 data with those for 2016.

A.2. Location quotient and average wage equations

1. Cluster i location quotient = $\dfrac{\dfrac{e_i}{e}}{\dfrac{E_i}{E}}$,

   where e_i = metro employment in cluster i, e = metro total employment, E_i = U.S. employment in cluster i and E = U.S. total employment.

2. Cluster average wage = $\dfrac{w_i}{e_i}$,

   where w_i = total wages paid in cluster i and e_i = employment in cluster i.

A.3. Additional data

Detailed cluster location quotient, employment, wage and demographic data are available at www.dallasfed.org/research/heart/tables.

Notes

[1] See U.S. Cluster Mapping Project, Institute for Strategy and Competitiveness, Harvard Business School; www.isc.hbs.edu/resources/Pages/data.aspx.
[2] See A.2. for the full equations.