The purpose of this simulation study is to examine different ways of evaluating an ODTR. Specifically, we examine how different estimators (g-comp, IPTW, IPTW-DR, TMLE, CV-TMLE) do with approximating the statistical parameter $E_0[Y_{d_0}]$ and the data-adaptive parameter $E_0[Y_{d_n}]$, under different SL ODTR estimators (ie different SL libraries, varying in "aggressiveness"). Here, d_0 is true optimal rule, d_n is estimate of optimal rule.

1 Description of DGP

$$\begin{aligned} W_1, W_2, W_3, W_4 &\sim Normal(\mu = 0, \sigma^2 = 1) \\ A &\sim Bernoulli(p = 0.5) \\ Y &\sim Bernoulli(p) \ . \end{aligned}$$

$$p = 0.5 * logit^{-1}(1 - W_1^2 + 3W_2 + 5W_3^2 A - 4.45A) + 0.5 logit^{-1}(-0.5 - W_3 + 2W_1W_2 + 3|W2|A - 1.5A) \ ,$$

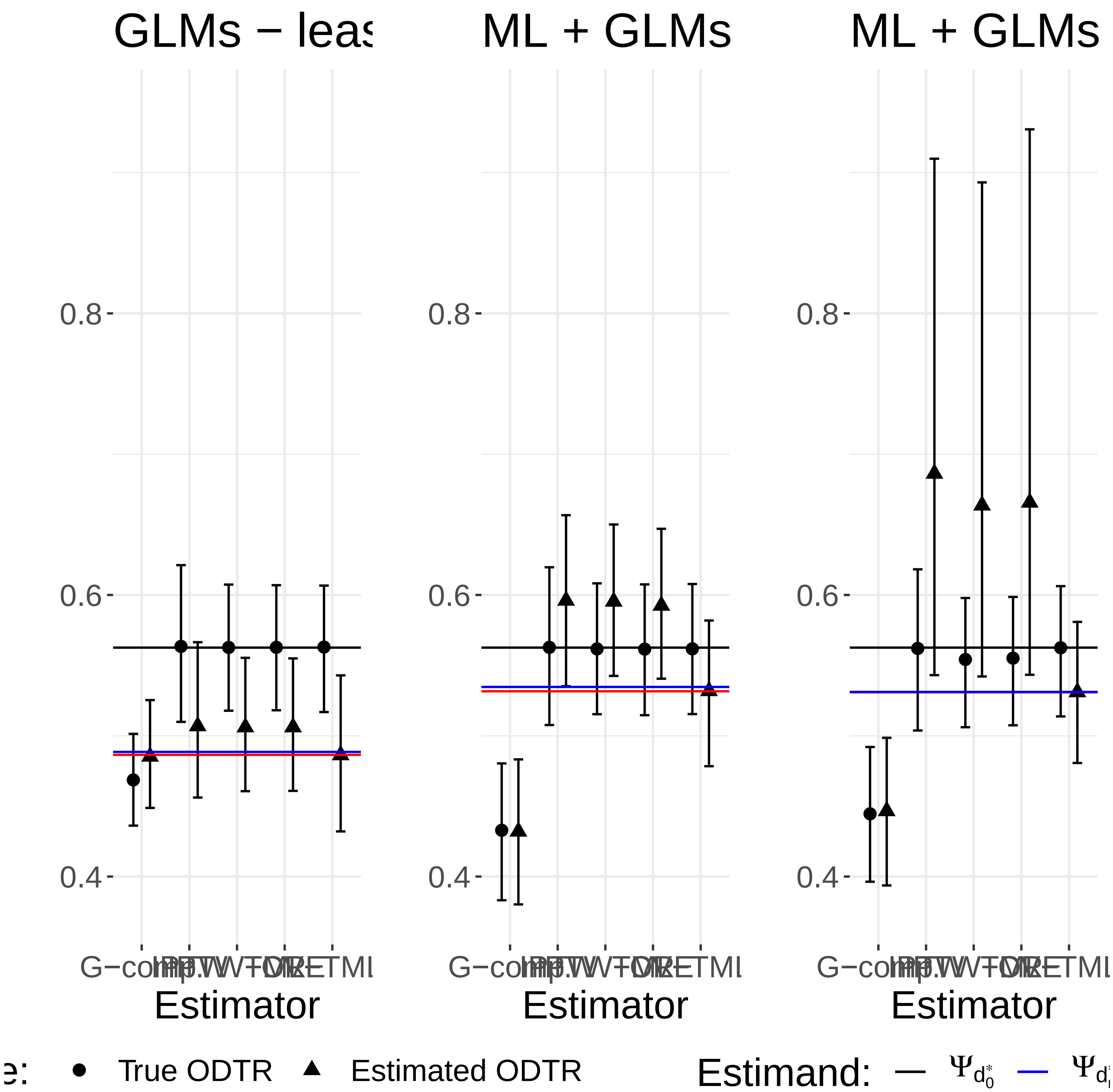
True blip function is:

$$\begin{aligned} B_0(W) =& 0.5[logit^{-1}(1 - W_1^2 + 3W_2 + 5W_3^2 - 4.45) + logit^{-1}(-0.5 - W_3 + 2W_1W_2 + 3|W2| - 1.5) \\ & - logit^{-1}(1 - W_1^2 + 3W_2) + logit^{-1}(-0.5 - W_3 + 2W_1W_2)] \ . \end{aligned}$$

2 Library legend

- Incorrect GLM
  - QAW.SL.library = linear model main terms W and A and interaction with W and A
  - blip.SL.library = linear model with main terms W
- GLMs
  - QAW.SL.library = linear model with W_j and A as main terms and W_j*A interaction for each j
  - blip.SL.library = linear model with main terms W_j for each j
- ML + GLMs not aggressive
  - QAW.SL.library = GLMs library AND SL.glm, SL.mean, SL.glm.interaction, SL.earth, SL.nnet, SL.svm, SL.rpart
  - blip.SL.library = GLMs library AND SL.glm, SL.mean, SL.glm.interaction, SL.earth, SL.nnet, SL.svm, SL.rpart
- ML + GLMs not aggressive
  - QAW.SL.library = ML + GLMs aggressive library AND SL.randomForest
  - blip.SL.library = ML + GLMs aggressive library AND SL.randomForest

3 Results




```
## pdf
##   2
```

```
make_table_EYdopt(EYdopt = EYdoptbin_glms, truevalues = DGP_bin_complex_true_values)
```

```
##                 Bias Variance    MSE Coverage      Estimator
## gcomp        -0.0773    4e-04 0.0064        -          gcomp
## IPTW         -0.0558    8e-04 0.0039      45%           IPTW
## IPTW_DR      -0.0565    6e-04 0.0038    30.1%        IPTW_DR
## TMLE         -0.0565    6e-04 0.0038    29.8%           TMLE
## LTMLE             NA       NA     NA        -          LTMLE
## CV.TMLE      -0.0764    9e-04 0.0067    14.7%        CV.TMLE
```

```
## gcomp_dopt0      -0.0940    3e-04 0.0091        -      gcomp_dopt0
## IPTW_dopt0        0.0009    8e-04 0.0008    95.3%       IPTW_dopt0
## IPTW_DR_dopt0     0.0001    5e-04 0.0005    93.7%    IPTW_DR_dopt0
## TMLE_dopt0        0.0002    5e-04 0.0005    93.7%       TMLE_dopt0
## LTMLE_dopt0           NA       NA     NA        -      LTMLE_dopt0
## CV.TMLE_dopt0     0.0004    5e-04 0.0005    93.7%    CV.TMLE_dopt0
## gcomp_sampspec   -0.0033    4e-04 0.0006        -   gcomp_sampspec
## IPTW_sampspec     0.0183    8e-04 0.0009    94.3%    IPTW_sampspec
## IPTW_DR_sampspec  0.0175    6e-04 0.0007    90.6% IPTW_DR_sampspec
## LTMLE_sampspec        NA       NA     NA        -   LTMLE_sampspec
## TMLE_sampspec     0.0175    6e-04 0.0007    90.7%    TMLE_sampspec
## CV.TMLE_sampspec -0.0002    9e-04 0.0005    94.3% CV.TMLE_sampspec
```

```
make_table_EYdopt(EYdopt = EYdoptbin_MLnotaggglms, truevalues = DGP_bin_complex_true_values)
```

```
##                     Bias Variance    MSE Coverage        Estimator
## gcomp            -0.1306    7e-04 0.0178        -            gcomp
## IPTW              0.0334    1e-03 0.0021    76.1%             IPTW
## IPTW_DR           0.0327    8e-04 0.0019    66.5%          IPTW_DR
## TMLE              0.0298    8e-04 0.0016    71.3%             TMLE
## LTMLE                 NA       NA     NA        -            LTMLE
## CV.TMLE          -0.0308    7e-04 0.0017      69%          CV.TMLE
## gcomp_dopt0      -0.1298    6e-04 0.0175        -      gcomp_dopt0
## IPTW_dopt0        0.0002    8e-04 0.0008    94.7%       IPTW_dopt0
## IPTW_DR_dopt0    -0.0009    6e-04 0.0006      94%    IPTW_DR_dopt0
## TMLE_dopt0       -0.0011    5e-04 0.0005    93.6%       TMLE_dopt0
## LTMLE_dopt0           NA       NA     NA        -      LTMLE_dopt0
## CV.TMLE_dopt0    -0.0009    5e-04 0.0005    93.2%    CV.TMLE_dopt0
## gcomp_sampspec   -0.1027    7e-04 0.0114        -   gcomp_sampspec
## IPTW_sampspec     0.0614    1e-03 0.0046    43.8%    IPTW_sampspec
## IPTW_DR_sampspec  0.0607    8e-04 0.0044    28.9% IPTW_DR_sampspec
## LTMLE_sampspec        NA       NA     NA        -   LTMLE_sampspec
## TMLE_sampspec     0.0578    8e-04 0.0040    30.4%    TMLE_sampspec
## CV.TMLE_sampspec  0.0002    7e-04 0.0005      94% CV.TMLE_sampspec
```

```
make_table_EYdopt(EYdopt = EYdoptbin_MLaggglms, truevalues = DGP_bin_complex_true_values)
```

```
##                     Bias Variance    MSE Coverage        Estimator
## gcomp            -0.1161   0.0007 0.0142        -            gcomp
## IPTW              0.1236   0.0109 0.0262      31%             IPTW
## IPTW_DR           0.1010   0.0092 0.0194      33%          IPTW_DR
## TMLE              0.1031   0.0108 0.0214    33.6%             TMLE
## LTMLE                 NA       NA     NA        -            LTMLE
## CV.TMLE          -0.0316   0.0007 0.0017    68.6%          CV.TMLE
## gcomp_dopt0      -0.1180   0.0006 0.0146        -      gcomp_dopt0
## IPTW_dopt0       -0.0006   0.0009 0.0009      94%       IPTW_dopt0
## IPTW_DR_dopt0    -0.0084   0.0005 0.0006    90.1%    IPTW_DR_dopt0
## TMLE_dopt0       -0.0075   0.0005 0.0006    90.6%       TMLE_dopt0
## LTMLE_dopt0           NA       NA     NA        -      LTMLE_dopt0
## CV.TMLE_dopt0    -0.0001   0.0005 0.0005    93.6%    CV.TMLE_dopt0
## gcomp_sampspec   -0.0846   0.0007 0.0081        -   gcomp_sampspec
```

```
## IPTW_sampspec     0.1551   0.0109 0.0366    16.3%    IPTW_sampspec
## IPTW_DR_sampspec  0.1325   0.0092 0.0283    15.8% IPTW_DR_sampspec
## LTMLE_sampspec        NA       NA     NA        -   LTMLE_sampspec
## TMLE_sampspec     0.1346   0.0108 0.0307    15.7%    TMLE_sampspec
## CV.TMLE_sampspec  0.0001   0.0007 0.0005    94.8% CV.TMLE_sampspec
```

```
##                         Comparison                  Library Estimator    Bias
## gcomp_sampspec    EnYdn for E0Ydn                      GLMs   G-comp. -0.0033
## IPTW_sampspec     EnYdn for E0Ydn                      GLMs      IPTW  0.0183
## IPTW_DR_sampspec  EnYdn for E0Ydn                      GLMs   IPTW-DR  0.0175
## TMLE_sampspec     EnYdn for E0Ydn                      GLMs      TMLE  0.0175
## CV.TMLE_sampspec  EnYdn for E0Ydn                      GLMs   CV-TMLE -0.0002
## gcomp_sampspec1   EnYdn for E0Ydn ML + GLMs not aggressive   G-comp. -0.1027
## IPTW_sampspec1    EnYdn for E0Ydn ML + GLMs not aggressive      IPTW  0.0614
## IPTW_DR_sampspec1 EnYdn for E0Ydn ML + GLMs not aggressive   IPTW-DR  0.0607
## TMLE_sampspec1    EnYdn for E0Ydn ML + GLMs not aggressive      TMLE  0.0578
## CV.TMLE_sampspec1 EnYdn for E0Ydn ML + GLMs not aggressive   CV-TMLE  0.0002
## gcomp_sampspec2   EnYdn for E0Ydn     ML + GLMs aggressive   G-comp. -0.0846
## IPTW_sampspec2    EnYdn for E0Ydn     ML + GLMs aggressive      IPTW  0.1551
## IPTW_DR_sampspec2 EnYdn for E0Ydn     ML + GLMs aggressive   IPTW-DR  0.1325
## TMLE_sampspec2    EnYdn for E0Ydn     ML + GLMs aggressive      TMLE  0.1346
## CV.TMLE_sampspec2 EnYdn for E0Ydn     ML + GLMs aggressive   CV-TMLE  0.0001
##                   Variance    MSE Coverage
## gcomp_sampspec      0.0004 0.0006        -
## IPTW_sampspec       0.0008 0.0009    94.3%
## IPTW_DR_sampspec    0.0006 0.0007    90.6%
## TMLE_sampspec       0.0006 0.0007    90.7%
## CV.TMLE_sampspec    0.0009 0.0005    94.3%
## gcomp_sampspec1     0.0007 0.0114        -
## IPTW_sampspec1      0.0010 0.0046    43.8%
## IPTW_DR_sampspec1   0.0008 0.0044    28.9%
## TMLE_sampspec1      0.0008 0.0040    30.4%
## CV.TMLE_sampspec1   0.0007 0.0005      94%
## gcomp_sampspec2     0.0007 0.0081        -
## IPTW_sampspec2      0.0109 0.0366    16.3%
## IPTW_DR_sampspec2   0.0092 0.0283    15.8%
## TMLE_sampspec2      0.0108 0.0307    15.7%
## CV.TMLE_sampspec2   0.0007 0.0005    94.8%
##                     Comparison                  Library Estimator    Bias
## gcomp_dopt0    EnYd0 for E0Yd0                      GLMs   G-comp. -0.0940
## IPTW_dopt0     EnYd0 for E0Yd0                      GLMs      IPTW  0.0009
## IPTW_DR_dopt0  EnYd0 for E0Yd0                      GLMs   IPTW-DR  0.0001
## TMLE_dopt0     EnYd0 for E0Yd0                      GLMs      TMLE  0.0002
## CV.TMLE_dopt0  EnYd0 for E0Yd0                      GLMs   CV-TMLE  0.0004
## gcomp_dopt01   EnYd0 for E0Yd0 ML + GLMs not aggressive   G-comp. -0.1298
## IPTW_dopt01    EnYd0 for E0Yd0 ML + GLMs not aggressive      IPTW  0.0002
## IPTW_DR_dopt01 EnYd0 for E0Yd0 ML + GLMs not aggressive   IPTW-DR -0.0009
## TMLE_dopt01    EnYd0 for E0Yd0 ML + GLMs not aggressive      TMLE -0.0011
## CV.TMLE_dopt01 EnYd0 for E0Yd0 ML + GLMs not aggressive   CV-TMLE -0.0009
## gcomp_dopt02   EnYd0 for E0Yd0     ML + GLMs aggressive   G-comp. -0.1180
```

```
## IPTW_dopt02    EnYd0 for E0Yd0     ML + GLMs aggressive      IPTW -0.0006
## IPTW_DR_dopt02 EnYd0 for E0Yd0     ML + GLMs aggressive   IPTW-DR -0.0084
## TMLE_dopt02    EnYd0 for E0Yd0     ML + GLMs aggressive      TMLE -0.0075
## CV.TMLE_dopt02 EnYd0 for E0Yd0     ML + GLMs aggressive   CV-TMLE -0.0001
##                Variance    MSE Coverage
## gcomp_dopt0       3e-04 0.0091        -
## IPTW_dopt0        8e-04 0.0008    95.3%
## IPTW_DR_dopt0     5e-04 0.0005    93.7%
## TMLE_dopt0        5e-04 0.0005    93.7%
## CV.TMLE_dopt0     5e-04 0.0005    93.7%
## gcomp_dopt01      6e-04 0.0175        -
## IPTW_dopt01       8e-04 0.0008    94.7%
## IPTW_DR_dopt01    6e-04 0.0006      94%
## TMLE_dopt01       5e-04 0.0005    93.6%
## CV.TMLE_dopt01    5e-04 0.0005    93.2%
## gcomp_dopt02      6e-04 0.0146        -
## IPTW_dopt02       9e-04 0.0009      94%
## IPTW_DR_dopt02    5e-04 0.0006    90.1%
## TMLE_dopt02       5e-04 0.0006    90.6%
## CV.TMLE_dopt02    5e-04 0.0005    93.6%
##               Comparison                  Library Estimator    Bias Variance
## gcomp    EnYdn for E0Yd0                     GLMs   G-comp. -0.0773   0.0004
## IPTW     EnYdn for E0Yd0                     GLMs      IPTW -0.0558   0.0008
## IPTW_DR  EnYdn for E0Yd0                     GLMs   IPTW-DR -0.0565   0.0006
## TMLE     EnYdn for E0Yd0                     GLMs      TMLE -0.0565   0.0006
## CV.TMLE  EnYdn for E0Yd0                     GLMs   CV-TMLE -0.0764   0.0009
## gcomp1   EnYdn for E0Yd0 ML + GLMs not aggressive   G-comp. -0.1306   0.0007
## IPTW1    EnYdn for E0Yd0 ML + GLMs not aggressive      IPTW  0.0334   0.0010
## IPTW_DR1 EnYdn for E0Yd0 ML + GLMs not aggressive   IPTW-DR  0.0327   0.0008
## TMLE1    EnYdn for E0Yd0 ML + GLMs not aggressive      TMLE  0.0298   0.0008
## CV.TMLE1 EnYdn for E0Yd0 ML + GLMs not aggressive   CV-TMLE -0.0308   0.0007
## gcomp2   EnYdn for E0Yd0     ML + GLMs aggressive   G-comp. -0.1161   0.0007
## IPTW2    EnYdn for E0Yd0     ML + GLMs aggressive      IPTW  0.1236   0.0109
## IPTW_DR2 EnYdn for E0Yd0     ML + GLMs aggressive   IPTW-DR  0.1010   0.0092
## TMLE2    EnYdn for E0Yd0     ML + GLMs aggressive      TMLE  0.1031   0.0108
## CV.TMLE2 EnYdn for E0Yd0     ML + GLMs aggressive   CV-TMLE -0.0316   0.0007
##             MSE Coverage
## gcomp    0.0064        -
## IPTW     0.0039      45%
## IPTW_DR  0.0038    30.1%
## TMLE     0.0038    29.8%
## CV.TMLE  0.0067    14.7%
## gcomp1   0.0178        -
## IPTW1    0.0021    76.1%
## IPTW_DR1 0.0019    66.5%
## TMLE1    0.0016    71.3%
## CV.TMLE1 0.0017      69%
## gcomp2   0.0142        -
## IPTW2    0.0262      31%
## IPTW_DR2 0.0194      33%
## TMLE2    0.0214    33.6%
```

```
## CV.TMLE2 0.0017    68.6%
```

4 Summary of Results Above

- **$E_n[Y_{d_0}]$ to estimate $E_0[Y_{d_0}]$**: these results speak to performance of estimators of $E_0[Y_d]$ for some given d (i.e., not about how well we estimate the rule, but about how well we estimate the performance of a given rule, which here it happens to be d_0). Estimator results:
  - g-comp: biased
    - Note: this differs from estimation of, e.g., $E[Y_1]$ in RCT (or using any treatment rule that isn't a function of covariates), where g-comp using a misspecified glm is a TMLE, and therefore unbiased
  - IPTW: less efficient (more variability), including less efficient than unadjusted
    - Note: this again differs from estimation of, e.g., $E[Y_1]$ in RCT, where IPTW using estimated weights we gain efficiency
  - IPTW-DR and TMLE: unbiased, EXCEPT if Q estimated aggressively, then bias enough for coverage to drop to $\sim 90\%$
    - Small variance gain compared to unadjusted (though suspect this gain would be bigger if covariates were more predictive of outcome?)
  - CV-TMLE: unbiased (even with more aggressive library for Q)
  - Unadjusted: unbiased
    - Small variance price vs the DR estimators, but without risk of bias due to overfitting Q
    - Very little difference compared to CV-TMLE
- **$E_n[Y_{d_n}]$ to estimate $E_0[Y_{d_0}]$**: these results speak to not only how good of a job we do evaluating the rule (i.e., as above), but also how well we estimate the rule
  - None do well. This is all due to fall off in estimating d_0, ie d_n not converging to d_0 fast enough.
    - See ODTR paper just submitted- how to do a better job on d_0 (including at finite sample sizes, even if cant get all the way there, how to get closer)
- **$E_n[Y_{d_n}]$ to estimate $E_0[Y_{d_n}]$**: these results are of interest if going after a data-adaptive target parameter
  - Note: the estimators are targeting different data adaptive parameters. Data-adaptive parameter here for CV-TMLE is the average of the folds, for the others, it is the d_n learned on the whole sample
    - However, here they are pretty similar
  - There is a real price in bias paid by not using sample splitting to evaluate performance
    - For all of the other estimators besides CV-TMLE, will overestimate how well the estimated rule does
    - As the library used to estimate Q gets more aggressive:
      - The estimated rule gets closer to the true rule
      - The price paid (in terms of bias) by not using CV-TMLE increases

Big picture summary:

- For large sample sizes, small price and many benefits to using CV-TMLE with aggressive library to estimate $E_0[Y_{d_0}]$ and $E_0[Y_{d_n}]$
  - But, is this true if truth is simple? If sample size is small, worry is in that case pay a price. We want a method that as sample size increases, goes towards the more complex; when sample size limited, data can't support, will go towards simple (CI-based ODTR might help here)