

2020 Wesleyan/CT Data Collaborative
Economic Development in Connecticut (Comparative, Exploratory Study)
Conner Sexton

Data Overview:

The primary data sources used for this study were the U.S. Census Bureau and the Bureau of Economic Analysis. Geographic areas are measured by Metro/Micro -politan Statistical Areas (MSA) and the data is time series (2010 - 2018). In some cases, the Torrington, CT micropolitan area is excluded due to lack of data. U.S. Census Bureau data was retrieved from the data.census.gov tool and BEA data was queried using their [developer API](#) in conjunction with the R-package, [bea.R](#).

[CBSA Guidelines](#)

March 2020 release of 2010 U.S. Census Bureau CBSA designations (metro/micro - politan statistical areas).

CBSA Code	CBSA Title	State
14860	Bridgeport-Stamford-Norwalk	Connecticut
25540	Hartford-East Hartford-Middletown	Connecticut
35300	New Haven-Milford	Connecticut
35980	Norwich-New London	Connecticut
45860	Torrington (Micro)	Connecticut
49340	Worcester	Connecticut
35620	New York-Newark-Jersey City	New York
14460	Boston-Cambridge-Newton	Massachusetts

[Bea Tables](#)

Data: Real GDP

We used BEA's API tool to query data on Real GDP (GDP adjusted for inflation) for our MSA's of interest, from 2010 to 2018. BEA's *Regional Data Set* and *CAGDP9* table does not include data on Micropolitan Statistical Areas, therefore Torrington is not included.

[US Census Bureau](#)

Data: [Population Statistics](#)

Year-to-year population changes across our Metro and Micro -politan areas of interest from 2010 to 2019. More detailed descriptions on specific variables included in this data can be found through the hyperlink above.

Data: American Community Survey (1-Year Estimates)

Household Income

Household income data was retrieved from the U.S. Census Bureau's *American Community Survey* and was collapsed into the following income brackets, as defined by the U.S. Census Bureau:

Bracket Label	Bracket Range	Notes
poverty	\$24,999 and below	Below or near poverty line
low_income	\$25,000 to 34,999	
middle_class	\$35,000 to 99,999	
upper_middle_class	\$100,000 to 199,999	
upper_class	\$200,000 or more	

Age

Age data was retrieved from the U.S. Census Bureau's *American Community Survey* and reflects the percentage of an MSA's population that falls within each age category. For clarity, the age categories were collapsed into the following variables:

Age Label	Age Range
age_one	19 and below
age_two	20 to 34
age_three	35 to 54
age_four	55 to 64
age_five	65 to 84

Educational Attainment

Educational attainment data was retrieved from the U.S. Census Bureau's *American Community Survey* and reflects the percentage of an MSA's population across different ages and levels of educational attainment. We decided to use the following variables:

Education Label	Age	Educational Attainment
y_adult_nohs	18 - 24 (young adult)	less than high school
y_adult_hs	18 - 24 (young adult)	high school diploma
y_adult_college_high	18 - 24 (young adult)	college diploma or more

Education Label	Age	Educational Attainment
adult_nohs	25 and over (adult)	less than high school
adult_hs	25 and over (adult)	high school diploma
adult_col_high	25 and over (adult)	college diploma or more

Employment (by industry)

Employment by industry data was retrieved from the U.S. Census Bureau's *American Community Survey* and reflects the percentage of an MSA's population 16 years and older employed across different industries. We decided to use the following variables:

Employment Label	Description
total_emp	Total Employment
agri	Agriculture, forestry, fishing and hunting, and mining
construction	Construction
wholesale	Wholesale trade
retail	Retail trade
transportation	Transportation and warehousing, and utilities
information	Information
finance	Finance and insurance, and real estate and rental and leasing
professional	Professional, scientific, and management, and administrative and waste management
education	Educational services, and health care and social assistance
ammanities	Arts, entertainment, and recreation, and accommodation and food services

Data: Statistics of U.S. Businesses (SUSB)

We used SUSB data reported by the U.S. Census to obtain numbers on establishments across our MSA's of interest. This data is organized in a year-to-year manner, with 2016 as the most current available year. The data we collected reflects numbers from 2010 to 2016. The two main areas of interest were number of establishments by NAICS code and number of establishments by establishment size (labels displayed below). Additional data management included taking the number of establishments in each NAICS/size category and dividing this by the total by year/MSA, respectively.

NAICS codes can be found [here](#).

Establishment sizes:

Label	Description
1	Total
2	0-4
3	5-9
4	10-19
5	<20
6	20-99
7	100-499
8	<500
9	500 and greater