# Computed Values

This document computes and reports the quantities reported in the text of the `.tex` file.

## "The Big Improvements from an Easy Solution"

```
# load packages
library(tidyverse)
library(kableExtra)

# load simulations
sim_df <- read_rds("simulations/simulations.rds") %>%
  glimpse()
```

```
## Observations: 342
## Variables: 14
## $ n            <dbl> 30, 30, 30, 30, 30, 30, 40, 40, 40, 40, 40, 40, 5...
## $ k            <dbl> 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3...
## $ b0           <dbl> -1.0, -1.0, -0.5, -0.5, 0.0, 0.0, -1.0, -1.0, -0....
## $ b1           <dbl> 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5,...
## $ n_sims       <dbl> 1e+05, 1e+05, 1e+05, 1e+05, 1e+05, 1e+05, 1e+05, ...
## $ true_coef    <dbl> 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5,...
## $ method       <fct> ML, PML, ML, PML, ML, PML, ML, PML, ML, PML, ML, ...
## $ ev           <dbl> 0.7060483, 0.5159187, 0.6327240, 0.5128711, 0.634...
## $ mc_error_ev  <dbl> 0.0027100703, 0.0019201452, 0.0016560016, 0.00123...
## $ bias         <dbl> 0.206048268, 0.015918676, 0.132723967, 0.01287107...
## $ percent_bias <dbl> 41.2096537, 3.1837352, 26.5447934, 2.5742145, 26....
## $ var          <dbl> 0.73444811, 0.36869574, 0.27423413, 0.15254398, 0...
## $ mse          <dbl> 0.77689664, 0.36894546, 0.29184704, 0.15270812, 0...
## $ prop_ones    <dbl> 0.2832472, 0.2832472, 0.4015349, 0.4015349, 0.496...
```

### Bias

```
# find largest bias scenario (will be for ML)
max_bias_df <- sim_df %>%
  filter(percent_bias == max(percent_bias))
# make table for this scenario for ML and PML
sim_df %>%
  filter(n == max_bias_df$n & k == max_bias_df$k & b0 == max_bias_df$b0) %>%
  mutate(`Bias` = round(bias, 2),
         `% Bias` = scales::percent(percent_bias/100, accuracy = 1)) %>%
  select(`Sample Size` = n,
         `Number of Other Covariates` = k,
         `Intercept` = b0,
         `Method` = method,
         `Bias`,
         `% Bias`) %>%
  kable(format = "latex",
        caption = "Largest Bias.",
```

```
        booktabs = TRUE) %>%
  kable_styling(latex_options = c("striped", "hold_position"))
```

Table 1: Largest Bias.

| Sample Size | Number of Other Covariates | Intercept | Method | Bias | % Bias |
|---|---|---|---|---|---|
| 30 | 9 | -0.5 | ML | 0.61 | 122% |
| 30 | 9 | -0.5 | PML | 0.06 | 12% |

```
# smallest bias
sim_df %>%
  filter(method == "ML") %>%
  top_n(-1, percent_bias) %>%
  mutate(`Bias` = round(bias, 2),
         `% Bias` = scales::percent(percent_bias/100, accuracy = 1)) %>%
  select(`Sample Size` = n,
         `Number of Other Covariates` = k,
         `Intercept` = b0,
         `Method` = method,
         `Bias`,
         `% Bias`)  %>%
  kable(format = "latex",
        caption = "Smallest Bias",
        booktabs = TRUE) %>%
  kable_styling(latex_options = c("striped", "hold_position"))
```

Table 2: Smallest Bias

| Sample Size | Number of Other Covariates | Intercept | Method | Bias | % Bias |
|---|---|---|---|---|---|
| 210 | 3 | -1 | ML | 0.01 | 3% |

**Variance**

```
# largest bias
sim_df %>%
  select(n, k, b0, method, var) %>%
  spread(method, var) %>%
  mutate(var_infl = ML/PML - 1,
         var_infl_chr = scales::percent(var_infl, accuracy = 1)) %>%
  filter(n %in% c(30, 60, 210) & b0 == -1)  %>%
  select(-var_infl) %>%
  select(`Sample Size` = n,
         `Number of Other Covariates` = k,
         `Intercept` = b0,
         `Variance Inflation` = var_infl_chr) %>%
  kable(format = "latex",
        caption = "The Variance Inflation in Several Scenarios",
        booktabs = TRUE,
        digits = 2) %>%
  kable_styling(latex_options = c("striped", "hold_position"))
```

Table 3: The Variance Inflation in Several Scenarios

| Sample Size | Number of Other Covariates | Intercept | Variance Inflation |
|---|---|---|---|
| 30 | 3 | -1 | 99% |
| 30 | 6 | -1 | 178% |
| 30 | 9 | -1 | 436% |
| 60 | 3 | -1 | 27% |
| 60 | 6 | -1 | 58% |
| 60 | 9 | -1 | 117% |
| 210 | 3 | -1 | 6% |
| 210 | 6 | -1 | 10% |
| 210 | 9 | -1 | 14% |

**MSE**

```
# largest bias
mse_df <- sim_df %>%
  select(n, k, b0, method, mse) %>%
  spread(method, mse) %>%
  mutate(mse_infl = ML/PML - 1,
         mse_infl_chr = scales::percent(mse_infl, accuracy = 1))

# largest mse inflation
top_n(mse_df, 1, mse_infl)
```

```
##    n k   b0       ML       PML mse_infl mse_infl_chr
## 1 30 9 -0.5 1.998806 0.2435322 7.207563         721%
```

```
# percent of mse infl larger than 100%
mse_df %>%
  summarize(`100%` = mean(mse_infl > 1),
            `50%` = mean(mse_infl > 0.5),
            `25%` = mean(mse_infl > 0.25)) %>%
  gather(`MSE Infl. Threshold`, `% Above`, `100%`:`25%`) %>%
  mutate(`% Above` = scales::percent(`% Above`, accuracy = 1))  %>%
  kable(format = "latex",
        caption = "Percent of MSE Infl. Larger than 100\\%, 50\\%, and 25\\%.",
        booktabs = TRUE) %>%
  kable_styling(latex_options = c("striped", "hold_position"))
```

Table 4: Percent of MSE Infl. Larger than 100%, 50%, and 25%.

| MSE Infl. Threshold | % Above |
|---|---|
| 100% | 11% |
| 50% | 21% |
| 25% | 45% |

```
# relative contribution of variance
rc_df_bias <- sim_df %>%
  select(n, k, b0, method, bias) %>%
  spread(method, bias) %>%
  rename(ml_bias = ML, pml_bias = PML)
```

```
rc_df_var <- sim_df %>%
  select(n, k, b0, method, var) %>%
  spread(method, var) %>%
  rename(ml_var = ML, pml_var = PML)

rc_df <- left_join(rc_df_bias, rc_df_var) %>%
  mutate(var_contrib = 100*ml_var/(pml_var + pml_bias^2),
         bias_contrib = 100*(ml_bias^2)/(pml_var + pml_bias^2),
         rel_contrib = var_contrib/bias_contrib)
```

```
## Joining, by = c("n", "k", "b0")
```

```
rc_df %>%
  filter(n %in% c(30, 210)) %>%
  group_by(n) %>%
  summarize(min = round(min(rel_contrib), 0),
            max = round(max(rel_contrib), 0)) %>%
  mutate(n = paste0("N = ", n)) %>%
  rename(N = n, `Minimum` = min, `Maximum` = max) %>%
  kable(format = "latex",
        caption = "Relative contribution of variance to MSE.",
        booktabs = TRUE) %>%
  kable_styling(latex_options = c("striped", "hold_position"))
```

Table 5: Relative contribution of variance to MSE.

| N | Minimum | Maximum |
|---|---|---|
| N = 30 | 4 | 17 |
| N = 210 | 27 | 166 |

## "The Substantive Importance of the Big Improvements"

```
# load coefficient estimates
coef_df <- read_rds("ge-replication/coefficient-estimates.rds")

# create table of percent changes
coef_df %>%
  select(var_name_print, model_name, est) %>%
  spread(model_name, est) %>%
  mutate(percent_decrease = scales::percent(PML/ML - 1, accuracy = 1)) %>%
  rename(`Variable Name` = var_name_print, `ML Est.` = ML, `PML Est.` = PML, `% Decrease` = percent_dec
  kable(format = "latex",
        digits = 2,
        caption = "Percent Difference in ML and PML Estimates",
        booktabs = TRUE) %>%
  kable_styling(latex_options = c("striped", "hold_position"))
```

```
# load coefficient estimates
osf_df <- read_rds("ge-replication/out-of-sample-fit.rds")

# create table of percent changes
osf_df %>%
```

Table 6: Percent Difference in ML and PML Estimates

| Variable Name | ML Est. | PML Est. | % Decrease |
|---|---|---|---|
| Intercept | -26.18 | -16.54 | -37% |
| Death-Qualified Jury | 2.47 | 1.82 | -26% |
| Capital Punishment Proportional to Offense | 5.62 | 3.67 | -35% |
| Particularizing Circumstances | 2.11 | 1.50 | -29% |
| Aggravating Factors | 2.17 | 1.40 | -35% |
| State Psychiatric Examination | 4.66 | 3.05 | -35% |
| Conservative Political Environment | 4.12 | 2.41 | -41% |
| Court Change | 1.10 | 0.56 | -49% |
| State Appellant | 2.70 | 1.88 | -30% |
| Inexperienced Defense Counsel | 2.27 | 1.52 | -33% |
| Repeat Player State | 3.60 | 2.37 | -34% |
| Amicus Brief from Solicitor General | 3.21 | 1.67 | -48% |

```
  spread(method, score) %>%
  mutate(percent_decrease = scales::percent(PML/ML - 1, accuracy = 1)) %>%
  rename(`Score Type` = score_type, `% Decrease` = percent_decrease) %>%
  kable(format = "latex",
        digits = 2,
        caption = "Percent Decrease in ML and PML Out-of-Sample Fit",
        booktabs = TRUE) %>%
  kable_styling(latex_options = c("striped", "hold_position"))
```

Table 7: Percent Decrease in ML and PML Out-of-Sample Fit

| Score Type | ML | PML | % Decrease |
|---|---|---|---|
| Brier Score | 0.17 | 0.16 | -7% |
| Log Score | 0.89 | 0.53 | -41% |

```
# load coefficient estimates
pr_df <- read_rds("ge-replication/qi-probs.rds")

# create table of percent changes
pr_df %>%
  spread(method, prob) %>%
  rename(`State Type` = st) %>%
  kable(format = "latex",
        digits = 2,
        caption = "Probability of Conservative Decision",
        booktabs = TRUE) %>%
  kable_styling(latex_options = c("striped", "hold_position"))
```

Table 8: Probability of Conservative Decision

| State Type | ML | PML |
|---|---|---|
| Not a Repeat Player | 0.06 | 0.17 |
| Repeat Player | 0.68 | 0.69 |

```r
# load qi estimates
fd_df <- read_rds("ge-replication/qi-fd.rds") %>%
  mutate(QI = "First Difference")

rr_df <- read_rds("ge-replication/qi-rr.rds") %>%
  mutate(QI = "Risk Ratio")

bind_rows(fd_df, rr_df) %>%
  select(-lwr_90, -upr_90) %>%
  spread(method, est) %>%
  mutate(percent_decrease = scales::percent(PML/ML - 1, accuracy = 1)) %>%
  rename(`% Decrease` = percent_decrease) %>%
  kable(format = "latex",
        digits = 2,
        caption = "Quantities of Interest",
        booktabs = TRUE) %>%
  kable_styling(latex_options = c("striped", "hold_position"))
```

Table 9: Quantities of Interest

| QI | ML | PML | % Decrease |
|---|---|---|---|
| First Difference | 0.63 | 0.52 | -17% |
| Risk Ratio | 12.27 | 4.00 | -67% |