# Replication of Empirical Results in 'Finite-Sample Optimal Estimation and Inference on Average Treatment Effects Under Unconfoundedness'

Michal Kolesár

November 10, 2020

## 1 Empirical application: NSW

We use a solution path with 3000 steps. At the last step the value of $\delta$ is given by 668.48739.

Table 2:

| | Opt: RMSE | Opt: FLCI | Matching: RMSE | Matching: FLCI |
|---|---|---|---|---|
| $M$ | | | 1.00 | 18.00 |
| $\delta$ | 1.82 | 3.30 | | |
| Estimate | 0.96 | 0.94 | 1.39 | 1.26 |
| Max. Bias | 1.64 | 1.78 | 1.48 | 2.21 |
| SE | 1.01 | 0.94 | 1.09 | 0.89 |
| cv | 3.26 | 3.55 | 3.01 | 4.13 |
| CI lo | -2.33 | -2.38 | -1.88 | -2.41 |
| CI hi | 4.26 | 4.27 | 4.66 | 4.93 |
| SE (PATE) | 1.06 | 1.00 | 1.14 | 0.94 |
| CI lo (PATE) | -2.75 | -2.80 | -2.32 | -2.78 |
| CI hi (PATE) | 4.67 | 4.69 | 5.11 | 5.30 |
| Lindeberg wgt | 0.07 | 0.06 | 0.19 | 0.06 |
| Rel. efficiency | 101.15 | | 90.41 | 91.36 |

Efficiency of FLCI in baseline specification (Section 2.5)

```
ATEHonest::ATTEffBounds(op, C = 1, sigma2 = mean(sigma2))
## $onesided
## [1] 0.977761
##
## $twosided
## [1] 0.970476
```

Optimal estimator and CIs as a function of $C$. Solid lines mark the CATT CI, while dotted lines mark the PATT CI.

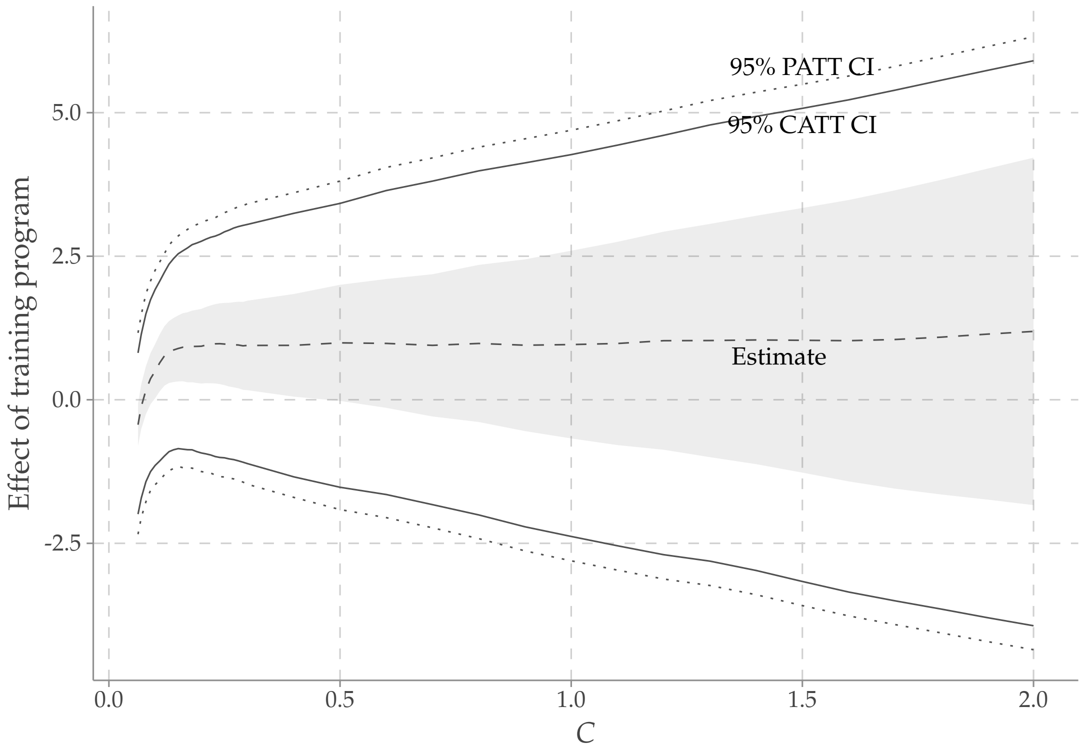


Figure 1: Reproduction of Figure 1 in the paper. Robust SE.

Range of length penalty of PATT relative to CATT in the four columns of Table 2: 1.102, 1.135

Range of point estimate for reasonable $C$:

```
range(oe$rmse$att[oe$rmse$C > 0.2 & oe$rmse$C < 2])
## [1] 0.939249 1.142967
```

## 1.1 Matching efficiency

Efficiency at $C = 1$:

| C | type | eff |
|---|---|---|
| 1 | RMSE | 90.4132 |
| 1 | FLCI | 85.9666 |

Minimum $C$ such that efficiency is at least 95%:

| C | type | eff |
|---|---|---|
| 1.5 | RMSE | 95.6341 |
| 2.8 | FLCI | 95.0211 |

### 1.2 Lindeberg weights

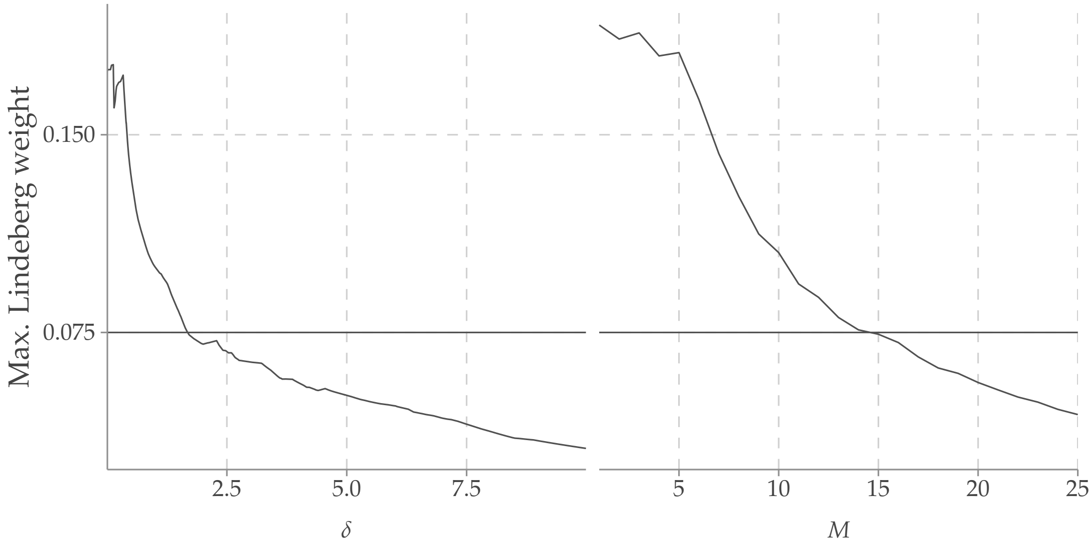


Figure 2: Lindeberg weights for optimal and matching estimator.

For optimal estiamtor, we need $\delta$ to be at least 1.686452. For matching estimator, we need at least 15 matches to be below the 0.075 cutoff.

### 1.3 Experimental data

```
Xe <- as.matrix(ATEHonest::NSWexper[, 2:10])
de <- ATEHonest::NSWexper$treated
ye <- ATEHonest::NSWexper$re78
s2 <- ATEHonest:::nnvar(ATEHonest::distMat(Xe, chol(solve(cov(Xe))),
    method = "euclidean"), de, ye, J = 2)
```

Difference in means estimate:

```
RCTvariance(ye, de, s2, ATEHonest::distMat(Xe, Amain, method = "manhattan",
    de))
##      att  maxbias Cond rob Cond hom CATT rob CATT hom PATE HC2 PATE EHW
## 1.794342 1.737712 0.611274 0.578961 0.641517 0.610806 0.670997 0.669315
## PATE hom
## 0.632853
```

### 1.4 Alternative choice of distance (Supplement)

Now consider the norm in Abadie and Imbens (2011), with $p = 2$, and C described in the text. Optimal FLCI:

```
oe2 <- ATEHonest::ATTOptEstimate(y = y, d = d, D0 = D2, C = 1/sqrt(2 *
    0.07^2), sigma2 = sigma2, opt.criterion = "FLCI")
## Increasing length of solution path to 100
```

```r
oe2
##
##
## |     | Estimate| Max. bias|       SE|CI                  |   delta|
## |:----|--------:|---------:|--------:|:-------------------|-------:|
## |CATT |  1.72259|   6.25009| 0.842248|(-5.91288, 9.35805) | 3.28971|
## |PATT |  1.72259|          | 0.913810|(-6.31854, 9.76371) |        |
oe2$e["rhl"]
##     rhl
## 7.63546
```

Now consider a matching estimator that matches on this norm:

```r
mp2 <- ATEHonest::ATTMatchPath(y, d, D2, M = 1)
## Now calculate worst-case bias under D0
mp2$ep$maxbias <- ATEHonest:::ATTbias(mp2$K[, d == 0], D0)
mp2$D0 <- D0
me2 <- ATEHonest::ATTMatchEstimate(mp2, sigma2 = sigma2, C = 1)
me2$e[c("hl", "rhl")]
##      hl     rhl
## 5.38073 3.37502
```

```r
meff2 <- MatchEfficiency(oe, mp2, mean(sigma2), mineff = 95)
knitr::kable(meff2$effC1, digits = 4, row.names = FALSE)
```

| C | type | eff |
|---|---|---|
| 1 | RMSE | 77.8138 |
| 1 | FLCI | 74.9252 |

```r
## Efficiency loss relative to using correct norm
meff$effC1$eff - meff2$effC1$eff
## [1] 12.5994 11.0414
## Maximum efficiency
max(meff2$d[meff2$d$type == "RMSE", "eff"])
## [1] 80.1513
```

Compilation of this document took: 2.7 mins.