

# SimResult

Jianhui Gao

7/1/2022

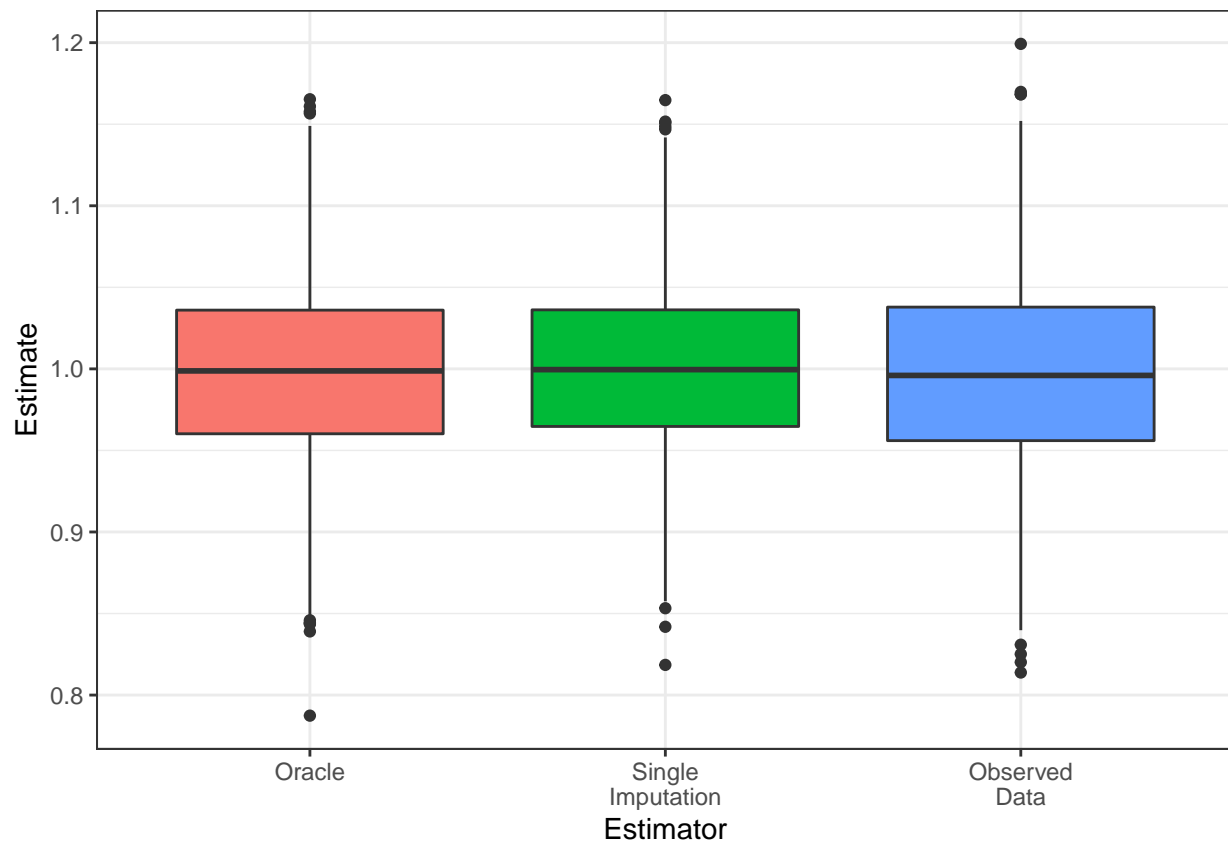
## Simulation Parameter

- sample size  $n = 1000$
- number of replicates = 1000
- $Y = G - 0.5X + 0.5Z + \epsilon$ , where  $G \sim \text{Bin}(2, \text{maf} = 0.25)$ ,  $\begin{bmatrix} X \\ Z \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right)$ ,  $\epsilon \sim N(0, 1)$ .
- Missing rate of  $Y$  is 20%.

## Single Imputation

### Beta known, model correctly specified

```
## Note: Using an external vector in selections is ambiguous.  
## i Use 'all_of(cov)' instead of 'cov' to silence this message.  
## i See <https://tidyselect.r-lib.org/reference/faq-external-vector.html>.  
## This message is displayed once per session.
```

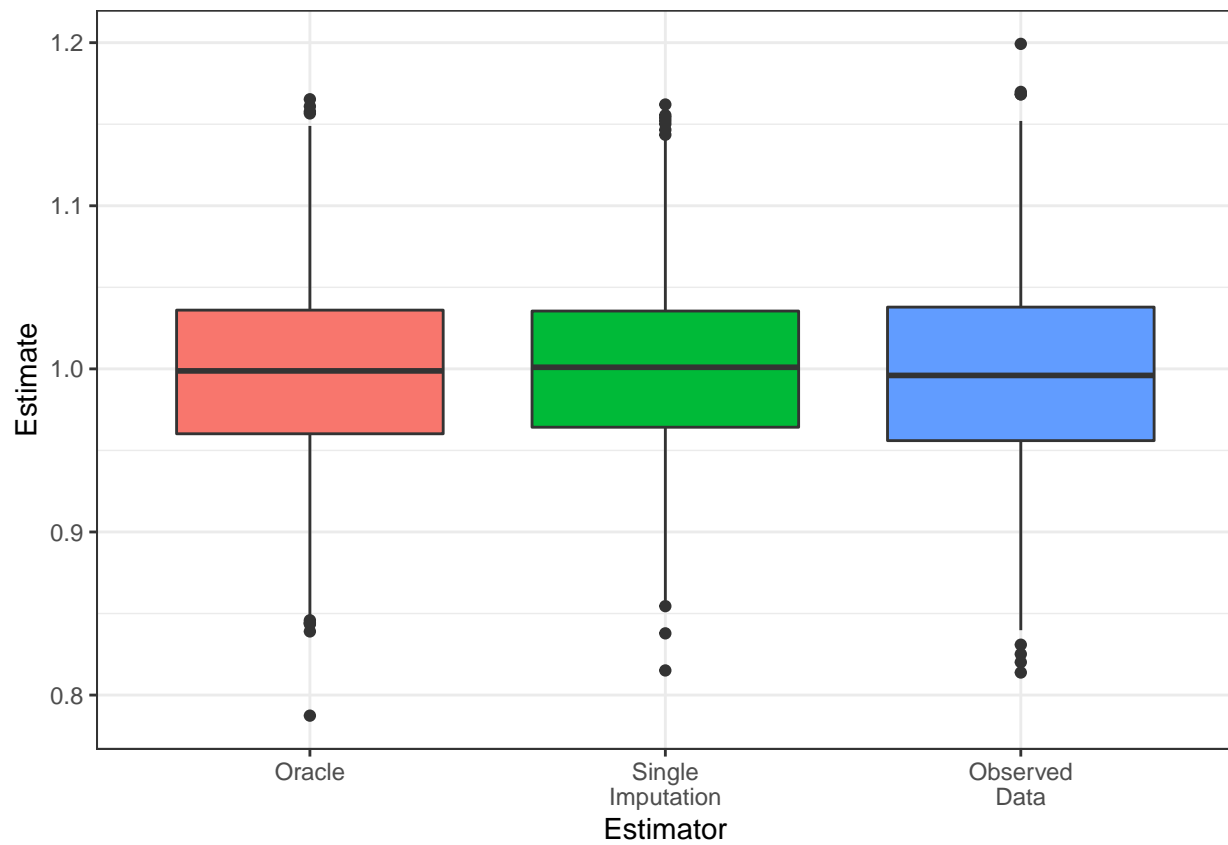


```
## # A tibble: 3 x 3
##   estimator  est    se
##   <chr>      <dbl> <dbl>
## 1 si        0.999 0.0518
## 2 obs       0.998 0.0579
## 3 oracle    0.998 0.0563
```

```
##           si      obs      oracle
## empirical 0.05772189 0.05991917 0.05383904
## estimated 0.05631670 0.05789101 0.05174918
```

Beta known, model incorrectly specified

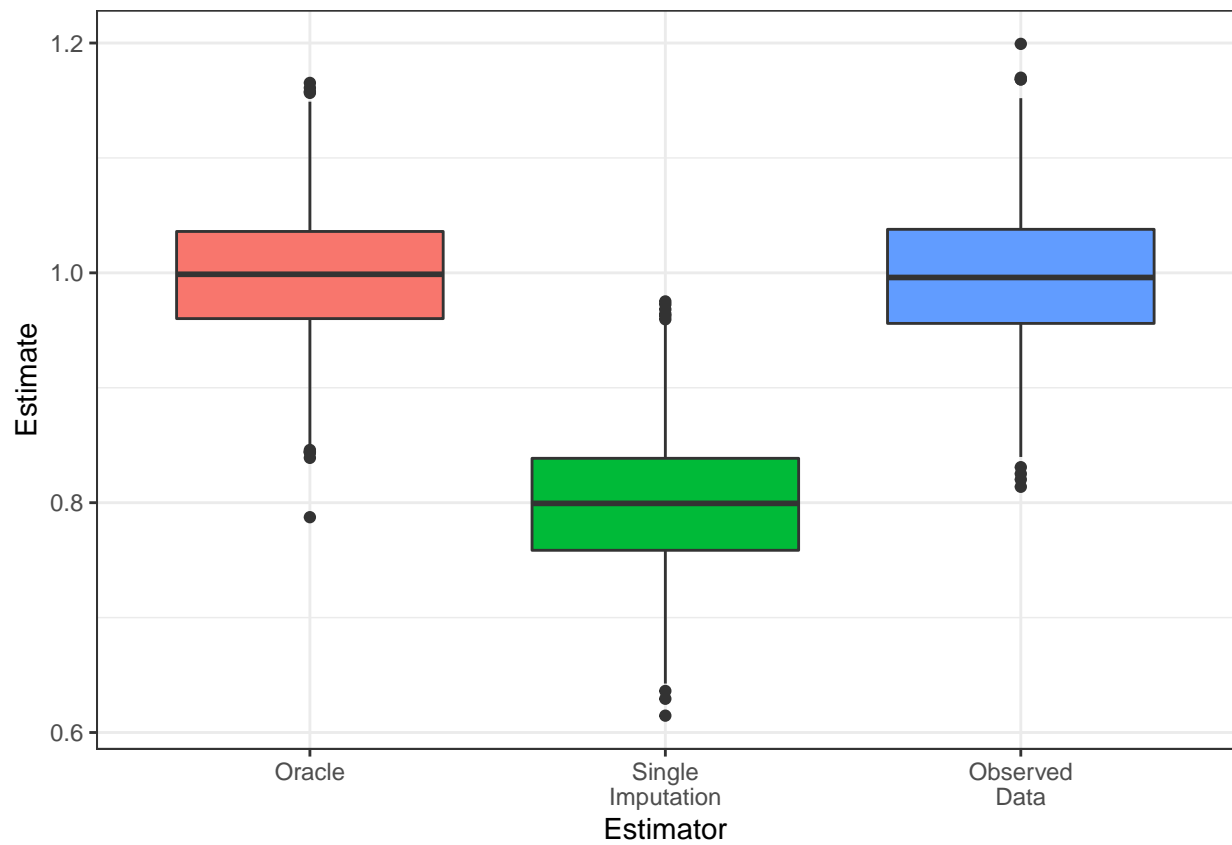
$$\hat{Y} \sim G, X$$



```
## # A tibble: 3 x 3
##   estimator  est    se
##   <chr>      <dbl> <dbl>
## 1 si        0.999 0.0528
## 2 obs       0.998 0.0579
## 3 oracle    0.998 0.0563

##           si      obs      oracle
## empirical 0.05772189 0.05991917 0.05469555
## estimated 0.05631670 0.05789101 0.05276199
```

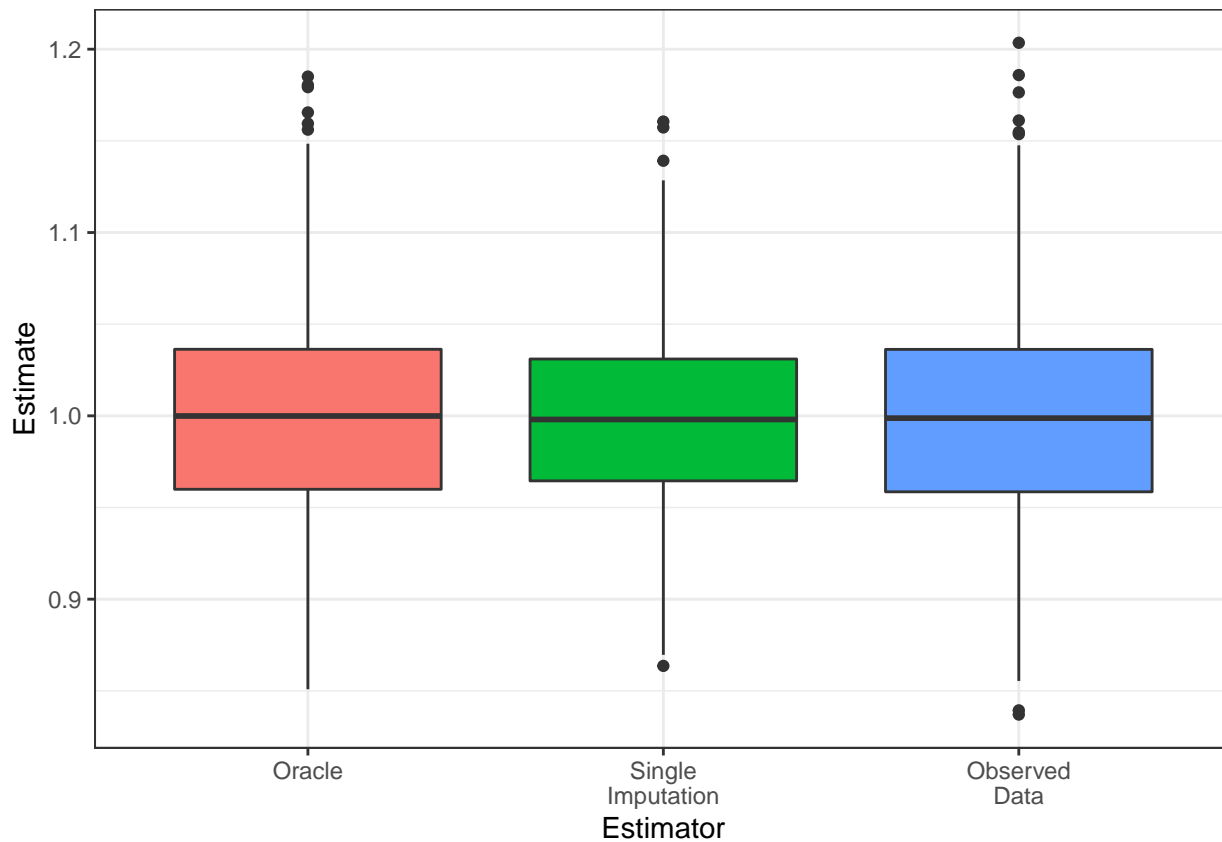
$\hat{Y} \sim X, Z$



```
## # A tibble: 3 x 3
##   estimator  est    se
##   <chr>      <dbl> <dbl>
## 1 si         0.799 0.0543
## 2 obs        0.998 0.0579
## 3 oracle     0.998 0.0563
```

```
##           si      obs      oracle
## empirical 0.05772189 0.05991917 0.06025671
## estimated 0.05631670 0.05789101 0.05424222
```

## Beta unkown and correctly specified

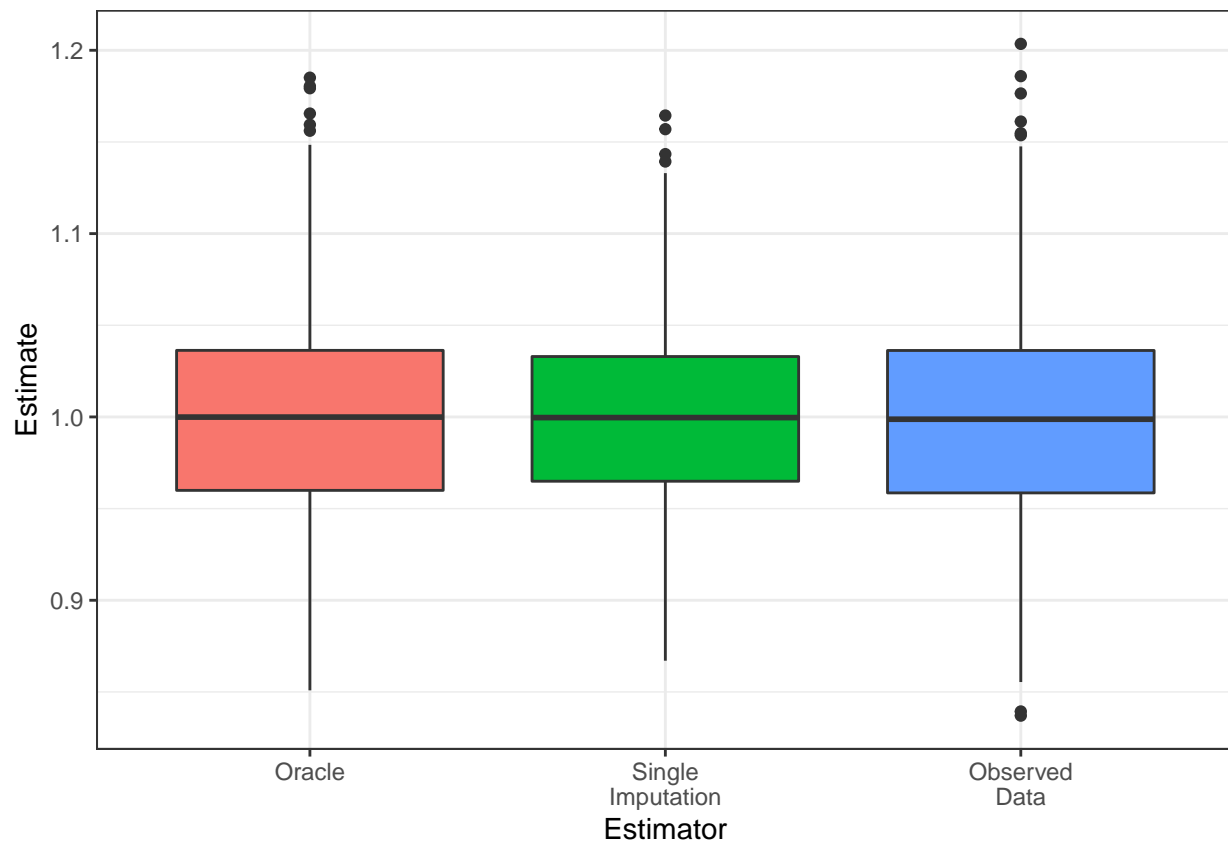


```
## # A tibble: 3 x 3
##   estimator  est    se
##   <chr>      <dbl> <dbl>
## 1 si        0.999 0.0569
## 2 obs       0.999 0.0578
## 3 oracle    0.999 0.0564
```

```
##           si      obs      oracle
## empirical 0.05542767 0.05657539 0.04743766
## estimated 0.05633309 0.05779937 0.05690282
```

## Beta unkown and incorrectly specified

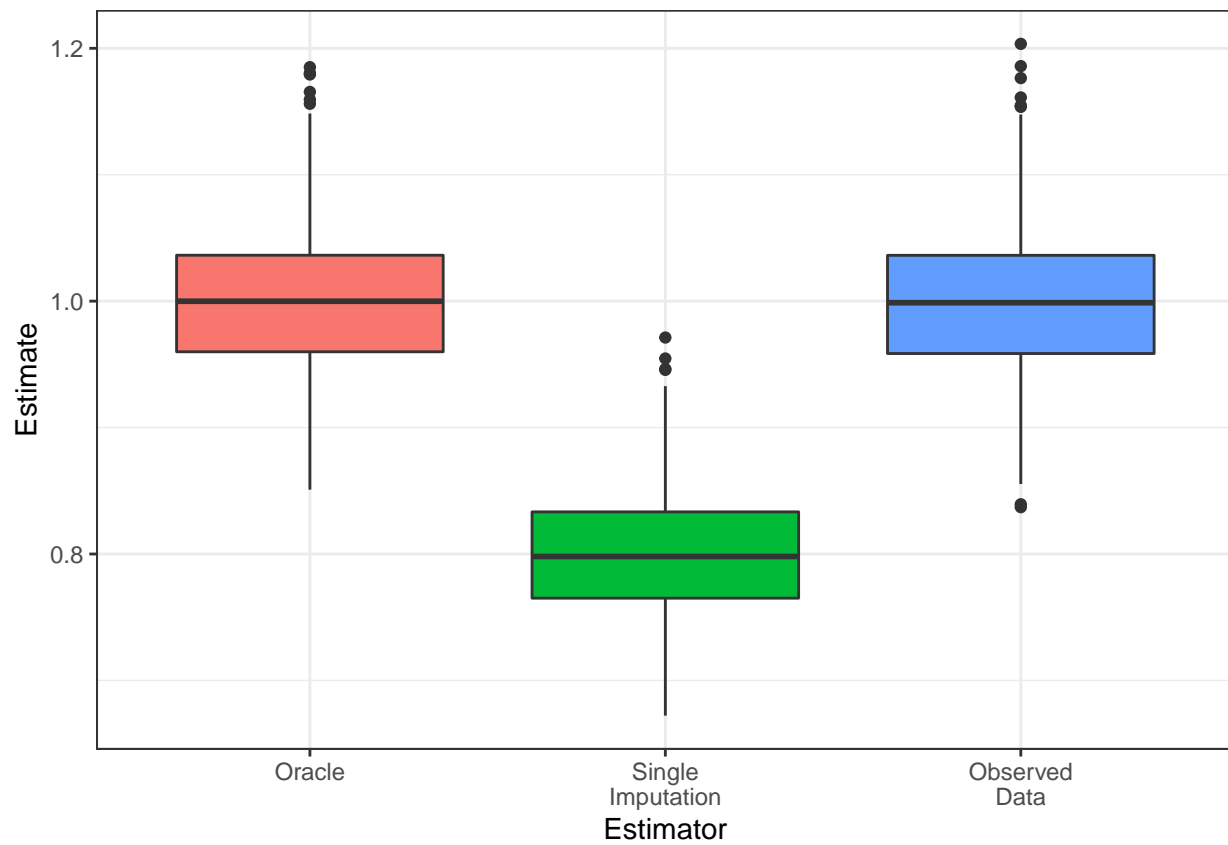
$$\hat{Y} \sim G, X$$



```
## # A tibble: 3 x 3
##   estimator  est    se
##   <chr>      <dbl> <dbl>
## 1 si        1.00  0.0594
## 2 obs       0.999 0.0578
## 3 oracle    0.999 0.0564

##           si      obs      oracle
## empirical 0.05542767 0.05657539 0.04871135
## estimated 0.05633309 0.05779937 0.05936070
```

$\hat{Y} \sim X, Z$

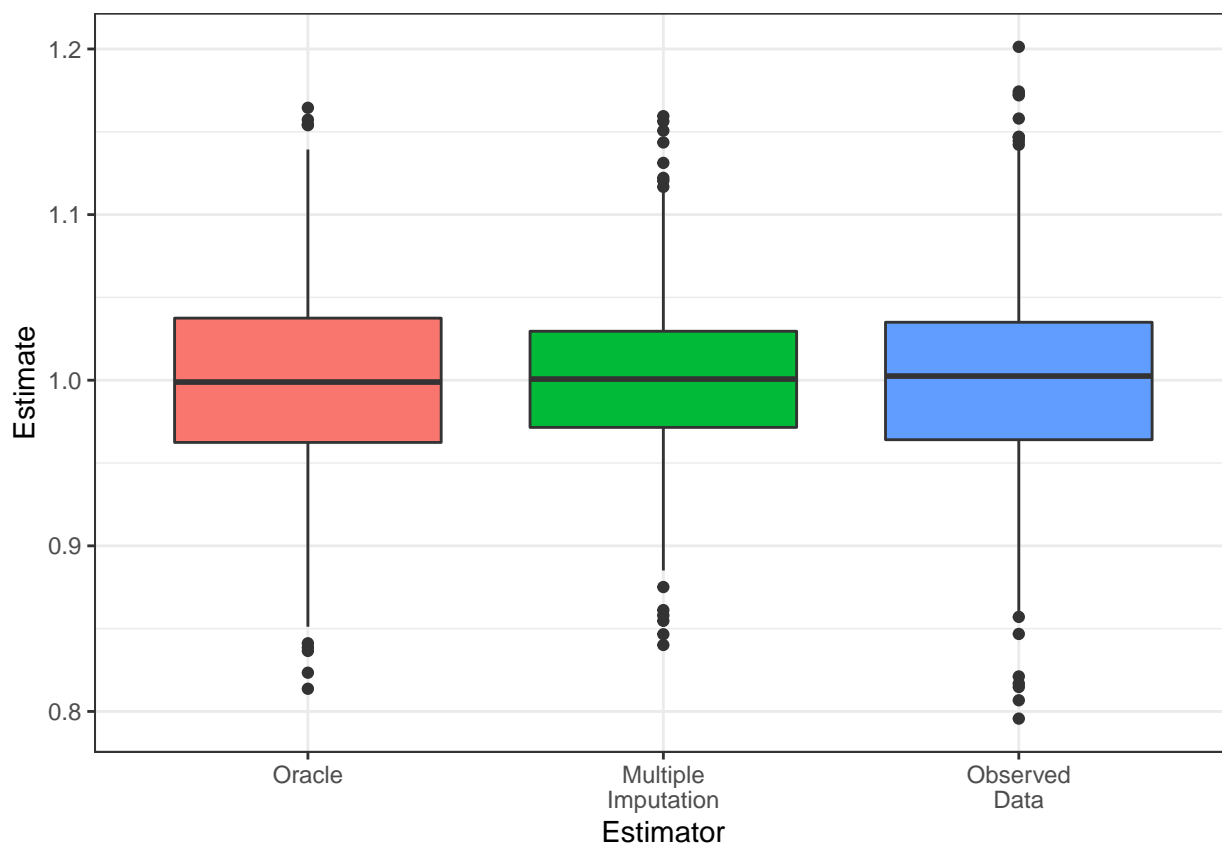


```
## # A tibble: 3 x 3
##   estimator  est    se
##   <chr>      <dbl> <dbl>
## 1 si        0.799 0.0618
## 2 obs       0.999 0.0578
## 3 oracle    0.999 0.0564
```

```
##           si      obs      oracle
## empirical 0.05542767 0.05657539 0.04821741
## estimated 0.05633309 0.05779937 0.06172049
```

# Multiple Imputation

Beta known, model correctly specified



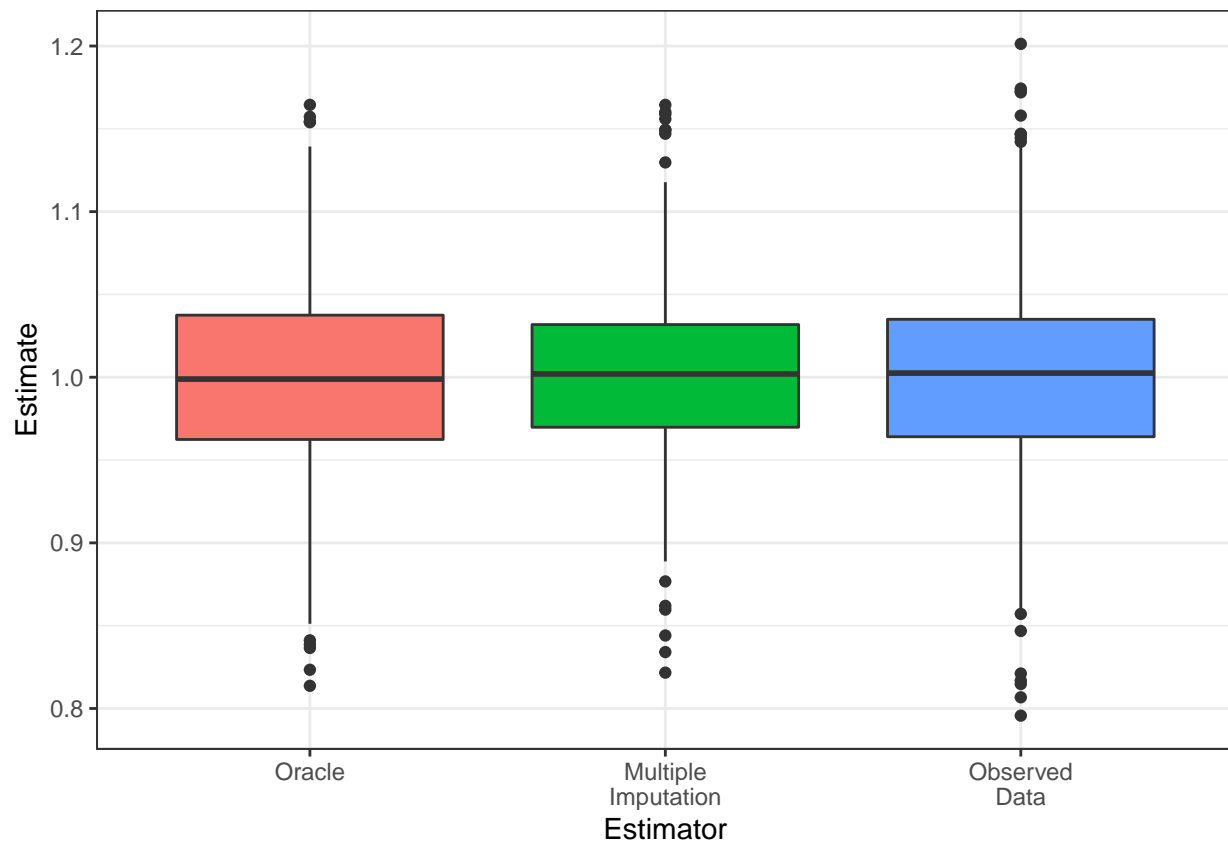
```
## # A tibble: 3 x 3
##   estimator  est    se
##   <chr>      <dbl> <dbl>
## 1 mi         1.00 0.0570
## 2 obs         1.00 0.0579
## 3 oracle     1.00 0.0564

##           oracle_est  obs_est  mi_est
## empirical 0.05522382 0.05728497 0.04592360
## estimated 0.05635012 0.05786868 0.05698082
```

Beta known, model incorrectly specified

$$\hat{Y} \sim G + X$$

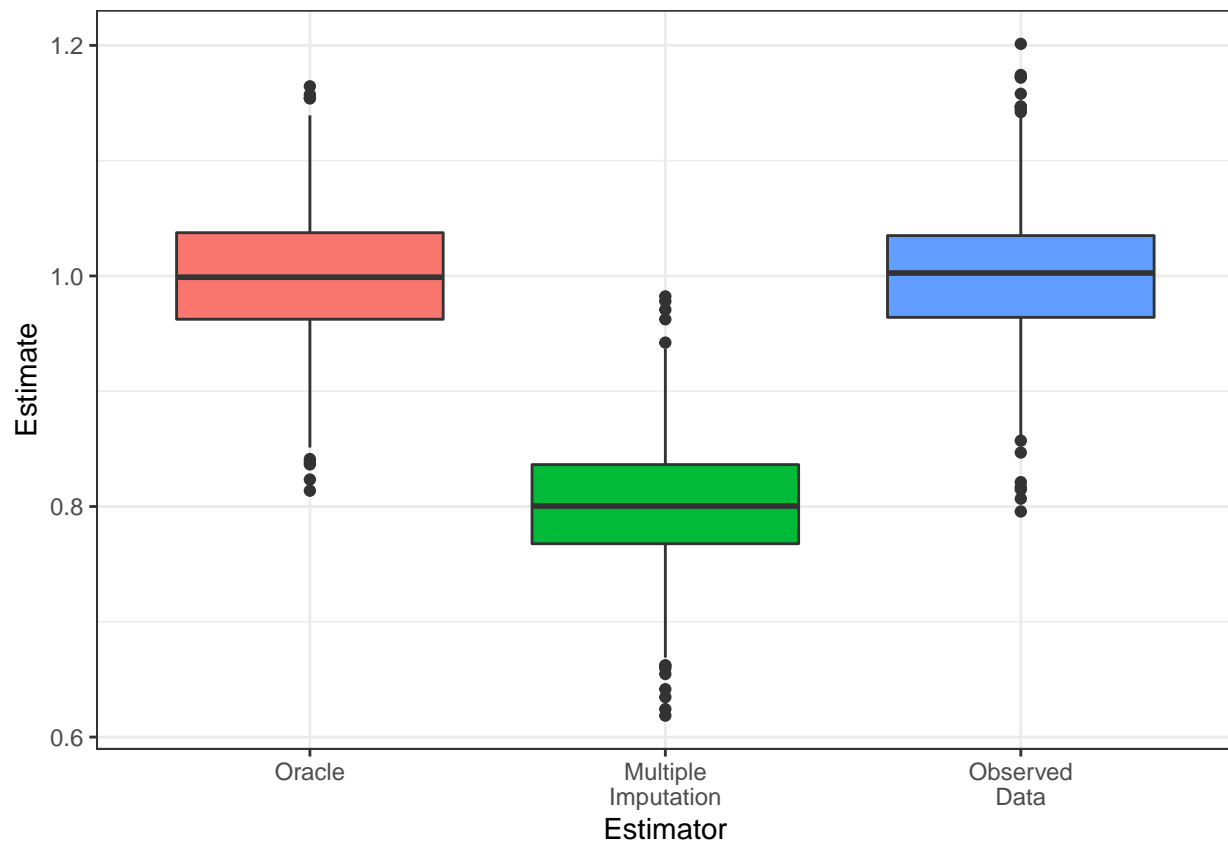




```
## # A tibble: 3 x 3
##   estimator  est    se
##   <chr>      <dbl> <dbl>
## 1 mi         1.00 0.0579
## 2 obs         1.00 0.0579
## 3 oracle     1.00 0.0564

##           oracle_est  obs_est  mi_est
## empirical 0.05522382 0.05728497 0.04790228
## estimated 0.05635012 0.05786868 0.05791107
```

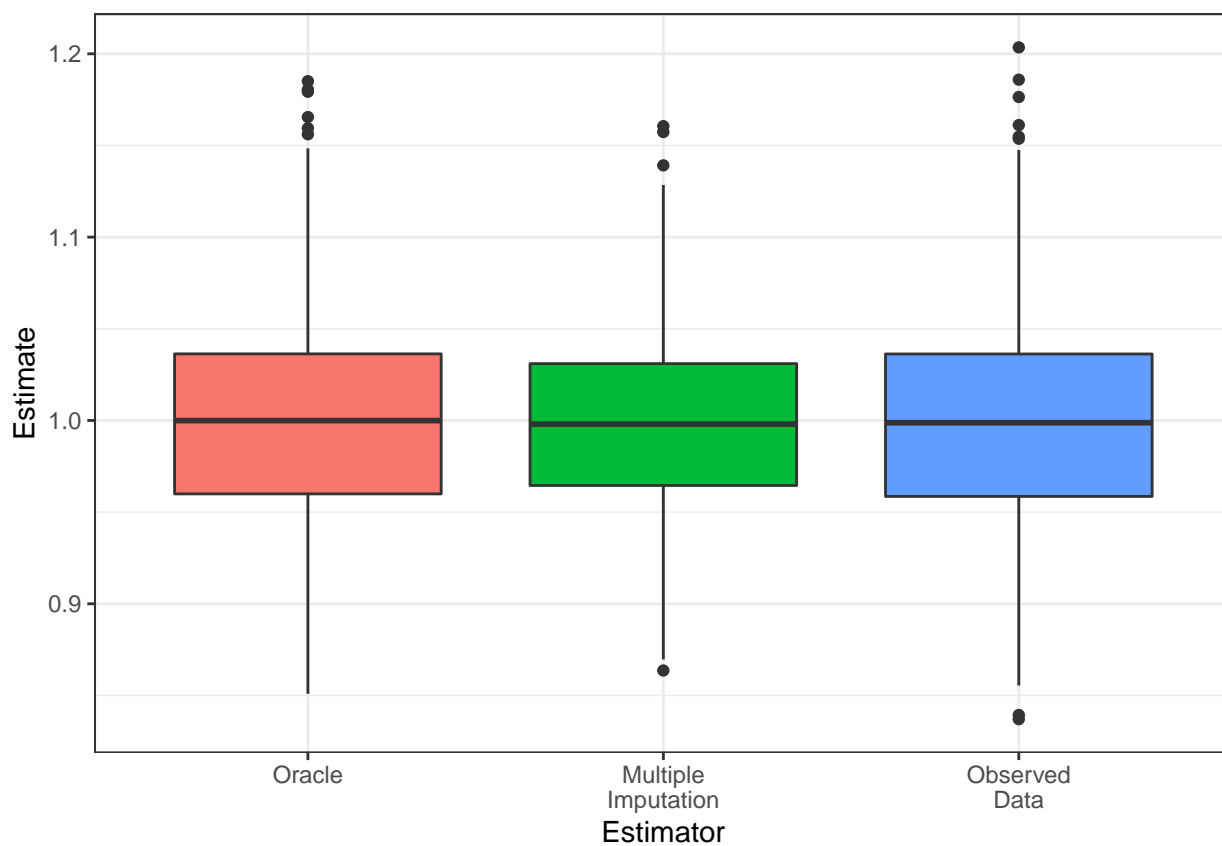
$$\hat{Y} \sim X + Z$$



```
## # A tibble: 3 x 3
##   estimator  est    se
##   <chr>      <dbl> <dbl>
## 1 mi        0.800 0.0593
## 2 obs       1.00  0.0579
## 3 oracle    1.00  0.0564

##           oracle_est  obs_est  mi_est
## empirical 0.05522382 0.05728497 0.05277274
## estimated 0.05635012 0.05786868 0.05927804
```

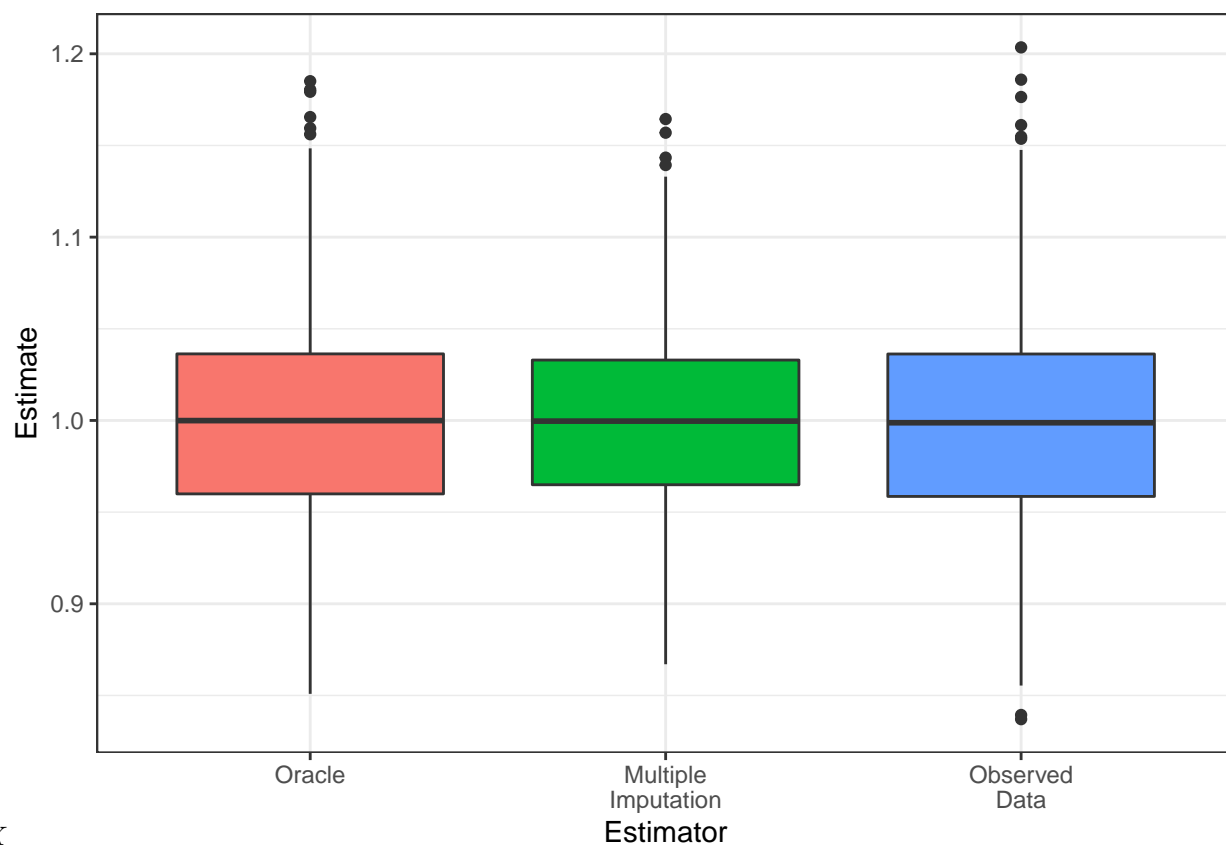
## Beta unknown, model correctly specified



```
## # A tibble: 3 x 3
##   estimator  est    se
##   <chr>      <dbl> <dbl>
## 1 mi         0.999 0.0569
## 2 obs         0.999 0.0578
## 3 oracle     0.999 0.0564
```

```
##           oracle_est  obs_est  mi_est
## empirical 0.05542767 0.05657539 0.04743766
## estimated 0.05633309 0.05779937 0.05690282
```

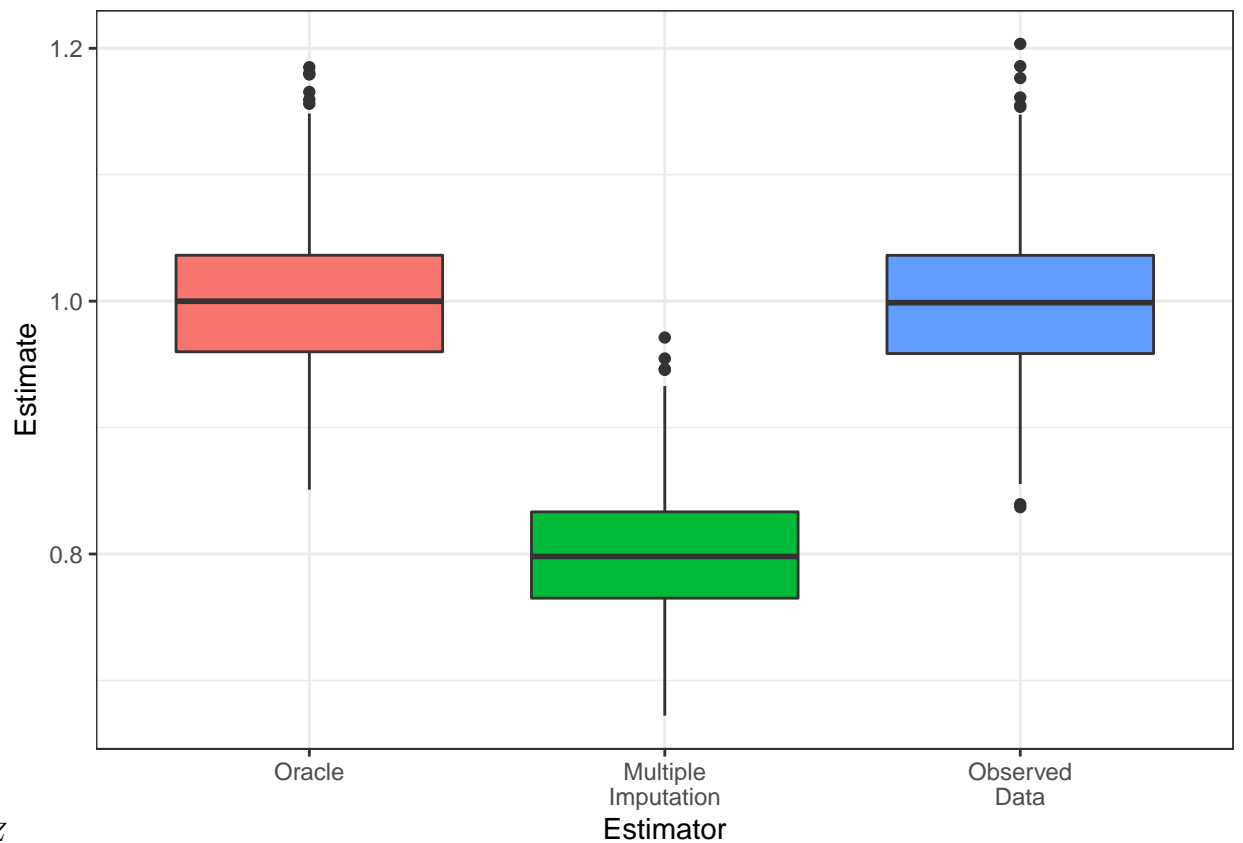
## Beta unknown, model incorrectly specified



$$\hat{Y} \sim G+X$$

```
## # A tibble: 3 x 3
##   estimator  est    se
##   <chr>      <dbl> <dbl>
## 1 mi         1.00  0.0594
## 2 obs        0.999 0.0578
## 3 oracle     0.999 0.0564
```

```
##           oracle_est  obs_est  mi_est
## empirical 0.05542767 0.05657539 0.04871135
## estimated 0.05633309 0.05779937 0.05936070
```



$$\hat{Y} \sim X+Z$$

```
## # A tibble: 3 x 3
##   estimator  est    se
##   <chr>      <dbl> <dbl>
## 1 mi        0.799 0.0618
## 2 obs       0.999 0.0578
## 3 oracle    0.999 0.0564
```

```
##           oracle_est  obs_est  mi_est
## empirical 0.05542767 0.05657539 0.04821741
## estimated 0.05633309 0.05779937 0.06172049
```