

Appendix

A.1. Methodology

This report uses industry cluster definitions developed by the StatsAmerica Innovation Project, funded by the U.S. Commerce Department’s Economic Development Administration and assembled by the Purdue Center for Regional Development and the Indiana Business Research Center.¹ The original 17 clusters and six manufacturing subclusters provide a comprehensive view of the interconnected upstream and downstream industries.²

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.

The StatsAmerica analysis focuses only on “traded” clusters, or industries that are export oriented; thus, some large and important industries were omitted. We altered some of the cluster definitions to create a more complete view of the industry mix in Texas and its metro areas. We included the Retail, Construction and Utilities NAICS supersectors, the Food Services sector (NAICS 722) and the Government sector that includes federal, state and local government workers, with the exception of those employed in education or health services. We took ambulatory health care services and health and personal care stores out of the Biomedical StatsAmerica cluster and created a separate Health Services cluster that includes public and private employment at hospitals, ambulatory health care services, and nursing and residential care facilities.

We combined the StatsAmerica Energy and Mining clusters and aggregated all of the mining and support activities subsectors up to the three-digit NAICS level. We modified StatsAmerica’s Education and Knowledge Creation cluster to include only public and private educational services. Additionally, to look at the manufacturing sector in more detail, we broke up the manufacturing grouping into its six subcluster components as defined by StatsAmerica.

For purposes of our cities 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 Midland–Odessa and Tyler–Longview, in each area, we combined the two MSAs into one.

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

TWC and BLS data may be suppressed at some levels of detail when the number of firms does not reach a certain threshold and the confidentiality of individual firms may be at risk. TWC data are only available quarterly, so annual employment data were calculated by taking average quarterly employment; annual total wages were calculated by summing quarterly wages. Thus, some 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 when 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/ substituted for missing quarters. Additionally, because of suppression issues, employment in some industries with fewer firms is potentially understated.

The detailed employment and wage data were aggregated into clusters based on the StatsAmerica cluster definitions using NAICS codes to match the raw data with the cluster definitions. For each cluster, the component industry annual employment and wage data were summed and excluded industries were subtracted. Average wage data for each cluster were calculated by taking total wages for the aggregated cluster and dividing by total employment in the cluster.

Location quotients (LQs) were calculated by taking cluster employment in each metro divided by total metro employment, over cluster employment in the U.S. divided by total U.S. employment.³ An LQ greater than 1, therefore, means that the cluster’s share of total

employment in the metro is greater than its share of total U.S. employment, indicating that 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 2017, population and median household income are used; other detail is from 2016. We compared those with data from the 2010 survey. In all three years, only one-year estimates were used for analysis.

The Kauffman Startup Activity Index measures business creation in the 40 largest metropolitan areas in the U.S. The index is based on three indicators: the rate of new entrepreneurs starting businesses, the percentage of new entrepreneurs not unemployed before starting a business and the number of newly established employer businesses per 1,000 employer firms.⁴

A.2. Changes to StatsAmerica Cluster Definitions

- Split Manufacturing grouping into individual sub-cluster components.
- Changed the Education and Knowledge Creation cluster as follows:
 - Removed NAICS 51111 (Newspaper Publishers), NAICS 51112 (Periodical Publishers), NAICS 51113 (Book Publishers), NAICS 516 (Internet Publishing and Broadcasting), which are already counted in Printing and Publishing, and NAICS 519 (Other Information Services).
 - Included both public and private employment in Educational Services (NAICS 61). In the 2016 edition, Education Services consisted only of private employment in NAICS 61.
- Added NAICS 519 to the Printing and Publishing cluster but removed NAICS 51911 and NAICS 51919 to avoid double counting.
- Combined the Energy and Mining clusters and aggregated all subsectors in NAICS 212 and 213.
- Removed NAICS 621 and NAICS 446 from the Bio-medical cluster and created a Health Services cluster that includes both government and private employment in NAICS 621, NAICS 622 and NAICS 623. In the 2016 edition, Health Service cluster was made up of only private employment in NAICS 621, NAICS

622 and NAICS 446.

- Added Retail (NAICS 44–45), Construction (NAICS 23) and Utilities (NAICS 22) supersectors.
- Added a Government sector, which includes total federal, state and local government workers, excluding those employed in public education and health care.
- This edition follows StatsAmerica's original definition of the Arts, Entertainment, Recreation and Visitor Industries cluster, which we refer to as the Recreation cluster. The first edition augmented the original definition by including all of NAICS 72.
- Added a Food Services cluster, composed of NAICS 722.

A.3. Location Quotient and Average Wage Equations

$$\frac{\sum e_i}{\sum e}$$

1. Cluster location quotient = $\frac{\sum e_i}{\sum E_i} \cdot \frac{\sum E}{\sum e}$,

where e_i = metro's cluster employment, e = metro's total employment, E_i = U.S. cluster employment and E = U.S. total employment.

2. Cluster average wage = $\frac{\sum x_i}{\sum e_i}$,

where x_i = total wages paid in each cluster and e_i = employees in each cluster.

A.4. Additional Data

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

Notes

¹ As used by Diane F. Primont and Bruce Domazlicky in "Industry Cluster Analysis for the Southeast Missouri Region," Center for Economic and Business Research, September 2008.

² Detailed cluster definitions can be found on the StatsAmerica website, www.statsamerica.org/innovation/about.html.

³ See A.3. for the full equations.

⁴ See www.kauffman.org/kauffman-index/reporting/startup-activity.