Inference in Shift-Share Designs

Michal Kolesár

2019-01-11

The package `BartikSE` implements confidence intervals proposed by Adão, Kolesár, and Morales (2018) for inference in shift-share least squares and instrumental variables regressions, in which the regressor of interest (or the instrument) has a shift-share structure, as in Bartik (1991). A shift-share variable has the structure $X_i = \sum_{s=1}^{S} w_{is}\mathcal{X}_s$, where i indexes regions, s indexes sectors, $\mathcal{X}_s$ are sectoral shifters (or shocks), and w_{is} are shares, such as initial share of region i's employment in sector s.

This vignette illustrates the use of the package using a dataset from Autor, Dorn, and Hanson (2013) (ADH hereafter). The dataset is included in the package as the list `ADH`. The first element of the list, `ADH$reg` is a data-frame with regional variables, the second element, `ADH$sic` is a vector of SIC codes for the sectors, and `ADH$W` is a matrix of shares. See `?ADH` for a description of the dataset.

We now replicate the first row of Table 6 in Adão, Kolesár, and Morales (2018). First we load the package, define the vector of controls, and define a vector of 3-digit SIC codes:

```
library("BartikSE")
ctrls <- "t2 + l_shind_manuf_cbp + l_sh_popedu_c + l_sh_popfborn +
          l_sh_empl_f + l_sh_routine33 + l_task_outsource + division"
sic <- floor(ADH$sic/10)
```

We cluster the standard errors at the 3-digit SIC code (using the option `sector_cvar`), and, following ADH, weight the data using the weights `ADH$reg$weights`. See `?lmBartik` and `?IVBartik` for full description of the options.

The first-stage regression:

```
lmBartik(as.formula(paste("shock ~ ", ctrls)), W = ADH$W,
    X = IV, data = ADH$reg, weights = weights, region_cvar = statefip,
    sector_cvar = sic, method = "all", residual_sector = TRUE)
#> Estimate: 0.6310409
#>
#> Inference:
#>                Std. Error      p-value  Lower CI  Upper CI
#> Homoscedastic 0.02732516 0.000000e+00 0.5774846 0.6845973
#> EHW           0.08700719 4.083400e-13 0.4605100 0.8015719
#> Reg. cluster  0.09142372 5.113909e-12 0.4518537 0.8102281
#> AKM           0.05286498 0.000000e+00 0.5274275 0.7346544
#> AKM0          0.07659602 1.310031e-03 0.5375408 0.8377916
```

The reduced-form and IV regressions:

```
lmBartik(as.formula(paste("d_sh_empl ~", ctrls)), W = ADH$W,
    X = IV, data = ADH$reg, region_cvar = statefip, weights = weights,
    sector_cvar = sic, method = "all", residual_sector = TRUE)
#> Estimate: -0.4885687
#>
#> Inference:
#>                Std. Error      p-value   Lower CI   Upper CI
#> Homoscedastic 0.06332778 1.221245e-14 -0.6126889 -0.3644485
#> EHW           0.11244360 1.392685e-05 -0.7089541 -0.2681833
#> Reg. cluster  0.07578147 1.140306e-10 -0.6370977 -0.3400398
```

```
#> AKM           0.16585182 3.221054e-03 -0.8136323 -0.1635051
#> AKM0          0.25797650 3.952061e-04 -1.2499065 -0.2386572
ivBartik(as.formula(paste("d_sh_empl ~", ctrls, "| shock")),
    W = ADH$W, X = IV, data = ADH$reg, region_cvar = statefip,
    weights = weights, sector_cvar = sic, method = "all",
    residual_sector = TRUE)
#> Estimate: -0.7742267
#>
#> Inference:
#>                Std. Error      p-value   Lower CI   Upper CI
#> Homoscedastic  0.1069532 4.523049e-13 -0.9838511 -0.5646022
#> EHW            0.1647892 2.623532e-06 -1.0972075 -0.4512459
#> Reg. cluster   0.1758096 1.063809e-05 -1.1188071 -0.4296462
#> AKM            0.2447514 1.559814e-03 -1.2539306 -0.2945228
#> AKM0           0.3396909 3.952061e-04 -1.7172769 -0.3857130
```

References

Adão, Rodrigo, Michal Kolesár, and Eduardo Morales. 2018. "Inference in Shift-Share Designs: Theory and Inference." https://arxiv.org/abs/1806.07928.

Autor, David H., David Dorn, and Gordon H. Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review* 103 (6): 2121–68.

Bartik, Timothy J. 1991. *Who Benefits from State and Local Economic Development Policies?* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.