

Assignment B-2

Estimate the determinants of the change in total population in US counties over 2010-2021. Your sample consists of one hundred counties.

1. First, run linear regressions for the two models below.

$$\text{I. } \hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4$$

$$\text{II. } \hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 + b_6x_6 + b_7x_7$$

Dependent variables:

The dependent variable is county growth in total population from 2010 to

2021. In the models, the dependent variable is $\frac{2021 \text{ total pop.} - 2010 \text{ total pop.}}{2010 \text{ total pop.}}$.

Independent variables:

$x_1 = 1$ if total 2010 county population is greater than 250,000
= 0 otherwise

x_2 - mean winter temperature within the state.

x_3 – county poverty rate in 2010.

x_4 – percent of county residents in 2010 who are foreign born.

x_5, x_6, x_7 represent dummy variables indicating the Census Region the county is in.

Use the US Census to construct the population, poverty, and foreign-born variables.

2. Turn in your SAS regression output as well as the constructed variables used to run the regressions. Also turn in your SAS programs.

3. Write a short paper interpreting the results, addressing the following points.
 - A. Justify the models. What effect do you expect each of the independent variables to have on the dependent variable? (Your expectations may differ from the estimated relationships).
 - B. Discuss your results. Do your results correspond with the expectations discussed in part A? Which results are statistically significant?
 - C. In what way do the relationships estimated in model I change when the regional dummy variables are added in model II?
 - D. Which results do you find most surprising? Explain.