IVCV vs TSLS Bias

Srikar Katta, Joe Rilling, Bear Zhang

4/15/2022

Data Generating Process

- Set n_{per} – the number of observations per treatment group
- Assign each of our N observations to K treatment groups, each group with n_{per} observations
  - For $i = 1, \ldots, N$
    - Let i_k represent unit i's treatment group
    - $U_i \sim \mathcal{N}(1, 25)$, the unobserved confounder
    - $\begin{pmatrix} X_{i,1} \\ X_{i,2} \end{pmatrix} \sim \mathcal{MN}\left(\begin{pmatrix} \gamma_k + 5U_i \\ \gamma_k \end{pmatrix}, \begin{pmatrix} 1 & 0.64 \\ 0.64 & 1 \end{pmatrix}\right)$, where γ_k is a group-level fixed effect
    - $Y_i = X_{i,1} + X_{i,2} + 5 \cdot U_i + \varepsilon_i, \varepsilon_i \sim \mathcal{N}(0, 1)$

Suppose we have P experiments, each with two treatment assignments (i.e., treated or control). Then, we would have $K = 2^P$ treatment groups. We assume that γ_k is the number of experiments that a unit belongs to the treatment group for. In other words, being assigned to the treatment group for an additional experiment would increase $\mathbb{E}[X_{i,1}]$ and $\mathbb{E}[X_{i,2}]$ by 1. Suppose unit i is treated in 10 experiments; then the fixed effect for unit i's treatment group is 10: $\gamma_{i_k} = 10$.

Additionally, X_1 is the endogeneous feature whose causal effect we are interested in. The treatment effect is 1, i.e., a 1 unit increase in X_1 will cause a 1 unit increase in Y, in expectation.

Experiments

We considered two different experimental setups: one in which we set n_{per} and increase K by increasing the number of experiments; and another in which we fix K and increase n_{per}. We consider $n_{per} \in \{1000, 2000, \ldots, 10000\}$ and $K \in \{2^1, 2^2, , \ldots, 2^{10}\}$. For each combination of n_{per} and K, we compare the results using the Instrumental Variables Cross Validation (IVCV) procedure and traditional Two Stage Least Squares (TSLS) using the squared error:

$$SE = \left(\hat{\beta} - 1\right)^2.$$

Results

Figures 1 and 2 display the bias of the TSLS and IVCV procedure for varying observations per group, n_{per}, and number of experiments, P. Please note that even though there are only 10 experiments, $K = 2^{10} = 1024$, so there are 1024 different treatment groups. When $n_{per} = 10000$, there are $1024 \cdot 10000 = 10240000$ total observations. If either TSLS or IVCV are asymptotically unbiased, then we should see the squared error of the procedure's tending to 0.

Interestingly, the IVCV bias has a smaller range of 3.25 to 4.25, while the TSLS bias varies from ≈ 0 to over 200. However, as n_{per} and K increase, the TSLS bias tends to 0, which cannot be said for the IVCV estimates: even when $n_{per} = 10000$ and $K = 2^{10} = 1024$, the IVCV estimate still has a bias around 4.

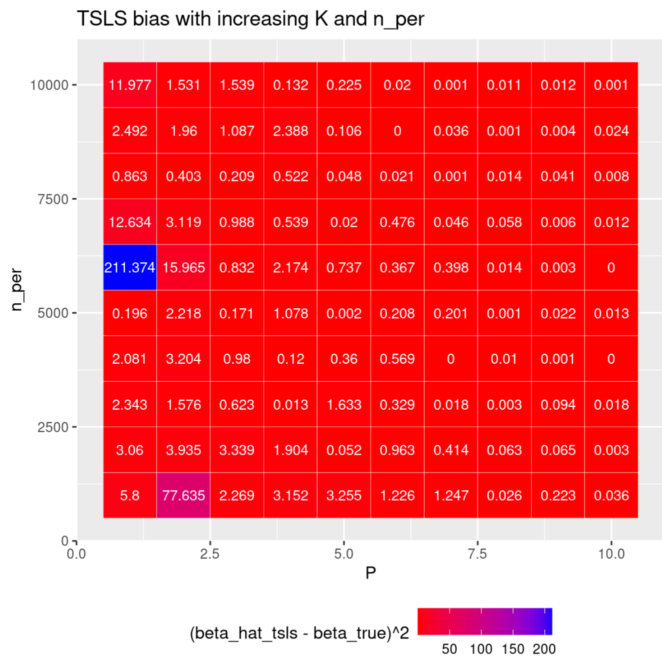

Figure 1: We increase the number of experiments P with binary treatments from 1 to 20 on the x-axis and increase n_per from 1000 to 10000 on the y-axis. For each (P, n_{per}) tuple, we display the squared error of the Two Stage Least Squares estimate, represented by color and shown in text. Brighter red blocks indicate less biased estimates while darker, blue blocks represent more biased estimates.

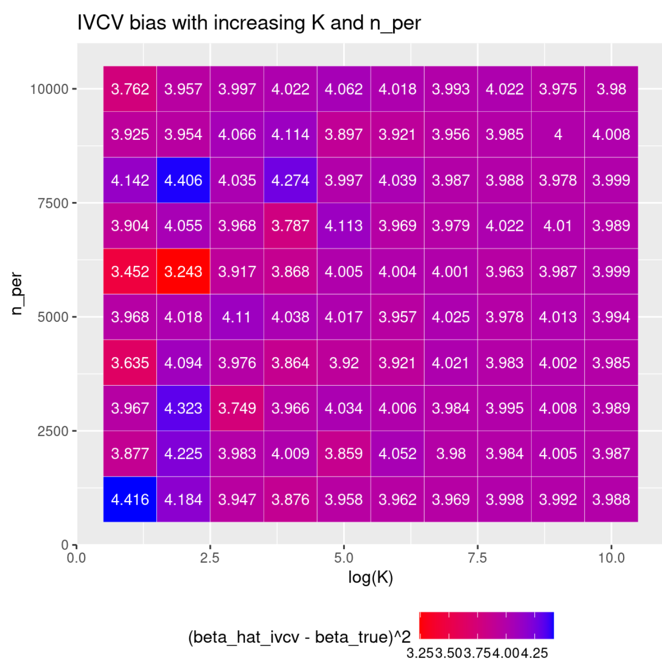

Figure 2: We increase the number of experiments P with binary treatments from 1 to 20 on the x-axis and increase n_per from 1000 to 10000 on the y-axis. For each (P, n_{per}) tuple, we display the squared error of the Instrumental Variables Cross Validation procedure estimate, represented by color and shown in text. Brighter red blocks indicate less biased estimates while darker, blue blocks represent more biased estimates.