Bartik instruments

Vasily Rusanov

January 29, 2019

Plan

- General setup and definition (using Borusyak et al. notation)
- Some examples
- Growth rates as instruments (Borusyak et al.)
- Industry shares as instruments (Goldsmith-Pinkham et al.)
- Conclusion

General setup

- $y_l = \beta x_l + \omega_l' \gamma + \varepsilon$, possibly a panel
- x_l is the endogenous regressor of interest, scalar at "locations" l

General setup

- $y_l = \beta x_l + \omega_l' \gamma + \varepsilon$, possibly a panel
- x_l is the endogenous regressor of interest, scalar at "locations" l
- Suppose we can decompose $x_l = \sum_{n=1}^{N} s_{ln} g_{ln}$, over "industries" n
- If $g_{ln} = g_n + \tilde{g}_{ln}$, then intuitive instrument: $z_l = \sum_n s_{ln} g_n$.
- Often not feasible: g_n is not observed. Use Leave-One-Out estimators.
- Decomposition $x_l = \sum_n s_{ln} g_{ln}$ is nice for intuition, but not necessary. Bartik is whenever the instrument is a weighted average: $z_l = \sum_n s_{ln} g_n$

Simple example: Acemoglu & Linn, Market Size in Innovation

- $y_l = \beta x_l + \omega_l' \gamma + \varepsilon$, instrument $z_l = \sum_n s_{ln} g_n$.

Simple example: Acemoglu & Linn, Market Size in Innovation

- $y_l = \beta x_l + \omega_l' \gamma + \varepsilon$, instrument $z_l = \sum_n s_{ln} g_n$.
- y_l: Number of new drugs approved in category l
  x_l: growth in consumption of drugs in l in the US
- Decompose $x_l = \sum_{n=1}^{N} s_{ln} g_{ln}$
  s_{ln} is the share of spending on drug in category l by age group n
  g_{ln} is the growth in spending of age group n on drug category l.

Simple example: Acemoglu & Linn, Market Size in Innovation

- $y_l = \beta x_l + \omega_l' \gamma + \varepsilon$, instrument $z_l = \sum_n s_{ln} g_n$.
- y_l: Number of new drugs approved in category l
  x_l: growth in consumption of drugs in l in the US
- Decompose $x_l = \sum_{n=1}^{N} s_{ln} g_{ln}$
  s_{ln} is the share of spending on drug in category l by age group n
  g_{ln} is the growth in spending of age group n on drug category l.
- Can decompose $g_{ln} = g_n + g_{ln}$, where g_n is simply the growth of population in group n.
- Plain Bartik is often not feasible as: g_n is not observed. Here, $g_l n$ is not observed either, but it's ok!

Famous example: Autor Dorn Hansen, China shock

- $y_l = \beta x_l + \omega_l' \gamma + \varepsilon$, instrument $z_l = \sum_n s_{ln} g_n$.
- y_l: % change in manufacturing employment
  x_l: % change in import exposure (employment-weighted change in US imports)

Famous example: Autor Dorn Hansen, China shock

- $y_l = \beta x_l + \omega_l' \gamma + \varepsilon$, instrument $z_l = \sum_n s_{ln} g_n$.
- y_l: % change in manufacturing employment
  x_l: % change in import exposure (employment-weighted change in US imports)
- By definition, $x_{lt} = \sum_{n=1}^{N} s_{lnt} g_{nt}$, where
  s_{lnt} is the share of people in state l who work in industry n at time t
  g_{nt} is the growth of imports to the US (not location specific)

Famous example: Autor Dorn Hansen, China shock

- $y_l = \beta x_l + \omega_l' \gamma + \varepsilon$, instrument $z_l = \sum_n s_{ln} g_n$.
- y_l: % change in manufacturing employment
  x_l: % change in import exposure (employment-weighted change in US imports)
- By definition, $x_{lt} = \sum_{n=1}^{N} s_{lnt} g_{nt}$, where
  s_{lnt} is the share of people in state l who work in industry n at time t
  g_{nt} is the growth of imports to the US (not location specific)
- Instrument x_l with $z_{lt} = \sum_{n=1}^{N} s_{lnt-1} \hat{g}_{nt}$, where $\hat{g}_{nt}$ is the growth of Chinese imports in industry n to 8 *other* countries.
- Note that $\sum_n s_{lnt} \neq 1$ because some people work in non-manufacturing (but still tradeables).

How to understand Bartik

- It's very broad; any instrument that has a structure $z_l = \sum_n s_{ln} g_n$ is Bartik

How to understand Bartik

- It's very broad; any instrument that has a structure $z_l = \sum_n s_{ln} g_n$ is Bartik
- We have s_{ln} and g_n. The assumptions required can be roughly classified into "shares are exogenous" or "growth rates" are exogenous.

How to understand Bartik

- It's very broad; any instrument that has a structure $z_l = \sum_n s_{ln} g_n$ is Bartik
- We have s_{ln} and g_n. The assumptions required can be roughly classified into "shares are exogenous" or "growth rates" are exogenous.
- Turns out, using $z_l = \sum_n s_{ln} g_n$ (a scalar) is numerically equivalent to using a vector g_n *or* a vector of $s_l n$ in weighted regressions, which motivates two sets of restrictions.

Borusyak Hull, Jaravel

- To tackle the controls, denote the matrix of k controls as D, and $M_D = I_{L\times k} - D(D'D)D$, and residualized outcomes as $Y^{\perp} = M_D Y$

- To tackle the controls, denote the matrix of k controls as D, and $M_D = I_{L\times k} - D(D'D)D$, and residualized outcomes as $Y^\perp = M_D Y$
- The Bartik IV estimator is $\hat{\beta} = \frac{\frac{1}{L}\sum_l z_l y_l^\perp}{\frac{1}{L}\sum_l z_l x_l^\perp}$

Borusyak Hull, Jaravel

- To tackle the controls, denote the matrix of k controls as D, and $M_D = I_{L\times k} - D(D'D)D$, and residualized outcomes as $Y^{\perp} = M_D Y$
- The Bartik IV estimator is $\hat{\beta} = \frac{\frac{1}{L}\sum_l z_l y_l^{\perp}}{\frac{1}{L}\sum_l z_l x_l^{\perp}}$
- $\hat{\beta}_{Bartik} = \frac{\frac{1}{L}\sum_l z_l y_l^{\perp}}{\frac{1}{L}\sum_l z_l x_l^{\perp}} = \frac{\sum_n g_n \left(\frac{1}{L}\sum_l s_{ln} y_l^{\perp}\right)}{\sum_n g_n \left(\frac{1}{L}\sum_l s_{ln} x_l^{\perp}\right)}$
- $\frac{1}{L}\sum_l s_{ln} y_l^{\perp}$ is the industy's n average residualized outcome, weighted by its importance in all locations l. Not a simple average.

Borusyak Hull, Jaravel

- To tackle the controls, denote the matrix of k controls as D, and $M_D = I_{L\times k} - D(D'D)D$, and residualized outcomes as $Y^{\perp} = M_D Y$
- The Bartik IV estimator is $\hat{\beta} = \dfrac{\frac{1}{L}\sum_l z_l y_l^{\perp}}{\frac{1}{L}\sum_l z_l x_l^{\perp}}$
- $\hat{\beta}_{Bartik} = \dfrac{\frac{1}{L}\sum_l z_l y_l^{\perp}}{\frac{1}{L}\sum_l z_l x_l^{\perp}} = \dfrac{\sum_n g_n \left(\frac{1}{L}\sum_l s_{ln} y_l^{\perp}\right)}{\sum_n g_n \left(\frac{1}{L}\sum_l s_{ln} x_l^{\perp}\right)}$
- $\frac{1}{L}\sum_l s_{ln} y_l^{\perp}$ is the industy's n average residualized outcome, weighted by its importance in all locations l. Not a simple average.
- Often $\sum_n s_{ln} = 1$. We usually don't think that $\sum_l s_{ln} = 1$.

Borusyak Hull, Jaravel

- To tackle the controls, denote the matrix of k controls as D, and $M_D = I_{L\times k} - D(D'D)D$, and residualized outcomes as $Y^{\perp} = M_D Y$
- The Bartik IV estimator is $\hat{\beta} = \frac{\frac{1}{L}\sum_l z_l y_l^{\perp}}{\frac{1}{L}\sum_l z_l x_l^{\perp}}$
- $\hat{\beta}_{Bartik} = \frac{\frac{1}{L}\sum_l z_l y_l^{\perp}}{\frac{1}{L}\sum_l z_l x_l^{\perp}} = \frac{\sum_n g_n \left(\frac{1}{L}\sum_l s_{ln} y_l^{\perp}\right)}{\sum_n g_n \left(\frac{1}{L}\sum_l s_{ln} x_l^{\perp}\right)}$
- $\frac{1}{L}\sum_l s_{ln} y_l^{\perp}$ is the industy's n average residualized outcome, weighted by its importance in all locations l. Not a simple average.
- Often $\sum_n s_{ln} = 1$. We usually don't think that $\sum_l s_{ln} = 1$.
- Bartik is num. equivalent to $\bar{y}_n^{\perp} = \beta \bar{x}_n^{\perp} + \varepsilon^{\perp}$, where $\bar{y}_n = \sum_l s_{ln} y_l / \sum_l s_{ln}$ and the regression is weighted by $\hat{s}_n$

Consistency when growth rates are instruments

- Under shocks-as-instruments approach, $\hat{\beta}_{Bartik}$ is consistent iff $\sum_n s_n g_n \phi_n \to 0$, where $s_n = \mathbb{E}(s_{ln})$ and $\phi_n = \mathbb{E}(s_{ln}\varepsilon_l)/\mathbb{E}(s_{ln})$. I don't understand it.
- The authors show that this holds under two assumptions:

Consistency when growth rates are instruments

- Under shocks-as-instruments approach, $\hat{\beta}_{Bartik}$ is consistent iff $\sum_n s_n g_n \phi_n \to 0$, where $s_n = \mathbb{E}(s_{ln})$ and $\phi_n = \mathbb{E}(s_{ln}\varepsilon_l)/\mathbb{E}(s_{ln})$.
  I don't understand it.
- The authors show that this holds under two assumptions:

A1 $\mathbb{E}(g_n|\phi_n) = \mu$, $\forall n$

A2 1) $\mathbb{E}[(g_n - \mu)(g_m - \mu)|\phi_n, \phi_m] = 0$, $\forall n$
2) $\sum_n s_n^2 \xrightarrow[L,N\to\infty]{} 0$

(shocks are random and no industry is too important)

Consistency when growth rates are instruments

- Under shocks-as-instruments approach, $\hat{\beta}_{Bartik}$ is consistent iff $\sum_n s_n g_n \phi_n \to 0$, where $s_n = \mathbb{E}(s_{ln})$ and $\phi_n = \mathbb{E}(s_{ln}\varepsilon_l)/\mathbb{E}(s_{ln})$.
  I don't understand it.
- The authors show that this holds under two assumptions:

A1 $\mathbb{E}(g_n|\phi_n) = \mu,\ \forall n$

A2 1) $\mathbb{E}[(g_n - \mu)(g_m - \mu)|\phi_n, \phi_m] = 0,\ \forall n$
2) $\sum_n s_n^2 \xrightarrow[L,N\to\infty]{} 0$

(shocks are random and no industry is too important)

- This can be modified to have clusters with random assignment within clusters and to the case where $\sum_n s_{ln} \neq 1$
- Similar for panel data.

Goldsmith-Pinkham, Sorkin, Swift

- $y_l = \beta x_l + \omega_l' \gamma + \varepsilon$, instrument $z_l = \sum_n s_{ln} g_n$.

Goldsmith-Pinkham, Sorkin, Swift

- $y_l = \beta x_l + \omega_l'\gamma + \varepsilon$, instrument $z_l = \sum_n s_{ln} g_n$.
- Vectorize $\underbrace{S}_{L\times N}$: matrix of shares, $\underbrace{G}_{L\times N}$: vector of growth rates, $\underbrace{Y}_{N\times 1}$: outcomes. Then $\underbrace{B}_{L\times 1} = ZG$ is the Bartik instrument.

- $y_l = \beta x_l + \omega_l' \gamma + \varepsilon$, instrument $z_l = \sum_n s_{ln} g_n$.
- Vectorize $\underbrace{S}_{L \times N}$: matrix of shares, $\underbrace{G}_{L \times N}$: vector of growth rates, $\underbrace{Y}_{N \times 1}$: outcomes. Then $\underbrace{B}_{L \times 1} = ZG$ is the Bartik instrument.
- We have $\hat{\beta}_{Bartik} = \dfrac{B'Y^{\perp}}{B'X^{\perp}} = \dfrac{G'Z'Y^{\perp}}{G'Z'X^{\perp}} = \dfrac{\overbrace{X^{\perp\prime} ZG}^{scalar} G'Z'Y^{\perp}}{X^{\perp\prime} ZGG'Z'X^{\perp}}$
- Using Barik is the same as running GMM with weight matrix $W = GG'$ and industry shares as instruments!

Goldsmith-Pinkham, Sorkin, Swift

- This can be extended to a panel, but notation is long
- They provide several consistency conditions similar to Borusyak et al., but this doesn't give much intuition
- Key contribution: the GMM estimator gives way to Rotemberg weights.

Rotemberg weights

- Let $\hat{C}(\hat{W}) = \hat{W} Z' X^{\perp}$ and $\hat{c}_n(\hat{W}) = \hat{W}_n Z' X^{\perp}$

Rotemberg weights

- Let $\hat{C}(\hat{W}) = \hat{W} Z' X^{\perp}$ and $\hat{c}_n(\hat{W}) = \hat{W}_n Z' X^{\perp}$
- Let $\hat{\beta}(\hat{W}) = \dfrac{\hat{C}(\hat{W})' Z' Y^{\perp}}{\hat{C}(\hat{W})' Z' X^{\perp}}$, $\hat{\alpha}_k(\hat{W}) = \dfrac{\hat{c}_k(\hat{W})' Z_k' X^{\perp}}{\hat{c}_k(\hat{W})' Z_k' X^{\perp}}$ and $\hat{\beta}_k = (Z_k' X^{\perp})^{-1} Z_k' Y^{\perp}$

Rotemberg weights

- Let $\hat{C}(\hat{W}) = \hat{W}Z'X^{\perp}$ and $\hat{c}_n(\hat{W}) = \hat{W}_n Z'X^{\perp}$
- Let $\hat{\beta}(\hat{W}) = \dfrac{\hat{C}(\hat{W})'Z'Y^{\perp}}{\hat{C}(\hat{W})'Z'X^{\perp}}$, $\hat{\alpha}_k(\hat{W}) = \dfrac{\hat{c}_k(\hat{W})'Z_k'X^{\perp}}{\hat{c}_k(\hat{W})'Z_k'X^{\perp}}$ and $\hat{\beta}_k = (Z_k'X^{\perp})^{-1}Z_k'Y^{\perp}$
- Then $\hat{\beta}(\hat{W}) = \sum_{n=1}^{N} \hat{\alpha}_k(\hat{W})\hat{\beta}_k$

Rotemberg weights

- Let $\hat{C}(\hat{W}) = \hat{W}Z'X^{\perp}$ and $\hat{c}_n(\hat{W}) = \hat{W}_n Z'X^{\perp}$
- Let $\hat{\beta}(\hat{W}) = \dfrac{\hat{C}(\hat{W})'Z'Y^{\perp}}{\hat{C}(\hat{W})'Z'X^{\perp}}$, $\hat{\alpha}_k(\hat{W}) = \dfrac{\hat{c}_k(\hat{W})'Z_k'X^{\perp}}{\hat{c}_k(\hat{W})'Z_k'X^{\perp}}$ and $\hat{\beta}_k = (Z_k'X^{\perp})^{-1}Z_k'Y^{\perp}$
- Then $\hat{\beta}(\hat{W}) = \sum_{n=1}^{N} \hat{\alpha}_k(\hat{W})\hat{\beta}_k$
- For the Bartik instruments, the weights are $\hat{\alpha}_k(\hat{W}) = \dfrac{\hat{X}_k^{Bartik}X^{\perp}}{\hat{X}^{Bartik}X^{\perp}}$ where $\hat{X}^{Bartik}$ is the residualized X after the first stage.
- This is a way to see whether some industries are more important than others. It also allows to estimate the bias.

Conclusion

- Little “theory” is needed to use Bartik.
  It’s an empirical tool, not a modeling technique.

Conclusion

- Little "theory" is needed to use Bartik.
  It's an empirical tool, not a modeling technique.
- You need those two papers to justify your instrument, not to define it.
  They way they define Bartik is the same.
- It's bad if one industry n gives a significant part of variation (under both approaches).