Making Instruments from Known Exogeneity

Joel Ferguson

11/15/2021

Non-random Exposure to Exogenous Shocks

The IV approach is predicated on knowing of some exogenous variation that is predictive of treatment, but is not systematically correlated with unobservable determinants of the outcome. Usually, we refer to this second condition as the "exclusion restriction" and it is sometimes stated as "the instrument only affects the outcome through its effect on treatment." I don't particularly like this expression because it glosses over the fact that the instrument *cannot* even be correlated with unobservable determinants. In other words, we may think that the only causal relationship between the instrument and the outcome passes through treatment, but if there are other determinants of the outcome that covary with the instrument, even if not in a causal way, the IV estimator will be inconsistent.

In a paper that I think could revolutionize the way we do IV, Borusyak and Hull (2020) point out that there is a wide class of situations where this type of problem arises with IV: non-random exposure to exogenous shocks. The idea is fairly straight-forward: often we know about some sort of variation that we believe to be exogenous, but the extent to which this variation matters for different units is not exogenous. A classic example is the shift-share instrument of Bartik (1991), in which non-random industry composition *shares* determine exposure to exogenous *shifts* in a variable such as international prices.

Borusyak and Hull show that if we take this non-randomness seriously and avoid assuming an IID sampling framework to allow for the complicated error dependencies that almost certainly arise in these settings, we get

$$\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{n}z_i\varepsilon_i\right] = \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{n}\mu_i\varepsilon_i\right]$$

where μ_i is the expected value of the instrument for unit i over the distribution of potential shock realizations. This is typically not equal to 0 because certain units may tend to be more or less exposed, and if the determinants of exposure are correlated with outcomes it would generate correlation between the expected instrument and the population error term; i.e. an exclustion restriction violation.

Fortunately, Borusyak and Hull show that there's a simple solution to this issue: we can **recenter** the instrument around μ_i for each observation, which mechanically resolves the (likely) exclusion restriction violation. In practice, this means that we need to know the distribution of potential shocks. So in the shift-share example above, we would need to know the distribution of potential price realizations. Importantly, using fixed effects is typically not enough; the set of fixed effects would have to span the variation in the μ_is.

Some shocks lend themselves more naturally to this sort of assumption than others. For international price shocks like in the shift share example, we might assume prices are some sort of unit-root process[1]. Dennis Egger, an econ PhD candidate on the job market this year, has a particularly clean implementation in his JMP, in which he utilizes the randomness of the placement of refugees in Switzerland together with refugee flows to find exogenous variation in the share of co-nationals in refugees' placements.

---

[1]This is the assumption in e.g. McGuirk and Burke (2021).

In the simulation below, I demonstrate the inconsistency introduced by non-random exposure and how the Borusyak and Hull methodology resolves it.

```r
N <- 3000 # Number of units
K <- 10 # Number of things determining exposure
G <- matrix(runif(N*K),
            nrow=N)

# Normalize G to be row of shares for each observation
G_sums <- sapply(c(1:N), function(x) sum(G[x,]))
G <- G/G_sums

# Generate exogenous variation
w_means <- runif(K,0,10) # Means vary by k in K
w <- matrix(sapply(w_means, function(x) rnorm(1,x,5)),nrow=K)

Z <- G%*%w
mu <- G%*%w_means
Z_rc <- Z-mu

# Choose unit-specific error means
mu <- as.vector(mu)
e1 <- sapply(mu,function(x) rnorm(1,x,0.5))
e2 <- sapply(e1,function(x) rnorm(1,x,0.01))

# Make treatment
D_weights <- matrix(runif(N*K),
            nrow=N)
D <- Z+e1

# Make Outcomes
Y <- D+e2 #True treatment effect=1

# Make a df
df <- tibble(
  "Y"=Y,
  "D"=D,
  "Z"=as.vector(Z),
  "Z_rc"=as.vector(Z_rc)
)

# OLS
feols(Y~D,
      data=df)
```

```
## OLS estimation, Dep. Var.: Y
## Observations: 3,000
## Standard-errors: IID
##              Estimate Std. Error  t value  Pr(>|t|)
## (Intercept)   1.49589   0.069125  21.6402 < 2.2e-16 ***
## D             1.37677   0.006291 218.8473 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 0.443309   Adj. R2: 0.941071
```

```
# Ordinary IV
feols(Y~1|
      D~Z,
      data=df)
```

```
## TSLS estimation, Dep. Var.: Y, Endo.: D, Instr.: Z
## Second stage: Dep. Var.: Y
## Observations: 3,000
## Standard-errors: IID
##              Estimate Std. Error  t value  Pr(>|t|)
## (Intercept)  3.57312    0.090096  39.6590 < 2.2e-16 ***
## fit_D        1.18641    0.008213 144.4598 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 0.506498   Adj. R2: 0.923075
## F-test (1st stage), D: stat =      9,811.8, p < 2.2e-16, on 1 and 2,998 DoF.
##              Wu-Hausman: stat = 5,961,432.3, p < 2.2e-16, on 1 and 2,997 DoF.
```

```
# Borusyak and Hull IV
feols(Y~1|
      D~Z_rc,
      data=df)
```

```
## TSLS estimation, Dep. Var.: Y, Endo.: D, Instr.: Z_rc
## Second stage: Dep. Var.: Y
## Observations: 3,000
## Standard-errors: IID
##              Estimate Std. Error t value  Pr(>|t|)
## (Intercept)  5.59722    0.162605 34.4221 < 2.2e-16 ***
## fit_D        1.00092    0.014861 67.3543 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 0.656119   Adj. R2: 0.870914
## F-test (1st stage), D: stat =  1,937.6, p < 2.2e-16, on 1 and 2,998 DoF.
##              Wu-Hausman: stat = 10,001.4, p < 2.2e-16, on 1 and 2,997 DoF.
```

As you can see, in the simulation above OLS is inconsistent because D_i is correlated with the error term in the relationship of interest. And because the way that shocks are aggregated to make D is correlated with this error term as well, traditional IV doesn't work despite the instrument just being an aggregate of exogenous shocks! However, once we recenter the instrument, we recover the true treatment effect.

Another nice aspect of the Borusyak and Hull methodology is that because we've already specified the shock distribution, we can perform randomization inference to obtain valid confidence intervals. The idea behind this is also fairly straight forward: we're simply asking how "likely is it that we would obtain such a large covariance between the (re-centered) instrument and the outcome under an alternative shock realization?" Code below shows how to actually implement this.

```
iterations <- 1000
covs <- rep(0,iterations)
for (i in c(1:iterations)){
  w_alt <- matrix(sapply(w_means, function(x) rnorm(1,x)),nrow=K)
  Z_alt <- G%*%w_alt
  Z_rc_alt <- Z_alt-mu
```
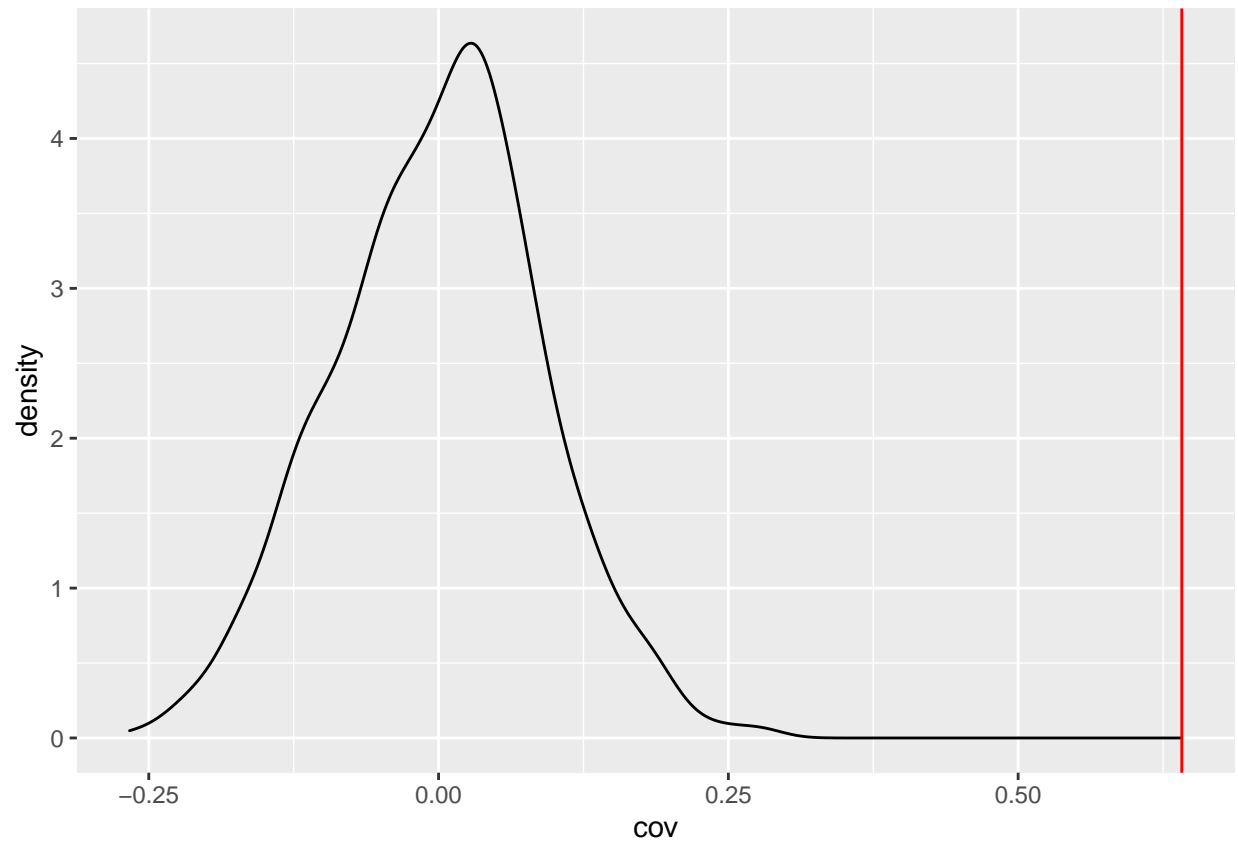
```
  covs[i] <- cov(Z_rc_alt,Y)
}

true_cov <- cov(Z_rc,Y)[1,1]

RI_df <- tibble("cov"=covs)
ggplot(data=RI_df)+
  geom_density(aes(x=cov))+
  geom_vline(aes(xintercept=true_cov),color="red")
```



```
print(paste("Randomization Inference p-value:",length(covs[abs(covs)>true_cov])/length(covs)))
```

```
## [1] "Randomization Inference p-value: 0"
```