Inference in Shift-Share Designs

Michal Kolesár

2018-08-06

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 teh 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: 0.631041
#>
#> Inference:
#>                Std. Error     p-value Lower CI Upper CI
#> Homoscedastic  0.0273252 0.00000e+00 0.577485 0.684597
#> EHW            0.0870072 4.08340e-13 0.460510 0.801572
#> Reg. cluster   0.0914237 5.11391e-12 0.451854 0.810228
#> AKM            0.0528650 0.00000e+00 0.527427 0.734654
#> AKM0           0.0765960 1.31003e-03 0.537541 0.837792
```

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: -0.488569
#>
#> Inference:
#>                Std. Error     p-value  Lower CI  Upper CI
#> Homoscedastic  0.0633278 1.22125e-14 -0.612689 -0.364449
#> EHW            0.1124436 1.39268e-05 -0.708954 -0.268183
#> Reg. cluster   0.0757815 1.14031e-10 -0.637098 -0.340040
```

```
#> AKM            0.1658518 3.22105e-03 -0.813632 -0.163505
#> AKM0           0.2579765 3.95206e-04 -1.249907 -0.238657
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: -0.774227
#>
#> Inference:
#>                Std. Error     p-value  Lower CI  Upper CI
#> Homoscedastic   0.106953 4.52305e-13 -0.983851 -0.564602
#> EHW             0.164789 2.62353e-06 -1.097207 -0.451246
#> Reg. cluster    0.175810 1.06381e-05 -1.118807 -0.429646
#> AKM             0.244751 1.55981e-03 -1.253931 -0.294523
#> AKM0            0.339691 3.95206e-04 -1.717277 -0.385713
```

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.