Web Appendix A - Proof of Theorem 1

We first recall Theorem 1 and then provide a proof that closely follows the one from Mao et al (2019).

**Theorem 1.** The estimator \( \hat{\tau}_t^{A-OW} \) is consistent for \( \tau_t^* \) when the treatment model is correctly specified, whether the outcome model is correctly specified or not. When the outcome model is correctly specified, but the treatment model is misspecified, \( \hat{\tau}_t^{A-OW} \) is consistent for

\[
\hat{\tau}_{t,t'} = \frac{\int_{X \in X} \{E[Y(t)|X] - E[Y(t')|X]\} f(X) h(X) \mu(dX)}{\int_{X \in X} f(X) h(X) \mu(dX)},
\]

where \( h(X) \) is the estimand of \( \frac{1}{\sum_{i=1}^k \hat{P}(T = t|X_i)} \) under the misspecified \( \hat{P}(T = t|X_i) \).

**Proof.** First, assume that the treatment model \( \hat{P}(T = t|X_i) \) is correctly specified. Then, the term \( \hat{\tau}_t^{A-OW} \) converges in probability to 0. These other terms can be written as

\[
\left\{ \frac{\sum_{i=1}^n h(X_i) \hat{E}(Y|T = t, X_i)}{\sum_{i=1}^n h(X_i)} - \frac{\sum_{i=1}^n I(T_i = t) \hat{E}(Y|T = t, X_i) w_i}{\sum_{i=1}^n I(T_i = t) w_i} \right\} - \left\{ \frac{\sum_{i=1}^n h(X_i) \hat{E}(Y|T = t', X_i)}{\sum_{i=1}^n h(X_i)} - \frac{\sum_{i=1}^n I(T_i = t') \hat{E}(Y|T = t', X_i) w_i}{\sum_{i=1}^n I(T_i = t') w_i} \right\}
\]

We have

\[
\frac{\sum_{i=1}^n h(X_i) \hat{E}(Y|T = t, X_i)}{\sum_{i=1}^n h(X_i)} \rightarrow P \frac{E[\hat{E}(Y|T = t, X_i) | h(X_i)]}{E[h(X_i)]},
\]

\[
\frac{\sum_{i=1}^n h(X_i) \hat{E}(Y|T = t', X_i)}{\sum_{i=1}^n h(X_i)} \rightarrow P \frac{E[\hat{E}(Y|T = t', X_i) | h(X_i)]}{E[h(X_i)]},
\]

where \( \hat{E}(Y|T, X_i) \) is the estimand of \( E(Y|T, X_i) \). When the outcome model is correctly specified, \( E(Y|T, X_i) = E(Y|T, X_i) \).

The overlap weights benefit from a balancing property, which entails that the distribution of the covariates \( X_i \) is, on average, the same in all groups and is the same as the distribution in the population weighted according to \( h(X_i) \) (Li et al. 2019). As such,

\[
E[I(T = t) w_i \eta(X_i)] = E[I(T = t') w_i \eta(X_i)] = E[h(X_i) \eta(X_i)]
\]

for any function \( \eta(X_i) \). As a result

\[
\frac{\sum_{i=1}^n I(T_i = t) \hat{E}(Y|T = t, X_i) w_i}{\sum_{i=1}^n I(T_i = t) w_i} \rightarrow P \frac{E[\hat{E}(Y|T = t, X_i) | h(X_i)]}{E[h(X_i)]},
\]

\[
\frac{\sum_{i=1}^n I(T_i = t') \hat{E}(Y|T = t', X_i) w_i}{\sum_{i=1}^n I(T_i = t') w_i} \rightarrow P \frac{E[\hat{E}(Y|T = t', X_i) | h(X_i)]}{E[h(X_i)]}.\]

Hence,

\[
\left\{ \frac{\sum_{i=1}^n h(X_i) \hat{E}(Y|T = t, X_i)}{\sum_{i=1}^n h(X_i)} - \frac{\sum_{i=1}^n I(T_i = t) \hat{E}(Y|T = t, X_i) w_i}{\sum_{i=1}^n I(T_i = t) w_i} \right\} - \left\{ \frac{\sum_{i=1}^n h(X_i) \hat{E}(Y|T = t', X_i)}{\sum_{i=1}^n h(X_i)} - \frac{\sum_{i=1}^n I(T_i = t') \hat{E}(Y|T = t', X_i) w_i}{\sum_{i=1}^n I(T_i = t') w_i} \right\} \rightarrow 0.
\]

This completes the proof the \( \hat{\tau}_t^{A-OW} \) converges in probability to 0 when the treatment model is correctly specified, whether the outcome model is correctly specified or not.
Now, assume that the treatment model is incorrectly specified, but the outcome model is correctly specified. The augmented overlap weight estimator can be written as

$$\hat{\tau}_{t,t'}^{A-O\text{W}} = \frac{\sum_{i=1}^{n} h(X_i) \left[ \hat{E}(Y|T = t, X_i) - \hat{E}(Y|T = t', X_i) \right]}{\sum_{i=1}^{n} h(X_i)}$$

$$+ \frac{\sum_{i=1}^{n} I(T_i = t) \left[ Y_i - \hat{E}(Y|T = t, X_i) \right] w_i}{\sum_{i=1}^{n} I(T_i = t) w_i}$$

$$- \frac{\sum_{i=1}^{n} I(T_i = t') \left[ Y_i - \hat{E}(Y|T = t', X_i) \right] w_i}{\sum_{i=1}^{n} I(T_i = t') w_i}.$$

Recall that $\hat{h}(X_i)$ denote the estimand of $h(X_i)$ under the misspecified outcome model. Using this notation, we can note that

$$\frac{\sum_{i=1}^{n} h(X_i) \left[ \hat{E}(Y|T = t, X_i) - \hat{E}(Y|T = t', X_i) \right]}{\sum_{i=1}^{n} h(X_i)} \xrightarrow{P} \frac{\mathbb{E}_X \left\{ \hat{h}(X) \left[ \mathbb{E}(Y|T = t, X) - \mathbb{E}(Y|T = t', X) \right] \right\}}{\mathbb{E}_X [h(X_i)]} = \hat{\tau}_{t,t'}^{\ast}$$

invoking the common exchangeability, consistency and positivity assumptions. Thus, it remains to show that the second and third term in $\hat{\tau}_{t,t'}^{A-O\text{W}}$ as written above converge in probability to 0. We have

$$\mathbb{E} \left\{ I(T_i = t) \left[ Y_i - \hat{E}(Y|T = t, X_i) \right] w_i \right\}$$

$$= \mathbb{E} \left\{ I(T_i = t) \left[ Y_i - \hat{E}(Y|T = t, X_i) \right] w_i \right\}$$

$$= \mathbb{E}_X \left( \mathbb{E} \left\{ I(T_i = t) \left[ Y_i - \hat{E}(Y|T = t, X_i) \right] w_i | X_i \right\} \right)$$

$$= \mathbb{E}_X \left( \mathbb{E} \left\{ I(T_i = t) w_i | X_i \right\} \mathbb{E}_X \left( \left[ Y_i - \hat{E}(Y|T = t, X_i) \right] | X_i \right) \right),$$

where the first equality is obtained by consistency, the second by the law of total probability, and the third is obtained by exchangeability, $Y \prod w_i | X_i (w_i$ is a surjective function of $X_i)$, $\mathbb{E}(Y|T = t, X_i) \prod T$ and $\mathbb{E}(Y|T = t, X_i) \prod w_i$ (because $\hat{E}(Y|T = t, X_i)$ is already conditional on $T$ and $X_i$). It now suffices to note that

$$\mathbb{E}_X (Y_i | X_i) = \mathbb{E}_X \left[ \hat{E}(Y|T = t, X_i) | X_i \right] = \mathbb{E}[Y^t]$$

to show that the second term converges in probability to 0. The proof for the third term is conducted in the same fashion. This completes the proof that when the outcome model is correctly specified, but the treatment model is incorrectly specified, $\hat{\tau}_{t,t'}^{A-O\text{W}} \xrightarrow{P} \hat{\tau}_{t,t'}^{\ast}$, thus completing the proof of Theorem 1. □
Web Appendix B - Additional simulation results

Figure 1: Distribution of the treatment probabilities according to treatment group and Scenarios with a weak ($T^-$) or a strong ($T^+$) association between covariates and treatment.

Legend: ---: $T = 1$, -- --: $T = 2$ and · · ·: $T = 3$
Table 1: Estimate of the treatment effect in the Scenario with a weak association between covariates and treatment, a weak association between covariates and outcome \((T - Y -)\), and a sample size of 500.

| Implementation | Approach | Bias \(\hat{\tau}_{21}\) | Bias \(\hat{\tau}_{31}\) | Std \(\hat{\tau}_{21}\) | Std \(\hat{\tau}_{31}\) | RMSE \(\hat{\tau}_{21}\) | RMSE \(\hat{\tau}_{31}\) | Coverage IC \(\hat{\tau}_{21}\) | Coverage IC \(\hat{\tau}_{31}\) |
|---------------|----------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| Cor.param.    | Crude    | 0.239            | 0.339            | 0.145            | 0.157            | 0.279            | 0.374            | 0.672            | 0.441            |
|               | stan     | 0.001            | 0.000            | 0.111            | 0.116            | 0.111            | 0.116            | 0.961            | 0.956            |
|               | IPW      | 0.004            | 0.002            | 0.144            | 0.148            | 0.144            | 0.148            | 0.991            | 0.985            |
|               | match    | 0.024            | 0.024            | 0.139            | 0.140            | 0.141            | 0.142            | 0.962            | 0.971            |
|               | BCM      | 0.001            | -0.004           | 0.136            | 0.139            | 0.136            | 0.139            | 0.965            | 0.971            |
|               | TMLE     | 0.001            | -0.001           | 0.115            | 0.119            | 0.115            | 0.119            | 0.959            | 0.959            |
|               | OW       | 0.001            | -0.002           | 0.123            | 0.128            | 0.123            | 0.128            | 0.987            | 0.983            |
|               | A-OV     | 0.001            | -0.000           | 0.113            | 0.118            | 0.113            | 0.118            | 0.946            | 0.939            |
| Inc.param.    | stan     | 0.046            | 0.054            | 0.131            | 0.138            | 0.139            | 0.148            | 0.927            | 0.936            |
|               | IPW      | 0.091            | 0.082            | 0.142            | 0.147            | 0.169            | 0.168            | 0.932            | 0.940            |
|               | match    | 0.041            | 0.038            | 0.136            | 0.144            | 0.142            | 0.149            | 0.949            | 0.955            |
|               | BCM      | 0.031            | 0.020            | 0.135            | 0.143            | 0.138            | 0.144            | 0.955            | 0.956            |
|               | TMLE     | 0.076            | 0.073            | 0.135            | 0.145            | 0.155            | 0.162            | 0.934            | 0.941            |
|               | OW       | 0.082            | 0.073            | 0.135            | 0.139            | 0.158            | 0.158            | 0.941            | 0.947            |
|               | A-OV     | 0.068            | 0.065            | 0.131            | 0.138            | 0.148            | 0.152            | 0.912            | 0.922            |
| M.Learning    | stan     | 0.004            | 0.002            | 0.114            | 0.119            | 0.114            | 0.119            | 0.943            | 0.944            |
|               | IPW      | 0.133            | 0.128            | 0.150            | 0.175            | 0.200            | 0.217            | 0.878            | 0.850            |
|               | match*   | 0.066            | 0.059            | 0.158            | 0.181            | 0.171            | 0.190            | 0.922            | 0.902            |
|               | BCM*     | -0.004           | -0.004           | 0.136            | 0.143            | 0.136            | 0.143            | 0.963            | 0.958            |
|               | TMLE     | 0.006            | 0.004            | 0.116            | 0.121            | 0.116            | 0.121            | 0.960            | 0.955            |
|               | OW       | 0.125            | 0.121            | 0.146            | 0.172            | 0.192            | 0.211            | 0.881            | 0.854            |
|               | A-OV     | 0.005            | 0.004            | 0.115            | 0.120            | 0.115            | 0.120            | 0.948            | 0.941            |

Legend: Cor.param=correct parametric models, Inc.param=incorrect parametric models, M.Learning=machine learning, Crude=Unadjusted, stan=standardization, IPW=inverse probability weighting, match=matching, BCM=bias-corrected matching, TMLE=targeted maximum likelihood, OW=overlap weights, A-OV=augmented overlap weights. *: 137 replications were discarded due to errors when running the code for match and BCM.
Table 2: Estimate of the treatment effect in the Scenario with a weak association between covariates and treatment, a weak association between covariates and outcome ($T - Y -$), and a sample size of 2000.

| Implementation | Approach | Bias $\hat{\tau}_{21}$ | Bias $\hat{\tau}_{31}$ | Std $\hat{\tau}_{21}$ | Std $\hat{\tau}_{31}$ | RMSE $\hat{\tau}_{21}$ | RMSE $\hat{\tau}_{31}$ | Coverage IC $\hat{\tau}_{21}$ | Coverage IC $\hat{\tau}_{31}$ |
|----------------|----------|--------------------------|--------------------------|------------------------|------------------------|------------------------|------------------------|--------------------------|------------------------|
| Cor.param.     | Crude    | 0.243 0.345              | 0.075 0.080              | 0.254 0.355            | 0.107 0.006            |                       |                       |                          |                       |
|                | stan     | 0.001 0.003              | 0.058 0.058              | 0.058 0.058            | 0.949 0.963            |                       |                       |                          |                       |
|                | IPW      | 0.001 0.001              | 0.073 0.070              | 0.073 0.070            | 0.989 0.995            |                       |                       |                          |                       |
|                | match    | 0.012 0.011              | 0.070 0.068              | 0.071 0.069            | 0.955 0.960            |                       |                       |                          |                       |
|                | BCM      | 0.002 0.001              | 0.070 0.068              | 0.070 0.068            | 0.953 0.963            |                       |                       |                          |                       |
|                | TMLE     | 0.001 0.002              | 0.059 0.059              | 0.059 0.059            | 0.961 0.958            |                       |                       |                          |                       |
|                | OW       | 0.001 0.001              | 0.065 0.064              | 0.065 0.064            | 0.982 0.988            |                       |                       |                          |                       |
|                | A-OW     | 0.001 0.002              | 0.059 0.058              | 0.059 0.059            | 0.941 0.943            |                       |                       |                          |                       |
| Inc.param.     | stan     | 0.047 0.058              | 0.070 0.069              | 0.084 0.089            | 0.887 0.867            |                       |                       |                          |                       |
|                | IPW      | 0.093 0.088              | 0.075 0.073              | 0.119 0.114            | 0.810 0.839            |                       |                       |                          |                       |
|                | match    | 0.024 0.024              | 0.069 0.069              | 0.073 0.073            | 0.941 0.955            |                       |                       |                          |                       |
|                | BCM      | 0.019 0.016              | 0.068 0.068              | 0.071 0.070            | 0.955 0.959            |                       |                       |                          |                       |
|                | TMLE     | 0.077 0.077              | 0.072 0.072              | 0.105 0.105            | 0.826 0.855            |                       |                       |                          |                       |
|                | OW       | 0.084 0.079              | 0.072 0.070              | 0.111 0.105            | 0.826 0.860            |                       |                       |                          |                       |
|                | A-OW     | 0.069 0.070              | 0.070 0.069              | 0.098 0.098            | 0.814 0.830            |                       |                       |                          |                       |
| M.Learning     | stan     | 0.001 0.004              | 0.060 0.059              | 0.060 0.059            | 0.931 0.941            |                       |                       |                          |                       |
|                | IPW      | 0.086 0.072              | 0.075 0.072              | 0.114 0.102            | 0.836 0.892            |                       |                       |                          |                       |
|                | match    | 0.025 0.014              | 0.072 0.069              | 0.077 0.071            | 0.950 0.964            |                       |                       |                          |                       |
|                | BCM      | -0.004 -0.003            | 0.069 0.068              | 0.069 0.068            | 0.963 0.966            |                       |                       |                          |                       |
|                | TMLE     | 0.002 0.004              | 0.061 0.060              | 0.061 0.060            | 0.950 0.964            |                       |                       |                          |                       |
|                | OW       | 0.078 0.064              | 0.072 0.069              | 0.106 0.094            | 0.853 0.906            |                       |                       |                          |                       |
|                | A-OW     | 0.001 0.004              | 0.060 0.059              | 0.060 0.060            | 0.934 0.947            |                       |                       |                          |                       |

Legend: Cor.param=correct parametric models, Inc.param=incorrect parametric models, M.Learning=machine learning, Crude=Unadjusted, stan=standardization, IPW=inverse probability weighting, match=matching, BCM=bias-corrected matching, TMLE=targeted maximum likelihood, OW=overlap weights, A-OW=augmented overlap weights
Table 3: Estimate of the treatment effect in the Scenario with a strong association between covariates and treatment, a weak association between covariates and outcome ($T + Y -$), and a sample size of 500.

| Implementation | Approach | Bias | Std | RMSE | Coverage IC |
|----------------|----------|------|-----|------|-------------|
|                |          | $\hat{\tau}_{21}$ | $\hat{\tau}_{31}$ | $\hat{\tau}_{21}$ | $\hat{\tau}_{31}$ | $\hat{\tau}_{21}$ | $\hat{\tau}_{31}$ |
|                |          | $\hat{\tau}_{21}$ | $\hat{\tau}_{31}$ | $\hat{\tau}_{21}$ | $\hat{\tau}_{31}$ | $\hat{\tau}_{21}$ | $\hat{\tau}_{31}$ |
|                | Crude    | 0.461 | 0.791 | 0.143 | 0.171 | 0.482 | 0.810 | 0.106 | 0.007 |
|                | stan     | -0.002 | -0.005 | 0.117 | 0.133 | 0.117 | 0.133 | 0.952 | 0.959 |
|                | IPW      | 0.052 | 0.055 | 0.360 | 0.384 | 0.363 | 0.388 | 0.874 | 0.889 |
|                | match    | 0.125 | 0.129 | 0.170 | 0.199 | 0.211 | 0.237 | 0.885 | 0.887 |
|                | BCM      | 0.018 | 0.006 | 0.166 | 0.195 | 0.167 | 0.195 | 0.948 | 0.939 |
|                | TMLE     | -0.000 | 0.003 | 0.145 | 0.181 | 0.145 | 0.181 | 0.949 | 0.935 |
|                | OW       | 0.004 | 0.001 | 0.153 | 0.168 | 0.153 | 0.168 | 0.952 | 0.961 |
|                | A-OW     | -0.001 | 0.001 | 0.138 | 0.155 | 0.138 | 0.155 | 0.944 | 0.946 |
|                | stan     | -0.132 | -0.046 | 0.145 | 0.160 | 0.196 | 0.167 | 0.841 | 0.929 |
|                | IPW      | 0.209 | 0.187 | 0.229 | 0.268 | 0.310 | 0.327 | 0.721 | 0.816 |
|                | match    | 0.153 | 0.156 | 0.169 | 0.198 | 0.228 | 0.252 | 0.850 | 0.873 |
|                | BCM      | 0.097 | 0.055 | 0.164 | 0.195 | 0.190 | 0.202 | 0.921 | 0.929 |
|                | TMLE     | 0.152 | 0.110 | 0.189 | 0.229 | 0.243 | 0.254 | 0.872 | 0.914 |
|                | OW       | 0.112 | 0.097 | 0.161 | 0.173 | 0.196 | 0.198 | 0.878 | 0.898 |
|                | A-OW     | 0.059 | 0.059 | 0.148 | 0.168 | 0.159 | 0.177 | 0.919 | 0.930 |
|                | stan     | -0.011 | -0.003 | 0.128 | 0.149 | 0.129 | 0.149 | 0.938 | 0.941 |
|                | IPW      | 0.176 | 0.163 | 0.237 | 0.269 | 0.295 | 0.314 | 0.768 | 0.845 |
|                | match*   | 0.145 | 0.149 | 0.170 | 0.195 | 0.223 | 0.246 | 0.863 | 0.880 |
|                | BCM*     | 0.006 | 0.014 | 0.162 | 0.191 | 0.162 | 0.191 | 0.940 | 0.948 |
|                | TMLE     | 0.011 | 0.010 | 0.149 | 0.186 | 0.149 | 0.186 | 0.940 | 0.928 |
|                | OW       | 0.085 | 0.074 | 0.158 | 0.170 | 0.180 | 0.186 | 0.911 | 0.916 |
|                | A-OW**   | -0.005 | -0.003 | 0.141 | 0.156 | 0.141 | 0.156 | 0.933 | 0.938 |

Legend: Cor.param=correct parametric models, Inc.param=incorrect parametric models, M.Learning=machine learning, Crude=Unadjusted, stan=standardization, IPW=inverse probability weighting, match=matching, BCM=bias-corrected matching, TMLE=targeted maximum likelihood, OW=overlap weights, A-OW=augmented overlap weights. *: 2 replications were discarded due to errors when running the code for match and BCM, **: in 9 replications, the confidence intervals could not be computed.
Table 4: Estimate of the treatment effect in the Scenario with a strong association between covariates and treatment, a weak association between covariates and outcome ($T + Y$), and a sample size of 2000.

| Implementation Approach | Bias $\hat{\tau}_{21}$ | Std $\hat{\tau}_{31}$ | RMSE $\hat{\tau}_{21}$ | Coverage IC $\hat{\tau}_{31}$ |
|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| Cor.param.              |                         |                         |                         |                         |
| Crude                   | 0.461 0.790             | 0.067 0.084             | 0.465 0.795             | 0.000 0.000             |
| stan                    | -0.000 0.000             | 0.059 0.068             | 0.059 0.068             | 0.957 0.958             |
| IPW                     | 0.012 0.014             | 0.328 0.332             | 0.328 0.332             | 0.854 0.889             |
| match                   | 0.075 0.072             | 0.088 0.101             | 0.116 0.124             | 0.861 0.896             |
| BCM                     | 0.015 0.010             | 0.086 0.099             | 0.087 0.099             | 0.956 0.952             |
| TMLE                    | 0.000 0.002             | 0.073 0.085             | 0.073 0.085             | 0.951 0.953             |
| OW                      | -0.000 0.000             | 0.072 0.081             | 0.072 0.081             | 0.975 0.970             |
| A-OW                    | -0.001 0.000             | 0.067 0.074             | 0.067 0.074             | 0.946 0.955             |
| Inc.param.              |                         |                         |                         |                         |
| stan                    | -0.129 -0.041            | 0.073 0.083             | 0.148 0.093             | 0.530 0.897             |
| IPW                     | 0.202 0.183             | 0.128 0.158             | 0.239 0.242             | 0.450 0.646             |
| match                   | 0.103 0.097             | 0.087 0.100             | 0.135 0.139             | 0.785 0.842             |
| BCM                     | 0.071 0.043             | 0.084 0.098             | 0.110 0.107             | 0.873 0.939             |
| TMLE                    | 0.151 0.112             | 0.099 0.117             | 0.181 0.162             | 0.615 0.830             |
| OW                      | 0.109 0.096             | 0.079 0.085             | 0.134 0.128             | 0.708 0.815             |
| A-OW                    | 0.057 0.059             | 0.072 0.083             | 0.092 0.102             | 0.865 0.884             |
| M.Learning              |                         |                         |                         |                         |
| stan                    | 0.001 0.017             | 0.064 0.076             | 0.064 0.077             | 0.940 0.931             |
| IPW                     | 0.101 0.096             | 0.198 0.204             | 0.222 0.225             | 0.728 0.808             |
| match                   | 0.089 0.084             | 0.089 0.101             | 0.126 0.131             | 0.813 0.862             |
| BCM                     | 0.002 0.007             | 0.086 0.098             | 0.086 0.099             | 0.953 0.956             |
| TMLE                    | 0.007 0.007             | 0.076 0.090             | 0.076 0.090             | 0.942 0.947             |
| OW                      | 0.043 0.039             | 0.080 0.084             | 0.091 0.093             | 0.908 0.937             |
| A-OW*                   | -0.005 -0.003            | 0.068 0.076             | 0.068 0.076             | 0.952 0.949             |

Legend: Cor.param=correct parametric models, Inc.param=incorrect parametric models, M.Learning=machine learning, Crude=Unadjusted, stan=standardization, IPW=inverse probability weighting, match=matching, BCM=bias-corrected matching, TMLE=targeted maximum likelihood, OW=overlap weights, A-OW=augmented overlap weights. *: in 1 replications, the confidence intervals could not be computed.
Table 5: Estimate of the treatment effect in the Scenario with a weak association between covariates and treatment, a strong association between covariates and outcome \((T - Y^+)\), and a sample size of 500.

| Implementation | Approach | 0.560 | 0.803 | 0.241 | 0.269 | 0.609 | 0.847 | 0.443 | 0.166 |
|----------------|----------|-------|-------|-------|-------|-------|-------|-------|-------|
| Cor.param.     | Crude    | 0.001 | 0.000 | 0.111 | 0.116 | 0.111 | 0.116 | 0.983 | 0.984 |
|                | stan     | 0.007 | 0.004 | 0.220 | 0.225 | 0.220 | 0.225 | 0.998 | 0.998 |
|                | match    | 0.052 | 0.057 | 0.162 | 0.164 | 0.170 | 0.173 | 0.984 | 0.985 |
|                | BCM      | 0.001 | -0.006| 0.149 | 0.153 | 0.149 | 0.153 | 0.990 | 0.999 |
|                | TMLE     | 0.001 | -0.001| 0.115 | 0.119 | 0.115 | 0.119 | 0.959 | 0.959 |
|                | OW       | 0.000 | -0.003| 0.153 | 0.159 | 0.153 | 0.159 | 0.999 | 0.997 |
|                | A-OW     | 0.001 | -0.000| 0.113 | 0.118 | 0.113 | 0.118 | 0.946 | 0.939 |
| Inc.param.     | stan     | 0.046 | 0.054 | 0.131 | 0.138 | 0.139 | 0.148 | 0.927 | 0.936 |
|                | IPW      | 0.091 | 0.082 | 0.142 | 0.147 | 0.169 | 0.168 | 0.932 | 0.940 |
|                | match    | 0.041 | 0.038 | 0.136 | 0.144 | 0.142 | 0.149 | 0.949 | 0.955 |
|                | BCM      | 0.031 | 0.020 | 0.135 | 0.143 | 0.138 | 0.144 | 0.955 | 0.956 |
|                | TMLE     | 0.076 | 0.073 | 0.135 | 0.145 | 0.155 | 0.162 | 0.934 | 0.941 |
|                | OW       | 0.082 | 0.073 | 0.135 | 0.139 | 0.158 | 0.158 | 0.941 | 0.947 |
|                | A-OW     | 0.068 | 0.065 | 0.131 | 0.138 | 0.148 | 0.152 | 0.912 | 0.922 |
| M.Learning     | stan     | 0.004 | 0.002 | 0.114 | 0.119 | 0.114 | 0.119 | 0.943 | 0.944 |
|                | IPW      | 0.133 | 0.128 | 0.150 | 0.175 | 0.200 | 0.217 | 0.878 | 0.850 |
|                | match*   | 0.077 | 0.076 | 0.160 | 0.186 | 0.178 | 0.201 | 0.913 | 0.895 |
|                | BCM*     | -0.001 | 0.003 | 0.139 | 0.150 | 0.139 | 0.150 | 0.962 | 0.957 |
|                | TMLE     | 0.006 | 0.004 | 0.116 | 0.121 | 0.116 | 0.121 | 0.960 | 0.955 |
|                | OW       | 0.125 | 0.121 | 0.146 | 0.172 | 0.192 | 0.211 | 0.881 | 0.854 |
|                | A-OW     | 0.005 | 0.004 | 0.115 | 0.120 | 0.115 | 0.120 | 0.948 | 0.941 |

Legend: Cor.param=correct parametric models, Inc.param=incorrect parametric models, M.Learning=machine learning, Crude=Unadjusted, stan=standardization, IPW=inverse probability weighting, match=matching, BCM=bias-corrected matching, TMLE=targeted maximum likelihood, OW=overlap weights, A-OW=augmented overlap weights.*: 2 replications were discarded due to errors when running the code for match and BCM.
Table 6: Estimate of the treatment effect in the Scenario with a weak association between covariates and treatment, a strong association between covariates and outcome (\(T - Y^+\)), and a sample size of 2000.

| Implementation | Approach | Bias \(\hat{\tau}_{21}\) | Bias \(\hat{\tau}_{31}\) | Std \(\hat{\tau}_{21}\) | Std \(\hat{\tau}_{31}\) | RMSE \(\hat{\tau}_{21}\) | RMSE \(\hat{\tau}_{31}\) | Coverage IC |
|----------------|----------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|--------------|
|                |          | 0.572                | 0.815                | 0.122                | 0.138                | 0.584                | 0.827                | 0.002        |
|                |          | 0.001                | 0.003                | 0.058                | 0.058                | 0.058                | 0.058                | 0.982        |
|                |          | 0.002                | 0.001                | 0.106                | 0.102                | 0.106                | 0.102                | 0.998        |
|                |          | 0.024                | 0.023                | 0.077                | 0.075                | 0.081                | 0.078                | 0.983        |
| Cor.param.     |          | 0.003                | -0.001               | 0.075                | 0.073                | 0.076                | 0.073                | 0.986        |
|                |          | 0.001                | 0.002                | 0.059                | 0.059                | 0.059                | 0.059                | 0.961        |
|                |          | 0.001                | -0.001               | 0.080                | 0.079                | 0.080                | 0.079                | 1.000        |
|                |          | 0.024                | 0.023                | 0.077                | 0.075                | 0.081                | 0.078                | 0.983        |
|                |          | 0.003                | 0.002                | 0.059                | 0.058                | 0.059                | 0.058                | 0.941        |
|                | stan     | 0.047                | 0.058                | 0.070                | 0.069                | 0.084                | 0.089                | 0.887        |
| Inc.param.     |          | 0.093                | 0.088                | 0.075                | 0.073                | 0.119                | 0.114                | 0.810        |
|                |          | 0.024                | 0.024                | 0.069                | 0.069                | 0.073                | 0.073                | 0.941        |
|                |          | 0.019                | 0.016                | 0.068                | 0.068                | 0.071                | 0.070                | 0.955        |
|                |          | 0.077                | 0.077                | 0.072                | 0.072                | 0.105                | 0.105                | 0.826        |
|                |          | 0.084                | 0.079                | 0.072                | 0.070                | 0.111                | 0.105                | 0.826        |
|                |          | 0.069                | 0.070                | 0.070                | 0.069                | 0.098                | 0.098                | 0.814        |
|                |          | 0.001                | 0.004                | 0.060                | 0.059                | 0.060                | 0.059                | 0.931        |
| M.Learning     | stan     | 0.086                | 0.072                | 0.075                | 0.072                | 0.114                | 0.102                | 0.836        |
|                |          | 0.025                | 0.014                | 0.072                | 0.069                | 0.077                | 0.071                | 0.950        |
|                |          | -0.004               | -0.003               | 0.069                | 0.068                | 0.069                | 0.068                | 0.963        |
|                |          | 0.002                | 0.004                | 0.061                | 0.060                | 0.061                | 0.060                | 0.950        |
|                |          | 0.078                | 0.064                | 0.072                | 0.069                | 0.106                | 0.094                | 0.853        |
|                |          | 0.001                | 0.004                | 0.060                | 0.059                | 0.060                | 0.060                | 0.934        |

Legend: Cor.param=correct parametric models, Inc.param=incorrect parametric models, M.Learning=machine learning, Crude=Unadjusted, stan=standardization, IPW=inverse probability weighting, match=matching, BCM=bias-corrected matching, TMLE=targeted maximum likelihood, OW=overlap weights, A-OW=augmented overlap weights
Table 7: Estimate of the treatment effect in the Scenario with a strong association between covariates and treatment, a strong association between covariates and outcome \((T + Y^+}\), and a sample size of 500.

| Implementation | Approach | Bias | Std | RMSE | Coverage IC |
|----------------|----------|------|-----|------|-------------|
|                |          | \(\hat{\tau}_{21}\) | \(\hat{\tau}_{31}\) | \(\hat{\tau}_{21}\) | \(\hat{\tau}_{31}\) | \(\hat{\tau}_{21}\) | \(\hat{\tau}_{31}\) |
| Cor.param.     | Crude    | 1.174 | 1.947 | 0.226 | 0.292 | 1.195 | 1.969 | 0.001 | 0.000 |
|                | stan     | -0.002 | -0.005 | 0.117 | 0.133 | 0.117 | 0.133 | 0.979 | 0.982 |
|                | IPW      | 0.117 | 0.128 | 0.693 | 0.713 | 0.703 | 0.725 | 0.841 | 0.868 |
|                | match    | 0.274 | 0.294 | 0.204 | 0.232 | 0.341 | 0.375 | 0.801 | 0.820 |
|                | BCM      | 0.036 | 0.007 | 0.182 | 0.210 | 0.185 | 0.210 | 0.966 | 0.970 |
|                | TMLE     | -0.000 | 0.003 | 0.145 | 0.181 | 0.145 | 0.181 | 0.949 | 0.932 |
|                | OW       | 0.009 | 0.003 | 0.191 | 0.206 | 0.191 | 0.206 | 0.980 | 0.985 |
|                | A-OV     | -0.001 | 0.001 | 0.138 | 0.155 | 0.138 | 0.155 | 0.944 | 0.946 |
| Inc.param.     | stan     | -0.132 | -0.046 | 0.145 | 0.160 | 0.196 | 0.167 | 0.841 | 0.929 |
|                | IPW      | 0.209 | 0.187 | 0.229 | 0.268 | 0.310 | 0.327 | 0.721 | 0.816 |
|                | match    | 0.153 | 0.156 | 0.169 | 0.198 | 0.228 | 0.252 | 0.850 | 0.873 |
|                | BCM      | 0.097 | 0.055 | 0.164 | 0.195 | 0.190 | 0.202 | 0.921 | 0.929 |
|                | TMLE     | 0.152 | 0.110 | 0.189 | 0.229 | 0.243 | 0.254 | 0.872 | 0.914 |
|                | OW       | 0.112 | 0.097 | 0.161 | 0.173 | 0.196 | 0.198 | 0.878 | 0.898 |
|                | A-OV     | 0.059 | 0.059 | 0.148 | 0.168 | 0.159 | 0.177 | 0.919 | 0.930 |
| M.Learning     | stan     | -0.011 | -0.003 | 0.128 | 0.149 | 0.129 | 0.149 | 0.938 | 0.941 |
|                | IPW      | 0.176 | 0.163 | 0.237 | 0.269 | 0.295 | 0.314 | 0.768 | 0.845 |
|                | match*   | 0.145 | 0.149 | 0.170 | 0.195 | 0.223 | 0.246 | 0.863 | 0.880 |
|                | BCM*     | 0.066 | 0.014 | 0.162 | 0.191 | 0.162 | 0.191 | 0.940 | 0.948 |
|                | TMLE     | 0.011 | 0.010 | 0.149 | 0.186 | 0.149 | 0.186 | 0.940 | 0.928 |
|                | OW       | 0.085 | 0.074 | 0.158 | 0.170 | 0.180 | 0.186 | 0.911 | 0.916 |
|                | A-OV**   | -0.005 | -0.003 | 0.141 | 0.156 | 0.141 | 0.156 | 0.933 | 0.938 |

Legend: Cor.param=correct parametric models, Inc.param=incorrect parametric models, M.Learning=machine learning, Crude=Unadjusted, stan=standardization, IPW=inverse probability weighting, match=matching, BCM=bias-corrected matching, TMLE=targeted maximum likelihood, OW=overlap weights, A-OV=augmented overlap weights.*: 3 replications were discarded due to errors when running the code for match and BCM, **: in 9 replications, the confidence intervals could not be computed.
Table 8: Estimate of the treatment effect in the Scenario with a strong association between covariates and treatment, a strong association between covariates and outcome ($T + Y +$), and a sample size of 2000.

| Implementation | Approach | Bias | Std | RMSE | Coverage IC |
|----------------|----------|------|-----|------|-------------|
|                |          | $\hat{\tau}_{21}$ | $\hat{\tau}_{31}$ | $\hat{\tau}_{21}$ | $\hat{\tau}_{31}$ | $\hat{\tau}_{21}$ | $\hat{\tau}_{31}$ |
| Cor.param.     | Crude    | 1.172 | 1.939 | 0.105 | 0.141 | 1.177 | 1.944 | 0.000 | 0.000 |
|                | stan     | -0.000 | 0.000 | 0.059 | 0.068 | 0.059 | 0.068 | 0.984 | 0.984 |
|                | IPW      | 0.024 | 0.027 | 0.708 | 0.709 | 0.708 | 0.710 | 0.835 | 0.862 |
|                | match    | 0.159 | 0.162 | 0.103 | 0.116 | 0.189 | 0.199 | 0.727 | 0.770 |
|                | BCM      | 0.029 | 0.016 | 0.094 | 0.107 | 0.098 | 0.108 | 0.978 | 0.978 |
|                | TMLE     | 0.001 | 0.002 | 0.073 | 0.085 | 0.073 | 0.085 | 0.947 | 0.953 |
|                | OW       | 0.001 | 0.000 | 0.090 | 0.101 | 0.090 | 0.101 | 0.988 | 0.990 |
|                | A-OW     | -0.001 | 0.000 | 0.067 | 0.074 | 0.067 | 0.074 | 0.946 | 0.955 |
| Inc.param.     | stan     | -0.129 | -0.041 | 0.073 | 0.083 | 0.148 | 0.093 | 0.530 | 0.897 |
|                | IPW      | 0.202 | 0.183 | 0.128 | 0.158 | 0.239 | 0.242 | 0.450 | 0.646 |
|                | match    | 0.103 | 0.097 | 0.087 | 0.100 | 0.135 | 0.139 | 0.785 | 0.842 |
|                | BCM      | 0.071 | 0.043 | 0.084 | 0.098 | 0.110 | 0.107 | 0.873 | 0.939 |
|                | TMLE     | 0.151 | 0.112 | 0.099 | 0.117 | 0.181 | 0.162 | 0.615 | 0.830 |
|                | OW       | 0.109 | 0.096 | 0.079 | 0.085 | 0.134 | 0.128 | 0.708 | 0.815 |
|                | A-OW     | 0.057 | 0.059 | 0.072 | 0.083 | 0.092 | 0.102 | 0.865 | 0.884 |
| M.Learning     | stan     | 0.001 | 0.017 | 0.064 | 0.076 | 0.064 | 0.077 | 0.940 | 0.931 |
|                | IPW      | 0.101 | 0.096 | 0.198 | 0.204 | 0.222 | 0.225 | 0.728 | 0.808 |
|                | match    | 0.089 | 0.084 | 0.089 | 0.101 | 0.126 | 0.131 | 0.813 | 0.862 |
|                | BCM      | 0.002 | 0.007 | 0.086 | 0.098 | 0.086 | 0.099 | 0.953 | 0.956 |
|                | TMLE     | 0.007 | 0.007 | 0.076 | 0.090 | 0.076 | 0.090 | 0.942 | 0.947 |
|                | OW       | 0.043 | 0.039 | 0.080 | 0.084 | 0.091 | 0.093 | 0.908 | 0.937 |
|                | A-OW     | -0.005 | -0.003 | 0.068 | 0.076 | 0.068 | 0.076 | 0.952 | 0.949 |

Legend: Cor.param=correct parametric models, Inc.param=incorrect parametric models, M.Learning=machine learning, Crude=Unadjusted, stan=standardization, IPW=inverse probability weighting, match=matching, BCM=bias-corrected matching, TMLE=targeted maximum likelihood, OW=overlap weights, A-OW=augmented overlap weights. *: in 1 replications, the confidence intervals could not be computed.
Figure 2: Ratio of the mean estimated standard error to the standard deviation of the estimates of $\tau_{21}$ for the correct parametric implementation according to approach (rows), sample size (dark gray = 500, gray = 1000, light gray = 2000), the strength of the association between covariates and treatment (columns; weak=$T^-$ or strong=$T^+$) and between the covariates and outcome (weak=$Y^-$ or strong=$Y^+$), and implementation (Cor = Correct parametric, Inc = Incorrect parametric, ML = machine learning).
Figure 3: Ratio of the mean estimated standard error to the standard deviation of the estimates of the plasmode simulations, according to approach (rows), model parameter (columns) and implementation (dark gray = parametric, light gray = machine learning)
Figure 4: Distribution of the “true” treatment probabilities according to treatment group in the complete sample.
Table 9: Characteristics of lightning-caused wildfires in Alberta, Canada, between 1996 and 2004 in provincial or public lands, according to initial intervention for suppressing the fire.

| Year       | Fire discovery | Air patrol | Lookout | Unplanned | Ecological region | Fuel type | PM Period | Month | Response time, mean(sd) | Active fires, mean(sd) | ln(initial size), mean(sd) |
|------------|----------------|------------|---------|-----------|-------------------|-----------|-----------|-------|------------------------|-------------------------|------------------------|
| 2003       | 248 (8.2)      | 61 (6.6)   | 38 (8.1) | 57 (10.6) | 738 (24.5)        | 753 (24.5)| 1061 (35.2)| 2077 (68.8) | 453 (18.0) | 422 (14.0) | 2646 (87.0) | 2939.45 (918.46) | 59.31 (61.56) | -2.84 (1.96) |
| 2004       | 277 (9.2)      | 151 (16.4) | 47 (10.1) | 82 (15.2) | 177 (7.5)        | 1482 (49.1)| 400 (43.5) | 697 (26.4) | 354 (11.7) | 272 (9.0) | 217 (7.2)  | 35.54 (45.5)  | 81.84 (89.55) | -1.72 (2.32) |
| 2005       | 355 (11.1)     | 95 (10.3)  | 78 (16.7) | 67 (12.4) | 78 (9.7)         | 535 (11.1) | 201 (41.5) | 35 (7.2)  | 198 (14.0) | 109 (3.6) | 149 (4.9)  | 10 (7.2)  | 1061 (35.2) | -2.36 (2.65) |
| 2006       | 242 (8.0)      | 79 (8.6)   | 31 (6.6)  | 22 (4.1)  | 73 (7.2)         | 235 (11.1) | 60 (14.0)  | 36 (3.9)  | 129 (14.0) | 123 (9.6) | 274 (9.1)  | 19 (4.9)  | 355 (11.1) | -2.17 (2.2)  |
| 2007       | 363 (12.0)     | 129 (14.0) | 46 (9.9)  | 74 (13.7) | 272 (9.0)        | 738 (24.5) | 217 (7.2)  | 36 (14.0) | 47 (5.6)  | 109 (3.6) | 274 (9.1)  | 73 (7.2)  | 363 (12.0) | -1.72 (2.32) |
| 2008       | 2012           | 89 (9.7)   | 24 (5.1)  | 31 (5.7)  | 177 (7.5)        | 235 (11.1) | 201 (41.5) | 35 (7.2)  | 129 (14.0) | 123 (9.6) | 274 (9.1)  | 19 (4.9)  | 363 (12.0) | -2.36 (2.65) |
| 2009       | 2013           | 79 (8.6)   | 24 (5.1)  | 31 (5.7)  | 177 (7.5)        | 235 (11.1) | 201 (41.5) | 35 (7.2)  | 129 (14.0) | 123 (9.6) | 274 (9.1)  | 19 (4.9)  | 363 (12.0) | -2.17 (2.2)  |
| 2010       | 2014           | 79 (8.6)   | 24 (5.1)  | 31 (5.7)  | 177 (7.5)        | 235 (11.1) | 201 (41.5) | 35 (7.2)  | 129 (14.0) | 123 (9.6) | 274 (9.1)  | 19 (4.9)  | 363 (12.0) | -2.17 (2.2)  |

Note: SMD were calculated using the `tableone` R package. References on formulas for computing SMD in the multivariate case are provided within the documentation.