Graduate Labor Economics

Lecture 20: Shift-Share Instruments

Brendan M. Price*
Federal Reserve Board

 Email: brendan.m.price@frb.gov. These notes are adapted from UC Davis course ECN 250A, which I taught in Winter 2018 and Spring 2019. Any views or opinions expressed here are my own and do not necessarily represent the views or policies of the Board of Governors of the Federal Reserve System or its staff. This version: March 21, 2020.

Today's lecture

- Emerging literature on shift-share instruments
  - Borusyak, Hull, and Jaravel (2018)
  - Goldsmith-Pinkham, Sorkin, and Swift (2018)
  - Adao, Kolesar, and Morales (2018)
  - Jaeger, Ruist, and Stuhler (2018)
- Transitioning into third year

Bartik shocks

- Classic Bartik IV: omnibus measure of local labor demand

$$B_{it} \equiv \sum_j \frac{L_{ij}}{L_i} \text{dlog}\, L_j$$

- Prominent citations:
  - Bartik (1991)
  - Blanchard and Katz (1992)
  - Autor and Duggan (2003): use $\text{dlog}\, L_{-i,j}$ to avoid mechanical bias
  - Notowidigdo (2019)
- Widely used, currently subject to much debate
  - Pros: easy to compute, always available, high-powered
  - Cons: black box, may not fully isolate demand shifts
- Alternative: shift-share IVs using specific shocks (e.g., ADH 2013)

Shift-share instruments are widely used

- Lots of influential examples:
  - Immigrant inflows (Card 2001)
  - Firm-level trade shocks (Hummels et al. 2014)
  - Pharmaceutical market volume (Acemoglu and Linn 2004)
  - Credit supply shocks (Greenstone, Mas, Nguyen, forthcoming)
- Same basic structure, similar econometric issues
- Papers listed on syllabus cite many more examples

Shares vs. shocks: the big debate

- Prominent debate over requirements for causal identification
- Goldsmith-Pinkham et al.: need exogenous shares
  - Seems unlikely to hold in many empirical settings
  - Implies a pessimistic view of shift-share instruments
- Borusyak et al.: okay to have exogenous shocks instead
  - Exogenous shares sufficient but not necessary
  - Can identify effects if shocks are "as good as random"
  - More optimistic view of shift-share instruments
- Both are general frameworks, but focus on the ADH application (with "locations" & "industries")

Borusyak, Hull, and Jaravel (2018): basic framework

- Start with cross-sectional case
- Basic estimating equation:

$$y_\ell = \beta x_\ell + w_\ell' \gamma + \varepsilon_\ell$$

  where x_ℓ is shock to location ℓ, w_ℓ are controls

- Concern: shock may be correlated with error term
- Solution: shift-share instrument

$$z_\ell = \sum_{n=1}^{N} s_{\ell n} g_n$$

  where $s_{\ell n}$ is exposure to sector n, g_n is exogenous shock to n

Numerical equivalence

- Main insight: location-level spec equivalent to industry-level spec

$$\begin{aligned}\hat{\beta} &= \frac{\frac{1}{L}\sum_{\ell=1}^{L}\left(\sum_{n=1}^{N} s_{\ell n} g_n\right) y_\ell^{\perp}}{\frac{1}{L}\sum_{\ell=1}^{L}\left(\sum_{n=1}^{N} s_{\ell n} g_n\right) x_\ell^{\perp}} \\ &= \frac{\sum_{n=1}^{N} g_n \left(\frac{1}{L}\sum_{\ell=1} L s_{ln} y_\ell^{\perp}\right)}{\sum_{n=1}^{N} g_n \left(\frac{1}{L}\sum_{\ell=1} L s_{ln} x_\ell^{\perp}\right)} \\ &= \frac{\sum_{n=1}^{N} \hat{s}_n g_n \overline{y}_n^{\perp}}{\sum_{n=1}^{N} \hat{s}_n g_n \overline{x}_n^{\perp}}\end{aligned}$$

  where $\hat{s}_n \equiv \frac{1}{L}\sum_{\ell=1}^{L} s_{\ell n}$ and $\overline{\nu} \equiv \frac{\sum_{\ell=1}^{L} s_{ln}\nu_\ell}{\sum_{\ell=1}^{L} s_{\ell n}}$

- We can think of this as a "change of basis"

The industry-level regression

- Equivalent industry-level regression:

$$\overline{y}_n^{\perp} = \alpha + \beta \overline{x}_n^{\perp} + \overline{\varepsilon}_n^{\perp}$$

- How to interpret $\overline{y}_n^{\perp}$?
  - Residualized growth rate of locations intensive in n
  - e.g., do textile-producing areas do better or worse than expected?
- How to interpret $\overline{x}_n^{\perp}$?
  - Residualized shock to locations intensive in n
  - e.g., are textile-producing areas more/less exposed to China shock?
- Instrument for $\overline{x}_n^{\perp}$ using quasi-random shocks g_n
  (e.g., imports from China within non-U.S. comparison countries)

Causal identification

- Shift-share IV is consistent if and only if

$$\sum_{n=1}^{N} s_n g_n \phi_n \to 0 \quad \text{as } N \text{ gets large}$$

  where:

  - s_n is average exposure to industry n
  - g_n is the shock to industry n
  - $\phi_n \equiv \frac{\mathbb{E}[s_{\ell n}\varepsilon_l]}{\mathbb{E}[s_{\ell n}]}$ is the average error term in n-intensive places

- Sufficient conditions for consistency:

  - Quasi-random shocks: $\mathbb{E}[g_n \mid \phi_n] = \mu$ for all n
  - Many independent shocks: mutually uncorrelated & $\sum_{n=1}^{N} s_n^2 \to 0$

When shift-share instruments fail

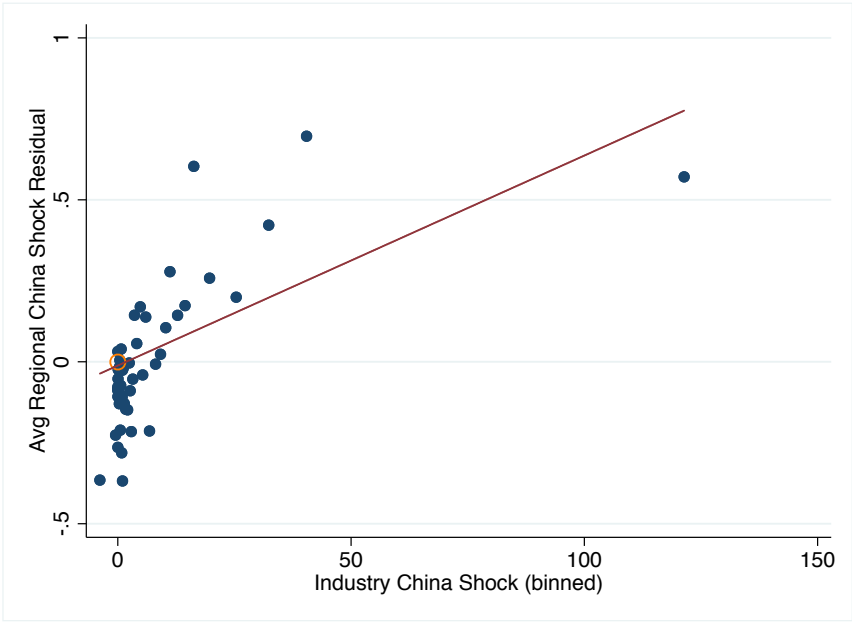
- Threats to quasi-randomness:
  - Unobserved industry shocks
  - Regional unobservables correlated with industrial composition
- Insufficient variation:
  - Risk of spurious correlation if a few industries dominate
  - Better to have lots of granular industries

Extensions

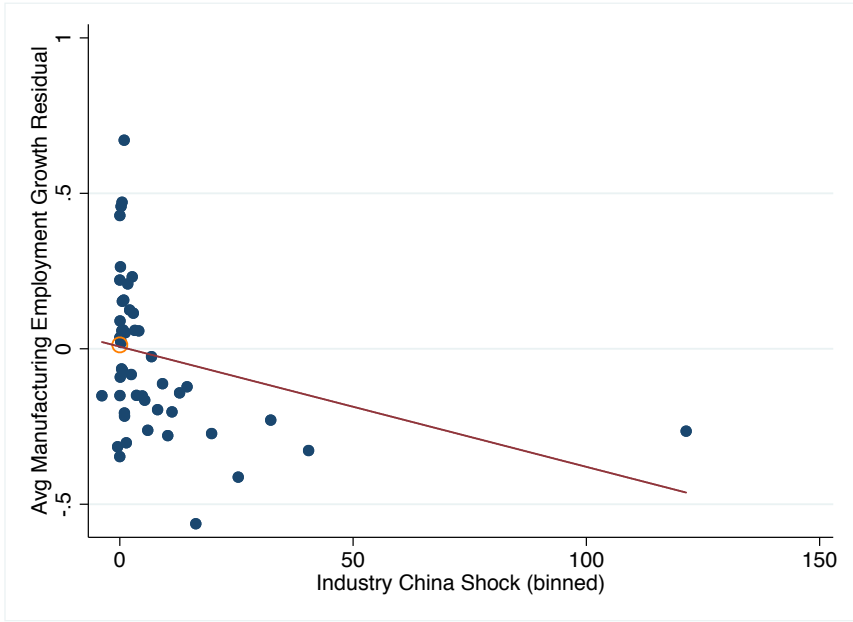
- Allowing for observables
  - Location regression: control for coarse sector shares
  - Industry regression: control for coarse sector dummies
- Adding a panel dimension
  - Include time-period fixed effects
  - Long panels can compensate for few industries
  - Often best to fix shares at baseline (bias-variance tradeoff)

ADH local imp. exposure renormalized into industry space

First stage



Reduced form



(Borusyak, Hull, and Jaravel (2018), Figure 1A)

Goldsmith-Pinkham, Sorkin, and Swift (2018)

- Shift-share is numerically equivalent to using *shares* as instruments
  - Bartik IV is a weighted average of industry-specific IV estimates
  - "Rotemberg weights" tell us which instruments drive the estimates
- Bartik is consistent if the baseline shares are exogenous
  - Need exogeneity *conditional on observables*
  - Example: quasi-random exposure to a common shock

The state of the debate

- Not yet clear how this debate will settle out
  - Will researchers claim to have exogenous shares? or shocks?
  - How skeptical will audiences/referees be of Bartik designs?
  - What diagnostics and specification tests will be expected?
- But it's a long-overdue debate
  - Clarifying the conditions under which Bartik IVs are valid
  - Providing new tools for assessing and interpreting them

Shift-share inference

- Adao, Kolesar, and Morales: classic shift-share understates SEs
  - Places with similar industry shares face similar China shocks
  - But they also face lots of *other* similar industry-based shocks
  - Cross-unit correlation in error term
- Clustering doesn't help much
  - Clustering by state accounts for spatial correlation
  - But it doesn't account for correlation based on industry structure
- Borusyak et al.: industry-level regression gets it right
  - Use heteroskedasticity-robust SEs
  - Asymptotically equivalent to Adao et al.'s suggested correction

Jaeger, Ruist, and Stuhler (2018): dynamic considerations

- Consider the "immigrant enclave" instrument
  - Baseline local immigrant shares by country of origin
  - New arrivals at national level by country of origin
- Tends to be highly correlated within locations over time
- Jaeger et al.: need to allow for dynamic adjustment process
  - Short run: quasi-fixed factors, immigrants depress wages
  - Long run: factor adjustment, GE mechanisms $\Longrightarrow$ wage recovery
- Potential solution: control for lagged immigrant shocks
  - Highly multicollinear $\Longrightarrow$ very challenging empirically

Transitioning into third year

- Last lecture $\implies$ life advice
- Transition from coursework to research is a challenging time

Modular tasks

- Challenge: research projects are big, sprawling, daunting
  - Not always clear where to allocate your time
  - Rarely get to feel like you've "finished" anything
- Advice: break projects into small, modular, manageable tasks
  - Today I'll read this paper
  - Today I'll clean this dataset
  - Today I'll work through an overly simplified model
  - Tomorrow I'll work on extending it
- Try to avoid letting tasks hang over till next day
  - Fixed cost of starting up
  - Not always possible (don't beat yourself up!)

The emotional side

- Impostor syndrome is more or less universal
  - Everybody feels behind
  - Everybody has intellectual insecurities
- Give self-care its due
  - Exercise, meditation, sleep, cooking, cleaning, vacation
  - Friends, family, support networks
  - Zero shame in seeking professional help
- Know your limits