

# Synthetic Differences in Differences Simulations

*Joshua Derenski*

## Simulations

### Parameters

- Number of Ls: 1
- Draws per L: 100
- Number of Units: 100
- Number of Control Units: 95
- Number of Times: 200
- Number of pre-treatment Times: 190
- Rank of L: 10
- Autocorrelation Parameter: 0
- True Effect Size for Constant Effect: 10
- Error Type: gaussian
- Error Variance (if Gaussian error): 16
- Degrees of freedom (if t-error): 5
- Exchangable: FALSE
- Penalized: TRUE
- Rank Estimation Method: threshold
- Scaling for  $L$ : 5
- Treatment Effect Type: constant
- Treatment Design: block\_treatment
- Lag Structure (if using staggered adoption structure): random
- Average Treatment Length (if using staggered adoption structure, with random adoption): 4
- Maximum lag: 4

## Our Method vs Competitors, Fixed Parameters

### Results

#### Signal to Noise Ratio

## [1] 21.79198

mse for DID

```
##      mse
## 18.70946
```

Se for mse for DID

```
##   se_mse
## 0.56006
```

mse for SC

```
##      mse
## 10.98789
```

Se for mse for SC

```
##      se_mse
## 0.3864514
```

mse for our Method (Explicit Tau)

```
##      mse
## 4.734002
```

Se for mse for our Method (Explicit Tau)

```
##      se_mse
## 0.2130038
```

mse for SDID

```
##      mse
## 10.90063
```

Se for mse for SDID

```
##      se_mse
## 0.3811718
```

mse For Our Method (Not Explicit Tau)

```
##      mse
## 4.733999
```

Se for mse for Our Method (Not Explicit Tau)

```
##      se_mse
## 0.2130036
```

mse For Oracle (Perfect L)

```
##      mse
## 3.409822
```

mse For Oracle (Perfect L)

```
##      se_mse
## 0.1552666
```

## Matrix Bias vs Reduction in Variance due to Averaging

For more general designs of  $W$  (like the block design scheme considered here) we allow a block in the bottom right hand corner of  $W$  to be non-zero. When implementing our method, we have two competing effects on estimation:

- The bias that's introduced by making more of the  $Y_{ijs}$  zero.
- The help we get with estimating  $\tau$  by being able to average over cells (because we assume  $\tau$  is the same for all units and times).

It would appear that accuracy increases for estimating  $\tau$  to a point, and then decreases when the bias introduced by replacement of cells with 0 in  $Y$  becomes too great.

Influence of  $N_0/N$  on Performance

Influence of  $\rho$  on Performance

Influence of  $\tau$  on Performance

Influence True Rank on Performance

Influence of Rank Error on Performance

