Next steps for Aim 2

Veronica Cole

Thursday, May 17, 2018

Issue 1: Logistic regression

Last week, we made the determination that Vermunt, rather than BCH, would be our three-step adjustment in the validity analyses. This was done because BCH and Vermunt yielded extremely similar results, except that BCH had a number of issues with convergence. Thus, our next task was to compare estimates of unadjusted logistic regression coefficients to those obtained using Vermunt's adjustment.

Though we will not consider all impact conditions and sample sizes, the following pages have graphs for cases both with and without impact, as well as different training data sample sizes. In general, neither impact nor DIF appears to affect the recovery of parameters, with the exception of the unadjusted logistic regression. However, the Vermunt adjustment does appear to yield significant gains in accuracy.

Issue 2: Detecting a spurious interaction

Because we are running this as a multiple groups latent class analysis, rather than a standard LCA with an interaction term, we need to calculate the interaction term manually. Assuming that frat membership is our predictor, the probability of membership to each class under the current formulation is:

$$P\left(AUD, Study1|Frat\right) = \frac{\exp\left(-\left(\gamma_{01} + \gamma_{11}Frat\right)\right)}{1 + \sum_{k=1}^{3}\exp\left(-\left(\gamma_{0k} + \gamma_{1k}Frat\right)\right)}$$

$$P\left(NoAUD, Study1|Frat\right) = \frac{\exp\left(-\left(\gamma_{02} + \gamma_{12}Frat\right)\right)}{1 + \sum_{k=1}^{3}\exp\left(-\left(\gamma_{0k} + \gamma_{1k}Frat\right)\right)}$$

$$P\left(AUD, Study2|Frat\right) = \frac{\exp\left(-\left(\gamma_{03} + \gamma_{13}Frat\right)\right)}{1 + \sum_{k=1}^{3}\exp\left(-\left(\gamma_{0k} + \gamma_{1k}Frat\right)\right)}$$

$$P\left(NoAUD, Study2|Frat\right) = \frac{1}{1 + \sum_{k=1}^{3}\exp\left(-\left(\gamma_{0k} + \gamma_{1k}Frat\right)\right)}$$

When we last discussed this topic, Dan pointed out that the interaction term (which we will call $\gamma^{\star}$) is given by:

$$\gamma^{\star} = \left(\gamma_{11} - \gamma_{12}\right) - \gamma_{13}$$

We can easily calculate this coefficient by simply reading in the output from Mplus. However, we also need to calculate its standard error. Am I correct that its variance would be:

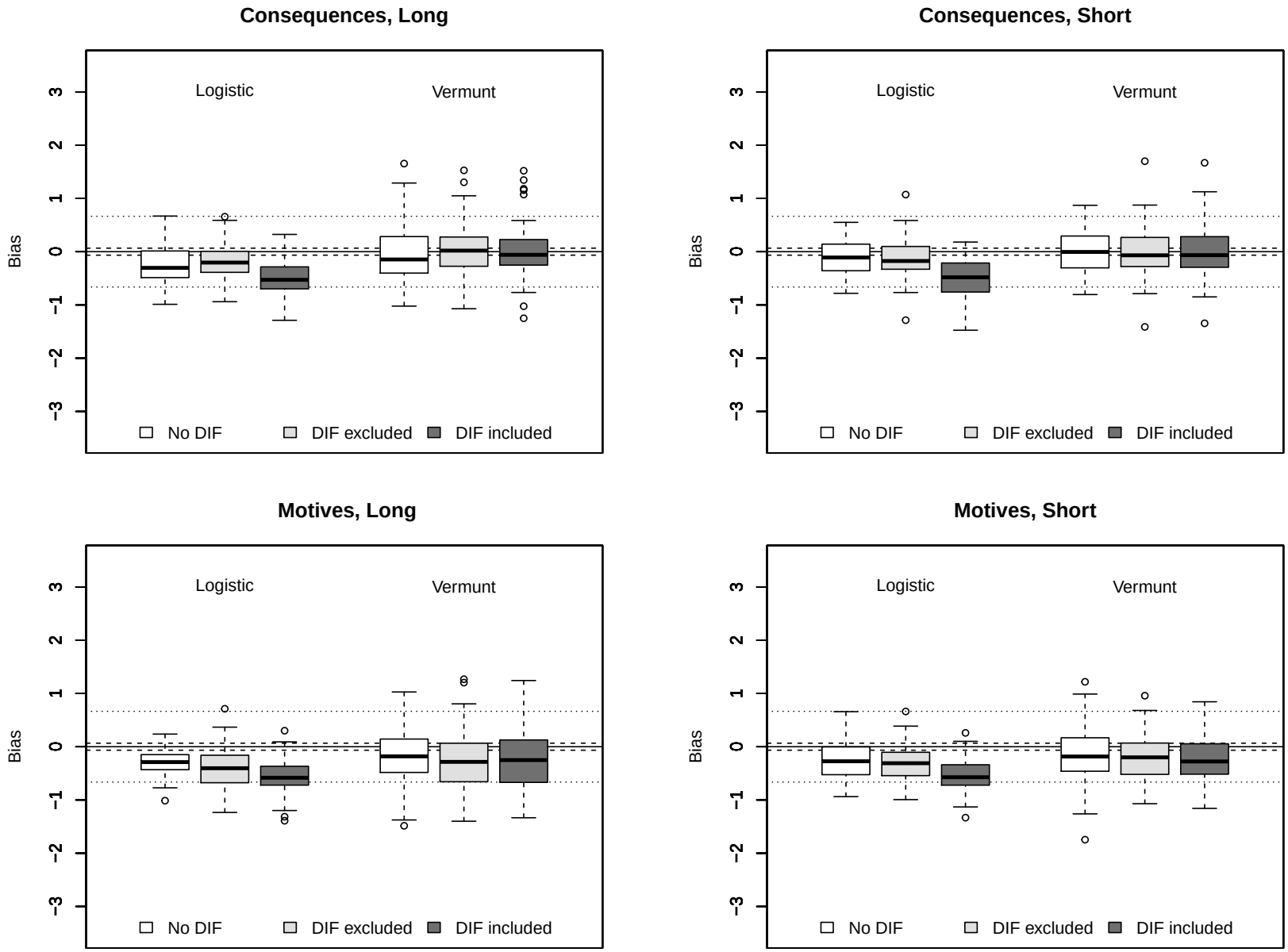
$$Var\left(\gamma^{\star}\right) = Var\left(\gamma_{11}\right) + Var\left(\gamma_{12}\right) + Var\left(\gamma_{13}\right) - 2Cov\left(\gamma_{11}, \gamma_{12}\right) - 2Cov\left(\gamma_{11}, \gamma_{13}\right) - 2Cov\left(\gamma_{12}, \gamma_{13}\right)$$

If so, the standard error can be easily calculated from the TECH3 output, which we already have.

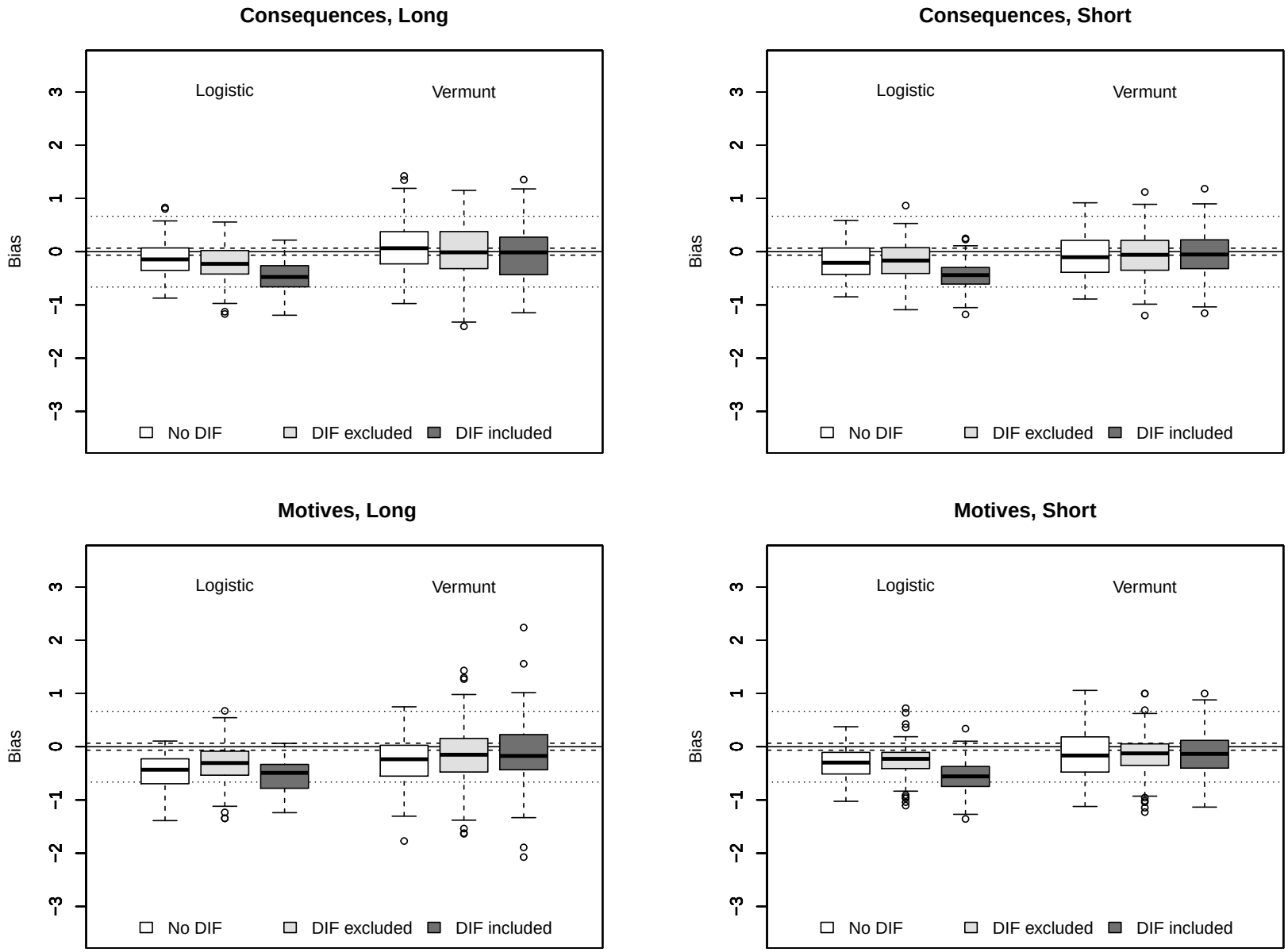
Next Steps

1. Smaller tasks
   (a) Calculate the null hypothesis rejection rate for the interaction term.
      i. Compare to a z distribution?
   (b) Confirm which sample size and impact conditions we would like to show in the bar graphs below.
   (c) Now that Vermunt is working, do we want to try using it to adjust validity analyses for random forest?
2. Bigger tasks
   (a) Rerun multiple group validity analyses with real (i.e., not fully synthetic) data.
   (b) Rerun reliability, both with real and synthetic data, in the multiple groups condition.
   (c) Write the paper.

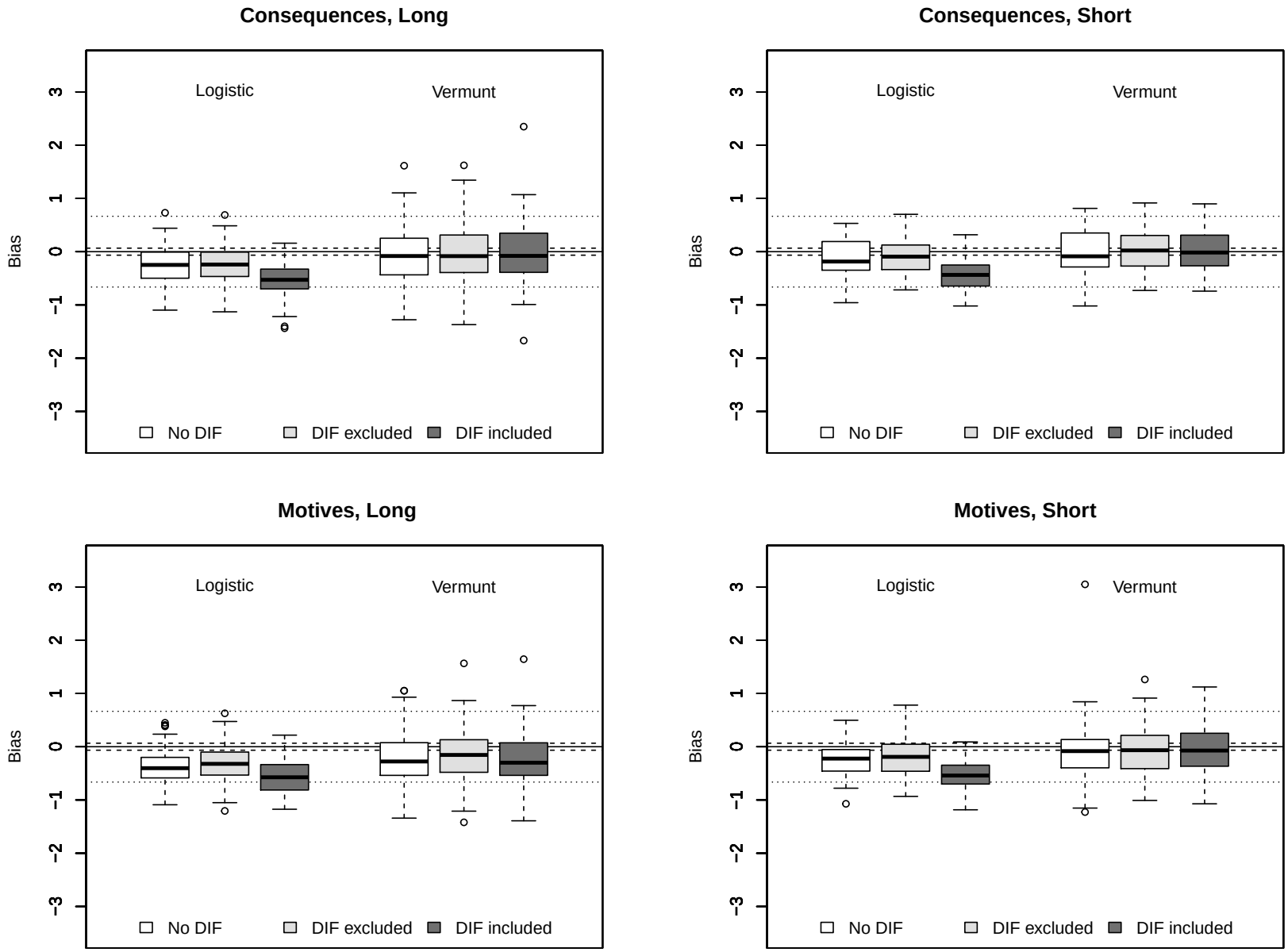
No impact, $N_1 = 25$.



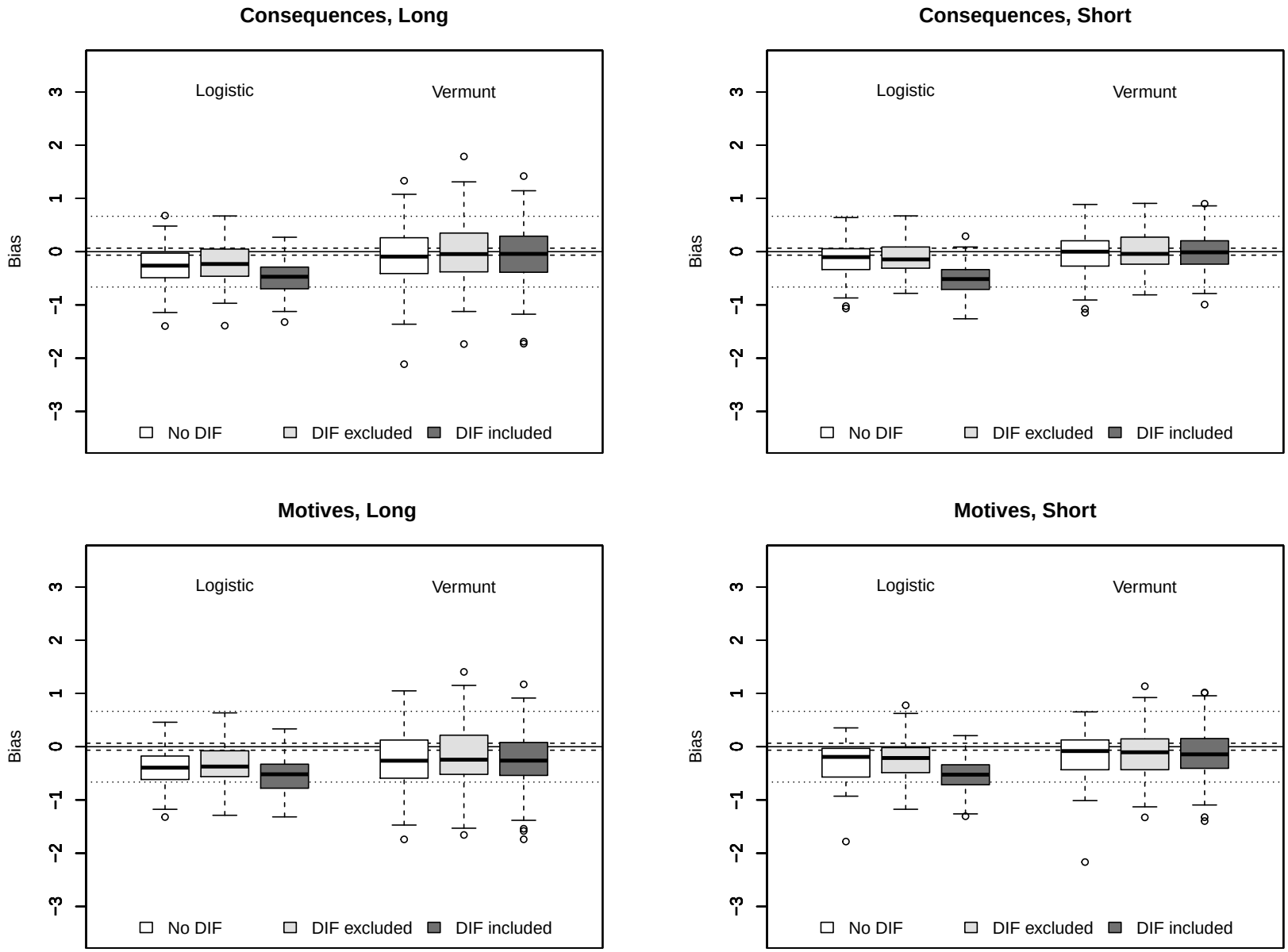
Impact, $N_1 = 25$**.**


Consequences, Long
Logistic
Vermunt
Bias
3
2
1
0
−1
−2
−3
No DIF
DIF excluded
DIF included
Consequences, Short
Logistic
Vermunt
Bias
No DIF
DIF excluded
DIF included
Motives, Long
Logistic
Vermunt
Bias
No DIF
DIF excluded
DIF included
Motives, Short
Logistic
Vermunt
Bias
No DIF
DIF excluded
DIF included

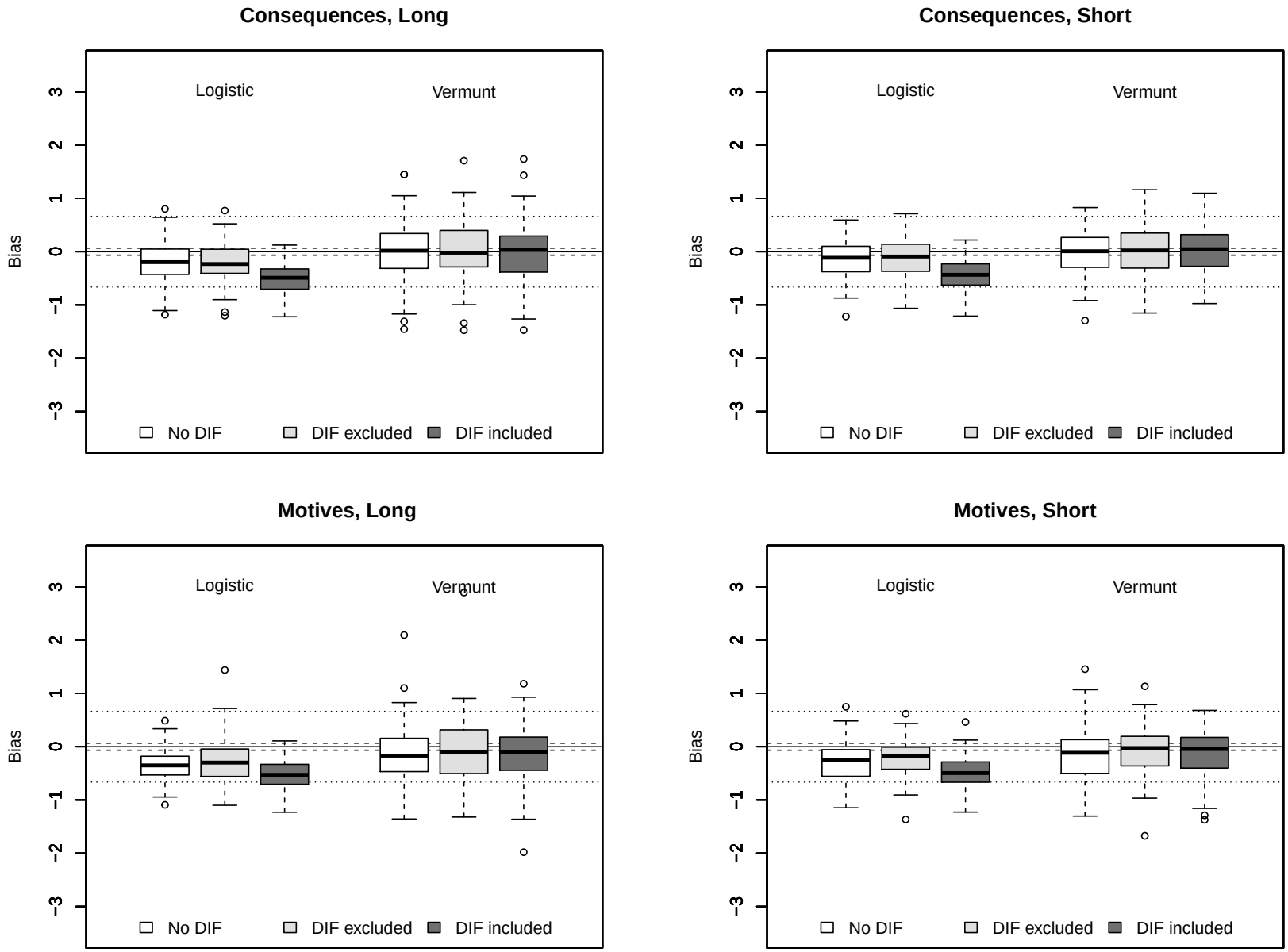
No impact, $N_1 = 50$**.**



Impact, $N_1 = 50$**.**


Consequences, Long
Logistic
Vermunt
Bias
3
2
1
0
−1
−2
−3
No DIF
DIF excluded
DIF included
Consequences, Short
Logistic
Vermunt
Bias
No DIF
DIF excluded
DIF included
Motives, Long
Logistic
Vermunt
Bias
No DIF
DIF excluded
DIF included
Motives, Short
Logistic
Vermunt
Bias
No DIF
DIF excluded
DIF included

No impact, $N_1 = 100$**.**



Impact, $N_1 = 100$**.**