Stata Exercise on Bartik Instruments
Graduate Urban Economics, SUFE

This exercise closely follows Goldsmith-Pinkham, Sorkin, and Swift "Bartik Instruments: What, When, Why, and How" (AER 2020). The goal is to simulate the simplest case of two industries and a single period of growth (or one observation of growth between two periods). We will use the following setup:

Wage growth in location l, y_l, is a function of employment growth, x_l, and an unobservable amenity shock, a_l

$$y_l = \beta x_l + \delta a_l + \epsilon_l \tag{1}$$

Employment growth in l can be decomposed into local growth in each of the two industries ($k \in \{1, 2\}$):

$$x_l = g_{l1} z_{l1} + g_{l2} z_{l2} \tag{2}$$

In turn, local growth is a function of national growth, g_1 and g_2, and an idiosyncratic local component:

$$g_{lk} = g_k + \tilde{g}_{lk} \tag{3}$$

The local amenity shock is just the sum of the idiosyncratic growth terms:

$$a_l = \tilde{g}_{l1} + \tilde{g}_{l2} \tag{4}$$

Finally, make the $\tilde{g}_{lk}$ and ϵ_l terms random normal variables with distribution:

$$\tilde{g}_{lk} \sim N(\mu_g, \sigma_g) \tag{5}$$
$$\epsilon_l \sim N(0, \sigma_e) \tag{6}$$

Simulate this set-up and make the following 7 parameters global variables so you can easily change them. I have listed the values I will use in my code so that we can compare.

1. Count of locations L; My values: $L = 100$
2. National growth rates g_1, g_2; My values: $g_1 = 0.1$, $g_2 = 0.05$
3. Coefficients: β,δ; My values: $\beta = 1.5$, $\delta = 1$
4. Distribution parameters: $\mu_g, \sigma_g, \sigma_e$; My values: $\mu_g = 0.025$, $\sigma_g = 0.01$, $\sigma_e = 0.005$

Questions:

1. Show that the Bartik instrument is equivalent to IV regression using industry shares
2. Show that *endogenous* shares violate Bartik identification assumption
3. If you have extra time: repeat exercise for two period / two industry case and show that first stage is equivalent to shares interacted with time period fixed effects