Extra slides: shift-share and simulated instruments overview

LPO 8852: Regression II

Sean P. Corcoran

Shift-share or "Bartik" instruments

Another type of instrumental variable that appears often in labor economics (and increasingly, in education research) is the shift-share or Bartik (1991) instrument. "Simulated instruments" are a related idea.

Two motivating examples:

1. Trade: what is the impact of Chinese imports on local manufacturing employment?

$$Employment_{\ell t} = \beta_0 + \beta_1 ImportExposure_{\ell t} + e_{\ell t}$$

2. Immigration: what is the impact of immigration on the wages of native workers?

$$\Delta \ln(nativewages)_{\ell t} = \beta_0 + \beta_1 Immigration_{\ell t} + e_{\ell t}$$

Shift-share or "Bartik" instruments

There are lots of potential sources of OVB here. The local manufacturing base and immigrant share likely covary with omitted factors related to the outcome (local employment, local wages).

Shift-share or "Bartik" instruments

There are lots of potential sources of OVB here. The local manufacturing base and immigrant share likely covary with omitted factors related to the outcome (local employment, local wages).

Can we identify an instrument where:

1. $Cov(z, x) \neq 0$?
2. $Cov(z, e) = 0$?

Could be challenging in this context.

Shift-share or "Bartik" instruments

The Bartik instrument in each case is:

1. Trade: let $z_{\ell,k,t-1}$ be the baseline industry k employment shares in locality ℓ and $g_{k,t}$ be the growth in imports nationally from China to the U.S. in industry k.

$$B_{\ell t} = \sum_{k=1}^{K} z_{\ell,k,t-1} g_{k,t}$$

Shift-share or "Bartik" instruments

The Bartik instrument in each case is:

1. Trade: let $z_{\ell,k,t-1}$ be the baseline industry k employment shares in locality ℓ and $g_{k,t}$ be the growth in imports nationally from China to the U.S. in industry k.

$$B_{\ell t} = \sum_{k=1}^{K} z_{\ell,k,t-1} g_{k,t}$$

2. Immigration: let $z_{\ell,j,t-1}$ be the baseline share of immigrants from source country j in locality ℓ and $m_{j,t}$ be the change in immigration nationally from country j..

$$B_{\ell t} = \sum_{j=1}^{J} z_{\ell,j,t-1} m_{k,t}$$

Shift-share or "Bartik" instruments

The z's are the *shares*—a baseline measure of exposure to Chinese imports or immigrant workers. The g's are the *shifts*.

The local "treatment" is a weighted average of the national "treatment," where the weights are a measure of *baseline* local exposure. Differential exposure to common shocks.

Shift-share or "Bartik" instruments

The z's are the *shares*—a baseline measure of exposure to Chinese imports or immigrant workers. The g's are the *shifts*.

The local "treatment" is a weighted average of the national "treatment," where the weights are a measure of *baseline* local exposure. Differential exposure to common shocks.

Two examples from education finance: Corcoran and Evans (2010) and Boustan et al. (2013), on the effect of income inequality on public spending.

- Income inequality is endogenous
- Shift-share instrument: updating local measures of income inequality using baseline inequality and national trends in the income distribution.

Shift-share or "Bartik" instruments

When is the shift-share instrument valid? Must satisfy the relevance and exclusion restriction $Cov(z, x) \neq 0$ and $Cov(z, e) = 0$.

The relevance (first-stage) condition is saying that *predicted* exposure (based on baseline shares) must be correlated with *actual* exposure.

Shift-share or "Bartik" instruments

Goldsmith-Pinkham, Sorkin, and Swift (2020) is an excellent resource on shift-share instruments. They show:

- Exclusion restriction generally comes down to the shares: can we assume baseline shares are uncorrelated with the *change* in x and y?
- While it might be hard to assume shares are uncorrelated with levels of the outcome, changes are more reasonable.
- With baseline data, can check for *parallel trends*, a la difference-in-difference. Prior to the "shock," did locality with different baseline shares have common trends?
- The estimator can be decomposed into a weighted combination of just-identified estimates, each using a single share as an instrument. Allows one to assess where the bulk of the variation is coming from, and conduct over-identification tests.

Simulated instruments

Like Bartik instruments, simulated instruments are often predictions of a "treatment" variable based on pre-existing conditions interacted with a policy "shock."

Simulated instruments: Example 1

Biasi (2019) School Finance Equalization and Inter-generational Mobility

- Education is intended to equalize opportunity
- School funding is historically unequal, inadequate
- Does equalization in public education funding increase inter-generational mobility?
- Think of mobility (M^r) as a child's percentile in the national income distribution given the income percentile of their parents (Chetty et al, 2014). Think of the slope (β) between per-pupil spending and local per-capita income as a measure of the progressivity of school spending.

Simulated instruments: Example 1

The regression of interest is:

$$M_{cb}^{r} = \delta^{r} \beta_{sb} + \theta_{c} + \tau_{b} + \epsilon_{cb}$$

for a child in commuting zone c and birth cohort b, state s.

Simulated instruments: Example 1

The regression of interest is:

$$M_{cb}^{r} = \delta^{r} \beta_{sb} + \theta_{c} + \tau_{b} + \epsilon_{cb}$$

for a child in commuting zone c and birth cohort b, state s.

Problem 1: β_{sb} is surely correlated with other time-varying determinants of M_{cb}^{r}.

Simulated instruments: Example 1

The regression of interest is:

$$M_{cb}^{r} = \delta^{r} \beta_{sb} + \theta_{c} + \tau_{b} + \epsilon_{cb}$$

for a child in commuting zone c and birth cohort b, state s.

Problem 1: β_{sb} is surely correlated with other time-varying determinants of M_{cb}^{r}.

Solution 1: court-ordered school finance reforms produced plausibly exogenous shifts in β_{sb}. SFRs can be used as an instrument for β_{sb}.

Simulated instruments: Example 1

Problem 2: SFRs may create incentives for households to sort across districts, sorting that affects property values, community composition, peer characteristics. These things are correlated with β_{sb} and may have their own effects on income mobility M.

Simulated instruments: Example 1

Problem 2: SFRs may create incentives for households to sort across districts, sorting that affects property values, community composition, peer characteristics. These things are correlated with β_{sb} and may have their own effects on income mobility M.

Solution 2: predict β_{sb} based on school finance formula relevant for cb, but using baseline district characteristics (property values, enrollment, income). The result is a *simulated* slope, an instrument for actual slope.

Interpretation: isolating the variation in school finance progressivity that occurs via the formula change, not other reasons.

Simulated instruments: Example 1

Finding: 2SLS finds a 1 sd reduction in slope (about $4,500 reduction in the difference in per-pupil revenues between the richest and poorest districts) leads to a 5.6 percentile increase in the income rank of children with parental income in the 10th percentile.

2SLS estimates are about 50% larger than OLS.

Simulated instruments: Example 2

Johnson and Tanner (2018) School Finance Reform in California

- California's Local Control Funding Formula (LCFF) was adopted in 2013, targeted revenues to districts based on the % disadvantaged in the district
- How did the funding change affect student outcomes?
- In contrast to Biasi (2019), only one state/reform, but expected variation in "treatment"

Uses school-by-cohort data, 1995-2016.

Simulated instruments: Example 2

The regression of interest is:

$$Y^{st} = \beta_0 + \beta_1 PPE_{st} + u_{st}$$

for a school s in year t.

Simulated instruments: Example 2

The regression of interest is:

$$Y^{st} = \beta_0 + \beta_1 PPE_{st} + u_{st}$$

for a school s in year t.

Problem: PPE_{st} is surely correlated with other time-varying determinants of Y_{st}: parent choices, district SES, residential location constraints (zoning, affordable housing), compensatory spending

Simulated instruments: Example 2

The regression of interest is:

$$Y^{st} = \beta_0 + \beta_1 PPE_{st} + u_{st}$$

for a school s in year t.

Problem: PPE_{st} is surely correlated with other time-varying determinants of Y_{st}: parent choices, district SES, residential location constraints (zoning, affordable housing), compensatory spending

Solution: isolate the change in PPE_{st} that comes about solely via the LCFF.

Simulated instruments: Example 2

Simulate the additional revenues expected from the LCFF, given the funding formula and *baseline* district characteristics (e.g., percent high-need students). This serves as an instrument for the actual revenue in post-LCFF years.

Are these valid instruments? Conditional on baseline disadvantage, the additional revenue received is determined entirely by the formula, not any other omitted variables.

Simulated instruments: Example 3

Jackson, Johnson, and Persico (2015) School Finance Reforms and Education/Economic Outcomes

- What is the impact of school finance equalization on long-run educational outcomes (attainment, earnings)?