Testing different models for *second high* confouding effects (fitting 2000 times)

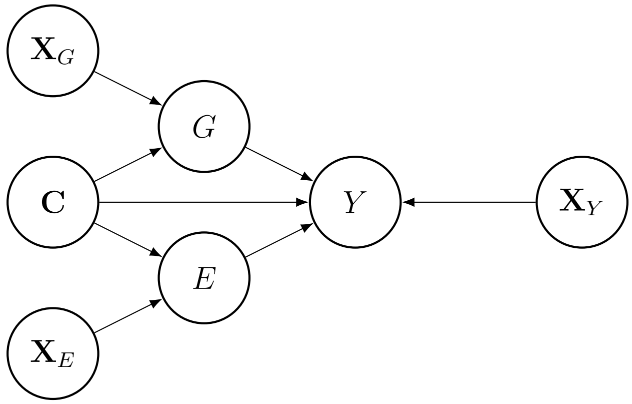


Figure 1: Directed Acyclic Graphs for causal interactions under two treatments G and E.

Let G and E denote the two binary treatments and Y be the binary outcome. Let $\mathbf{X}_G$, $\mathbf{X}_E$ be the set of instrumental variables for G and E, respectively. Let $\mathbf{X}_Y$ be the outcome-only covariates. Let $\mathbf{C}$ be the set of common confounders to both treatments.

Data generation

In this study, we will assume the outcome Y is binary and there are two treatments G and E. Let $\mathbf{X}_G = \{X_1, X_2\}$, $\mathbf{X}_E = \{X_3, X_4\}$, $\mathbf{X}_Y = \{X_5, X_6\}$ and $\mathbf{C} = \{X_7, X_8\}$. Suppose $\mathbf{X} = \{X_1, \ldots, X_8\}$ is generated to follow a multivariate normal distribution such that $\mathbf{X} \sim \mathcal{N}_8(\mathbf{0}, \mathbf{I}_8)$. To generate the treatment assignment probabilities, we firstly generate probabilities for each treatment individually:

$$\pi_{G=1} = P(G = 1 \mid \mathbf{Z}_G) = \frac{\exp\{f(\mathbf{Z}_G)\}}{1 + \exp\{f(\mathbf{Z}_G)\}}$$
$$\pi_{E=1} = P(E_i = 1 \mid \mathbf{Z}_E) = \frac{\exp\{f(\mathbf{Z}_E)\}}{1 + \exp\{f(\mathbf{Z}_E)\}}$$

where $f(\mathbf{Z}_G)$ and $f(\mathbf{Z}_E)$ are the function of $\mathbf{Z}_G$ and $\mathbf{Z}_E$ such that $\mathbf{Z}_G = \{X_1, X_2, X_7, X_8\}$ and $\mathbf{Z}_E = \{X_3, X_4, X_7, X_8\}$. We now suppose

$$f(\mathbf{Z}_G) = 0.8X_1 + 0.5X_2 + 0.3X_7 + 0.5X_8$$
$$f(\mathbf{Z}_E) = -0.5X_3 + 0.2X_4 + 0.3X_7 + 0.5X_8$$

Note that here it is unnecessary to add the intercept as the expectation for each covariate is 0, making the marginal probabilities $\pi_{G=1}$ and $\pi_{E=1}$ are approximately 0.5. With the above setups, the probability of assignment for each combination of treatments can be computed as

$$\theta_{00} = p_{G=0,E=0} = (1 - \pi_{G=1})(1 - \pi_{E=1})$$
$$\theta_{10} = p_{G=1,E=0} = \pi_{G=1}(1 - \pi_{E=1})$$
$$\theta_{01} = p_{G=0,E=1} = (1 - \pi_{G=1})\pi_{E=1}$$
$$\theta_{11} = p_{G=1,E=1} = \pi_{G=1}\pi_{E=1}$$

For each observation, the probability of receiving each pair of treatment follows a multinomial distribution such that $T_i \mid \mathbf{X} \sim \text{Multinomial}(1, \theta_{00}, \theta_{10}, \theta_{01}, \theta_{11})$, where $T_i \in \{(0,0), (1,0), (0,1), (1,1)\}$. In addition, denote $\mathbf{Z}_Y = \{\mathbf{X}_Y, \mathbf{C}\} = \{X_5, X_6, X_7, X_8\}$, the outcome is generated as

$$\begin{aligned}\Pr(Y = 1 \mid \mathbf{Z}_Y, G, E) = \text{logit}^{-1}(\beta_0 + &\\ &0.8X_5 + 0.5X_6 - 0.8X_7 + 0.6X_8 + \\ &0.3G + 0.4E + 0.8GE)\end{aligned}$$

The true value of conditional RERI in terms of odds ratio can be computed as $\exp\{0.3 + 0.4 + 0.8\} - \exp\{0.3\} - \exp\{0.4\} + 1 = 2.64$.

Using control only

$\beta_0 = -6$**: prevalence** ≈ 0.01

Model	β_3	$\widehat{SE}_{\beta_3}$	coverage	Abs. Bias	Bias(%)
Default LOM	3.804(9.460)	4.343	0.963	1.183	45.14
CLOM(X5X6)	3.904(8.135)	3.844	0.965	1.264	47.88
CLOM(X7X8)	3.636(9.080)	3.730	0.942	0.996	37.73
CLOM(X5678)	3.986(9.111)	4.280	0.953	1.346	50.98
MSLOM(X7X8)	3.586(8.817)	3.225	0.932	0.965	36.82
MSLOM(X1278)	3.619(9.370)	3.614	0.928	0.998	38.08
MSLOM(X3478)	3.644(8.753)	3.401	0.934	1.023	39.03
MSLOM(X5678)	3.660(8.905)	3.245	0.934	1.039	39.64
MSLOM(X123478)	3.700(9.619)	3.881	0.930	1.079	41.17
MSLOM(all)	3.639(10.080)	4.041	0.935	1.018	38.84
DR	4.250(11.590)	2.592	0.743	1.610	60.98

Conditional: $RERI_{OR} = 2.64$; Marginal: $RERI_{OR} = 2.621$, $RERI_{RR} = 2.535$

$\beta_0 = -5.5$**: prevalence** ≈ 0.018

Model	β_3	$\widehat{SE}_{\beta_3}$	coverage	Abs. Bias	Bias(%)
Default LOM	2.805(2.505)	1.600	0.947	0.297	11.84
CLOM(X5X6)	3.005(3.098)	1.731	0.950	0.365	13.83
CLOM(X7X8)	2.730(2.504)	1.630	0.918	0.090	3.41
CLOM(X5678)	3.104(4.331)	1.912	0.947	0.464	17.58
MSLOM(X7X8)	2.688(3.608)	1.658	0.923	0.180	7.18
MSLOM(X1278)	2.740(3.545)	1.943	0.925	0.232	9.25
MSLOM(X3478)	2.736(3.723)	1.766	0.920	0.228	9.09
MSLOM(X5678)	2.764(3.507)	1.700	0.931	0.256	10.21
MSLOM(X123478)	2.802(3.319)	2.073	0.930	0.294	11.72
MSLOM(all)	2.894(3.717)	2.127	0.933	0.386	15.39
DR	3.223(7.047)	1.044	0.700	0.583	22.08

Conditional: $RERI_{OR} = 2.64$; Marginal: $RERI_{OR} = 2.508$, $RERI_{RR} = 2.381$

$\beta_0 = -5$**: prevalence** ≈ 0.03

Model	β_3	$\widehat{SE}_{\beta_3}$	coverage	Abs. Bias	Bias(%)
Default LOM	2.500(1.151)	1.066	0.954	0.124	5.22
CLOM(X5X6)	2.723(1.243)	1.152	0.943	0.083	3.14
CLOM(X7X8)	2.489(1.199)	1.117	0.909	0.151	5.72
CLOM(X5678)	2.863(1.395)	1.281	0.950	0.223	8.45
MSLOM(X7X8)	2.394(1.243)	1.141	0.940	0.018	0.76
MSLOM(X1278)	2.455(1.521)	1.350	0.925	0.079	3.32
MSLOM(X3478)	2.422(1.327)	1.207	0.935	0.046	1.94
MSLOM(X5678)	2.495(1.305)	1.185	0.946	0.119	5.01
MSLOM(X123478)	2.499(1.637)	1.434	0.928	0.123	5.18
MSLOM(all)	2.605(1.705)	1.487	0.933	0.229	9.64
DR	2.913(1.594)	0.643	0.659	0.273	10.34

Conditional: $RERI_{OR} = 2.64$; Marginal: $RERI_{OR} = 2.376$, $RERI_{RR} = 2.192$

$\beta_0 = -4.5$**: prevalence** ≈ 0.05

Model	β_3	$\widehat{SE}_{\beta_3}$	coverage	Abs. Bias	Bias(%)
Default LOM	2.267(0.801)	0.780	0.956	0.036	1.61
CLOM(X5X6)	2.531(0.881)	0.861	0.926	0.109	4.13
CLOM(X7X8)	2.319(0.866)	0.838	0.878	0.321	12.16
CLOM(X5678)	2.749(1.024)	0.987	0.949	0.109	4.13
MSLOM(X7X8)	2.194(0.870)	0.840	0.942	0.037	1.66
MSLOM(X1278)	2.217(1.052)	0.990	0.934	0.014	0.63
MSLOM(X3478)	2.208(0.912)	0.889	0.933	0.023	1.03
MSLOM(X5678)	2.319(0.929)	0.883	0.949	0.088	3.94
MSLOM(X123478)	2.234(1.104)	1.047	0.933	0.003	0.13
MSLOM(all)	2.363(1.170)	1.100	0.945	0.132	5.92
DR	2.771(1.097)	0.482	0.648	0.131	4.96

Conditional: $RERI_{OR} = 2.64$; Marginal: $RERI_{OR} = 2.231$, $RERI_{RR} = 1.974$

$\beta_0 = -4$: **prevalence** ≈ 0.08

Model	β_3	$\widehat{SE}_{\beta_3}$	coverage	Abs. Bias	Bias(%)
Default LOM	2.104(0.605)	0.602	0.958	0.020	0.96
CLOM(X5X6)	2.417(0.680)	0.681	0.910	0.223	8.45
CLOM(X7X8)	2.227(0.678)	0.666	0.849	0.413	15.64
CLOM(X5678)	2.726(0.823)	0.810	0.960	0.086	3.26
MSLOM(X7X8)	2.078(0.683)	0.660	0.948	0.006	0.29
MSLOM(X1278)	2.095(0.828)	0.782	0.945	0.011	0.53
MSLOM(X3478)	2.195(4.765)	2.473	0.952	0.111	5.33
MSLOM(X5678)	2.235(0.744)	0.704	0.962	0.151	7.25
MSLOM(X123478)	2.103(0.868)	0.823	0.940	0.019	0.91
MSLOM(all)	2.263(0.940)	0.879	0.956	0.179	8.59
DR	2.746(0.893)	0.392	0.632	0.106	4.01

Conditional: $RERI_{OR} = 2.64$; Marginal: $RERI_{OR} = 2.084$, $RERI_{RR} = 1.737$

$\beta_0 = -3$: **prevalence** ≈ 0.18

Model	β_3	$\widehat{SE}_{\beta_3}$	coverage	Abs. Bias	Bias(%)
Default LOM	1.812(0.401)	0.399	0.954	0.003	0.17
CLOM(X5X6)	2.212(0.474)	0.474	0.796	0.428	16.21
CLOM(X7X8)	2.057(0.472)	0.466	0.707	0.583	22.08
CLOM(X5678)	2.685(0.612)	0.602	0.951	0.045	1.70
MSLOM(X7X8)	1.864(0.476)	0.455	0.949	0.049	2.70
MSLOM(X1278)	1.877(0.577)	0.538	0.948	0.062	3.42
MSLOM(X3478)	1.871(0.507)	0.481	0.953	0.056	3.09
MSLOM(X5678)	2.098(0.544)	0.504	0.938	0.283	15.59
MSLOM(X123478)	1.883(0.611)	0.567	0.948	0.068	3.75
MSLOM(all)	2.121(0.696)	0.632	0.951	0.306	16.86
DR	2.692(0.667)	0.284	0.620	0.052	1.97

Conditional: $RERI_{OR} = 2.64$; Marginal: $RERI_{OR} = 1.815$, $RERI_{RR} = 1.248$

$\beta_0 = -2$**: prevalence** ≈ 0.4

Model	β_3	$\widehat{SE}_{\beta_3}$	coverage	Abs. Bias	Bias(%)
Default LOM	1.635(0.312)	0.308	0.950	0.004	0.25
CLOM(X5X6)	2.091(0.384)	0.381	0.649	0.549	20.80
CLOM(X7X8)	1.955(0.383)	0.375	0.508	0.685	25.95
CLOM(X5678)	2.676(0.515)	0.506	0.947	0.036	1.36
MSLOM(X7X8)	1.752(0.389)	0.368	0.943	0.121	7.42
MSLOM(X1278)	1.759(0.462)	0.438	0.946	0.128	7.85
MSLOM(X3478)	1.754(0.410)	0.390	0.943	0.123	7.54
MSLOM(X5678)	2.082(0.478)	0.426	0.839	0.451	27.65
MSLOM(X123478)	1.761(0.486)	0.462	0.950	0.130	7.97
MSLOM(all)	2.097(0.595)	0.539	0.903	0.466	28.57
DR	2.681(0.556)	0.232	0.604	0.041	1.55

Conditional: $RERI_{OR} = 2.64$; Marginal: $RERI_{OR} = 1.631$, $RERI_{RR} = 0.810$

$\beta_0 = -1$**: prevalence** ≈ 0.6

Model	β_3	$\widehat{SE}_{\beta_3}$	coverage	Abs. Bias	Bias(%)
Default LOM	1.560(0.278)	0.284	0.959	0.000	0.00
CLOM(X5X6)	2.029(0.351)	0.355	0.570	0.611	23.14
CLOM(X7X8)	1.908(0.340)	0.350	0.438	0.732	27.73
CLOM(X5678)	2.668(0.475)	0.480	0.950	0.028	1.06
MSLOM(X7X8)	1.742(0.379)	0.355	0.938	0.182	11.67
MSLOM(X1278)	1.751(0.461)	0.426	0.937	0.191	12.24
MSLOM(X3478)	1.749(0.405)	0.377	0.935	0.189	12.12
MSLOM(X5678)	2.176(0.492)	0.429	0.739	0.616	39.49
MSLOM(X123478)	1.756(0.491)	0.451	0.934	0.196	12.56
MSLOM(all)	2.197(0.634)	0.549	0.832	0.637	40.83
DR	2.682(0.544)	0.211	0.554	0.042	1.59

Conditional: $RERI_{OR} = 2.64$; Marginal: $RERI_{OR} = 1.560$, $RERI_{RR} = 0.475$

Using both control and case

$\beta_0 = -6$**: prevalence** ≈ 0.01

Model	β_3	$\widehat{SE}_{\beta_3}$	coverage	Abs. Bias	Bias(%)
Default LOM	3.804(9.460)	4.343	0.963	1.183	45.14
CLOM(X5X6)	3.904(8.135)	3.844	0.965	1.264	47.88
CLOM(X7X8)	3.636(9.080)	3.730	0.942	0.996	37.73
CLOM(X5678)	3.986(9.111)	4.280	0.953	1.346	50.98
MSLOM(X7X8)	3.534(9.044)	3.160	0.930	0.913	34.83
MSLOM(X1278)	3.556(9.731)	3.605	0.930	0.935	35.67
MSLOM(X3478)	3.583(8.943)	3.398	0.932	0.962	36.70
MSLOM(X5678)	3.462(8.188)	3.144	0.931	0.841	32.09
MSLOM(X123478)	3.558(9.696)	3.877	0.928	0.937	35.75
MSLOM(all)	3.614(9.555)	3.861	0.930	0.993	37.89
DR	4.275(12.462)	2.655	0.745	1.635	61.93

Conditional: $RERI_{OR} = 2.64$; Marginal: $RERI_{OR} = 2.621$, $RERI_{RR} = 2.535$

$\beta_0 = -5.5$**: prevalence** ≈ 0.018

Model	β_3	$\widehat{SE}_{\beta_3}$	coverage	Abs. Bias	Bias(%)
Default LOM	2.805(2.505)	1.600	0.947	0.297	11.84
CLOM(X5X6)	3.005(3.098)	1.731	0.950	0.365	13.83
CLOM(X7X8)	2.730(2.504)	1.630	0.918	0.090	3.41
CLOM(X5678)	3.104(4.331)	1.912	0.947	0.464	17.58
MSLOM(X7X8)	2.613(4.036)	1.616	0.919	0.105	4.19
MSLOM(X1278)	2.636(2.969)	1.880	0.920	0.128	5.10
MSLOM(X3478)	2.649(3.759)	1.715	0.918	0.141	5.62
MSLOM(X5678)	2.585(3.109)	1.605	0.921	0.077	3.07
MSLOM(X123478)	2.705(3.105)	2.008	0.925	0.197	7.85
MSLOM(all)	2.711(3.575)	2.011	0.928	0.203	8.09
DR	3.244(7.633)	1.065	0.701	0.604	22.88

Conditional: $RERI_{OR} = 2.64$; Marginal: $RERI_{OR} = 2.508$, $RERI_{RR} = 2.381$

$\beta_0 = -5$**: prevalence** ≈ 0.03

Model	β_3	$\widehat{SE}_{\beta_3}$	coverage	Abs. Bias	Bias(%)
Default LOM	2.500(1.151)	1.066	0.954	0.124	5.22
CLOM(X5X6)	2.723(1.243)	1.152	0.943	0.083	3.14
CLOM(X7X8)	2.489(1.199)	1.117	0.909	0.151	5.72
CLOM(X5678)	2.863(1.395)	1.281	0.950	0.223	8.45
MSLOM(X7X8)	2.287(1.171)	1.096	0.928	0.089	3.75
MSLOM(X1278)	2.344(1.432)	1.294	0.920	0.032	1.35
MSLOM(X3478)	2.314(1.248)	1.158	0.929	0.062	2.61
MSLOM(X5678)	2.285(1.160)	1.096	0.929	0.091	3.83
MSLOM(X123478)	2.385(1.537)	1.371	0.920	0.009	0.38
MSLOM(all)	2.380(1.518)	1.371	0.923	0.004	0.17
DR	2.918(1.609)	0.648	0.662	0.278	10.53

Conditional: $RERI_{OR} = 2.64$; Marginal: $RERI_{OR} = 2.376$, $RERI_{RR} = 2.192$

$\beta_0 = -4.5$**: prevalence** ≈ 0.05

Model	β_3	$\widehat{SE}_{\beta_3}$	coverage	Abs. Bias	Bias(%)
Default LOM	2.267(0.801)	0.780	0.956	0.036	1.61
CLOM(X5X6)	2.531(0.881)	0.861	0.926	0.109	4.13
CLOM(X7X8)	2.319(0.866)	0.838	0.878	0.321	12.16
CLOM(X5678)	2.749(1.024)	0.987	0.949	0.109	4.13
MSLOM(X7X8)	2.066(0.806)	0.797	0.923	0.165	7.40
MSLOM(X1278)	2.086(0.973)	0.936	0.916	0.145	6.50
MSLOM(X3478)	2.078(0.845)	0.842	0.920	0.153	6.86
MSLOM(X5678)	2.063(0.797)	0.796	0.929	0.168	7.53
MSLOM(X123478)	2.101(1.021)	0.988	0.921	0.130	5.83
MSLOM(all)	2.097(1.009)	0.987	0.922	0.134	6.01
DR	2.774(1.097)	0.488	0.654	0.134	5.08

Conditional: $RERI_{OR} = 2.64$; Marginal: $RERI_{OR} = 2.231$, $RERI_{RR} = 1.974$

$\beta_0 = -4$**: prevalence** ≈ 0.08

Model	β_3	$\widehat{SE}_{\beta_3}$	coverage	Abs. Bias	Bias(%)
Default LOM	2.104(0.605)	0.602	0.958	0.020	0.96
CLOM(X5X6)	2.417(0.680)	0.681	0.910	0.223	8.45
CLOM(X7X8)	2.227(0.678)	0.666	0.849	0.413	15.64
CLOM(X5678)	2.726(0.823)	0.810	0.960	0.086	3.26
MSLOM(X7X8)	1.923(0.617)	0.616	0.933	0.161	7.73
MSLOM(X1278)	1.937(0.747)	0.725	0.927	0.147	7.05
MSLOM(X3478)	1.933(0.646)	0.649	0.938	0.151	7.25
MSLOM(X5678)	1.921(0.609)	0.615	0.933	0.163	7.82
MSLOM(X123478)	1.944(0.782)	0.762	0.924	0.140	6.72
MSLOM(all)	1.941(0.774)	0.762	0.928	0.143	6.86
DR	2.748(0.890)	0.399	0.642	0.108	4.09

Conditional: $RERI_{OR} = 2.64$; Marginal: $RERI_{OR} = 2.084$, $RERI_{RR} = 1.737$

$\beta_0 = -3$**: prevalence** ≈ 0.18

Model	β_3	$\widehat{SE}_{\beta_3}$	coverage	Abs. Bias	Bias(%)
Default LOM	1.812(0.401)	0.399	0.954	0.003	0.17
CLOM(X5X6)	2.212(0.474)	0.474	0.796	0.428	16.21
CLOM(X7X8)	2.057(0.472)	0.466	0.707	0.583	22.08
CLOM(X5678)	2.685(0.612)	0.602	0.951	0.045	1.70
MSLOM(X7X8)	1.658(0.407)	0.410	0.923	0.157	8.65
MSLOM(X1278)	1.667(0.488)	0.480	0.923	0.148	8.15
MSLOM(X3478)	1.664(0.433)	0.432	0.925	0.151	8.32
MSLOM(X5678)	1.656(0.397)	0.410	0.929	0.159	8.76
MSLOM(X123478)	1.671(0.516)	0.504	0.923	0.144	7.93
MSLOM(all)	1.669(0.509)	0.504	0.925	0.146	8.04
DR	2.695(0.663)	0.295	0.641	0.055	2.08

Conditional: $RERI_{OR} = 2.64$; Marginal: $RERI_{OR} = 1.815$, $RERI_{RR} = 1.248$

$\beta_0 = -2$: **prevalence** ≈ 0.4

Model	β_3	$\widehat{SE}_{\beta_3}$	coverage	Abs. Bias	Bias(%)
Default LOM	1.635(0.312)	0.308	0.950	0.004	0.25
CLOM(X5X6)	2.091(0.384)	0.381	0.649	0.549	20.80
CLOM(X7X8)	1.955(0.383)	0.375	0.508	0.685	25.95
CLOM(X5678)	2.676(0.515)	0.506	0.947	0.036	1.36
MSLOM(X7X8)	1.496(0.312)	0.319	0.924	0.135	8.28
MSLOM(X1278)	1.499(0.368)	0.373	0.931	0.132	8.09
MSLOM(X3478)	1.496(0.328)	0.336	0.920	0.135	8.28
MSLOM(X5678)	1.494(0.302)	0.319	0.932	0.137	8.40
MSLOM(X123478)	1.500(0.386)	0.391	0.932	0.131	8.03
MSLOM(all)	1.499(0.379)	0.392	0.932	0.132	8.09
DR	2.682(0.547)	0.248	0.638	0.042	1.59

Conditional: $RERI_{OR} = 2.64$; Marginal: $RERI_{OR} = 1.631$, $RERI_{RR} = 0.810$

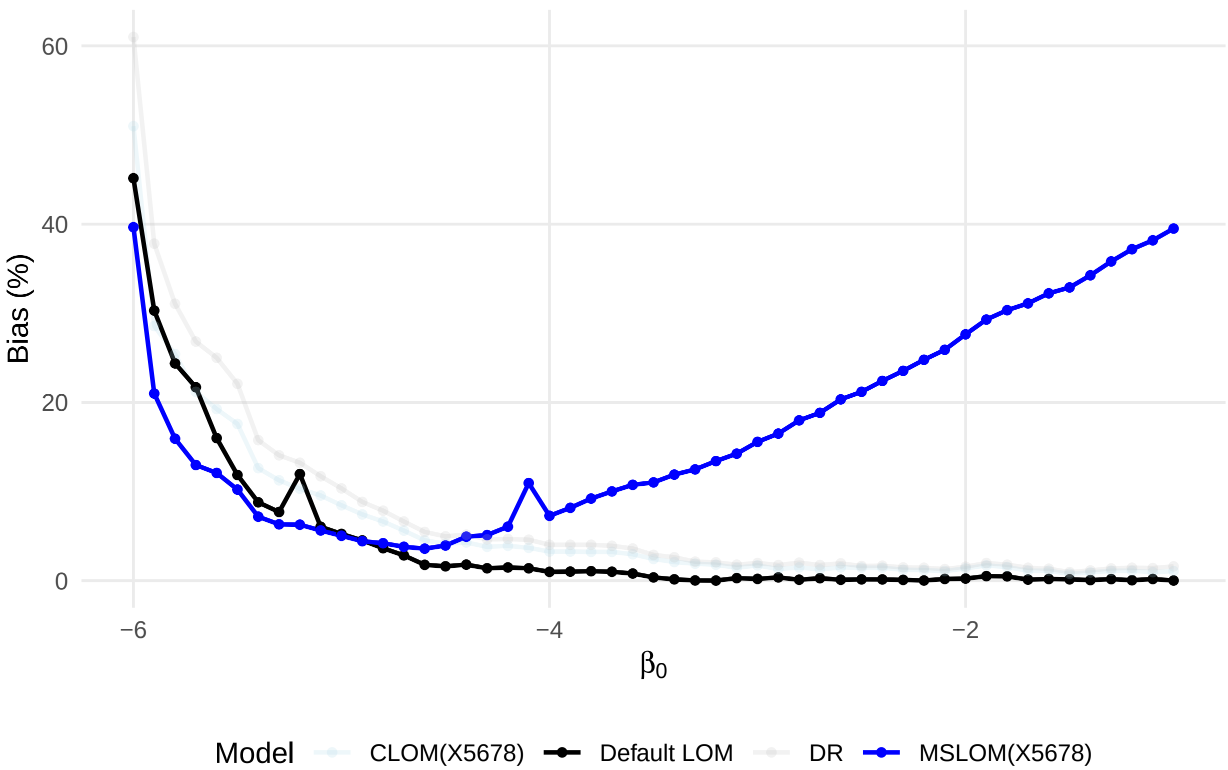
$\beta_0 = -1$: **prevalence** ≈ 0.6

Model	β_3	$\widehat{SE}_{\beta_3}$	coverage	Abs. Bias	Bias(%)
Default LOM	1.560(0.278)	0.284	0.959	0.000	0.00
CLOM(X5X6)	2.029(0.351)	0.355	0.570	0.611	23.14
CLOM(X7X8)	1.908(0.340)	0.350	0.438	0.732	27.73
CLOM(X5678)	2.668(0.475)	0.480	0.950	0.028	1.06
MSLOM(X7X8)	1.435(0.286)	0.297	0.926	0.125	8.01
MSLOM(X1278)	1.437(0.342)	0.347	0.927	0.123	7.88
MSLOM(X3478)	1.437(0.303)	0.312	0.927	0.123	7.88
MSLOM(X5678)	1.434(0.278)	0.297	0.926	0.126	8.08
MSLOM(X123478)	1.437(0.360)	0.363	0.928	0.123	7.88
MSLOM(all)	1.436(0.354)	0.364	0.936	0.124	7.95
DR	2.676(0.526)	0.236	0.625	0.036	1.36

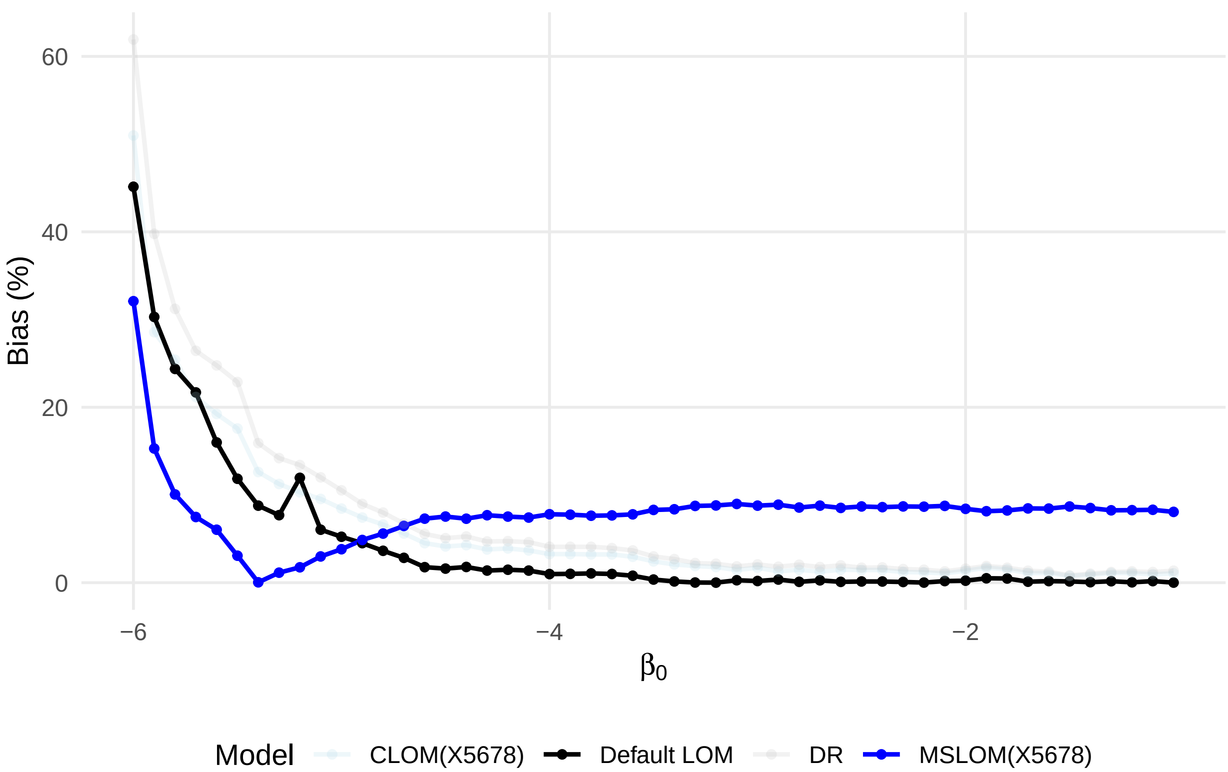
Conditional: $RERI_{OR} = 2.64$; Marginal: $RERI_{OR} = 1.560$, $RERI_{RR} = 0.475$

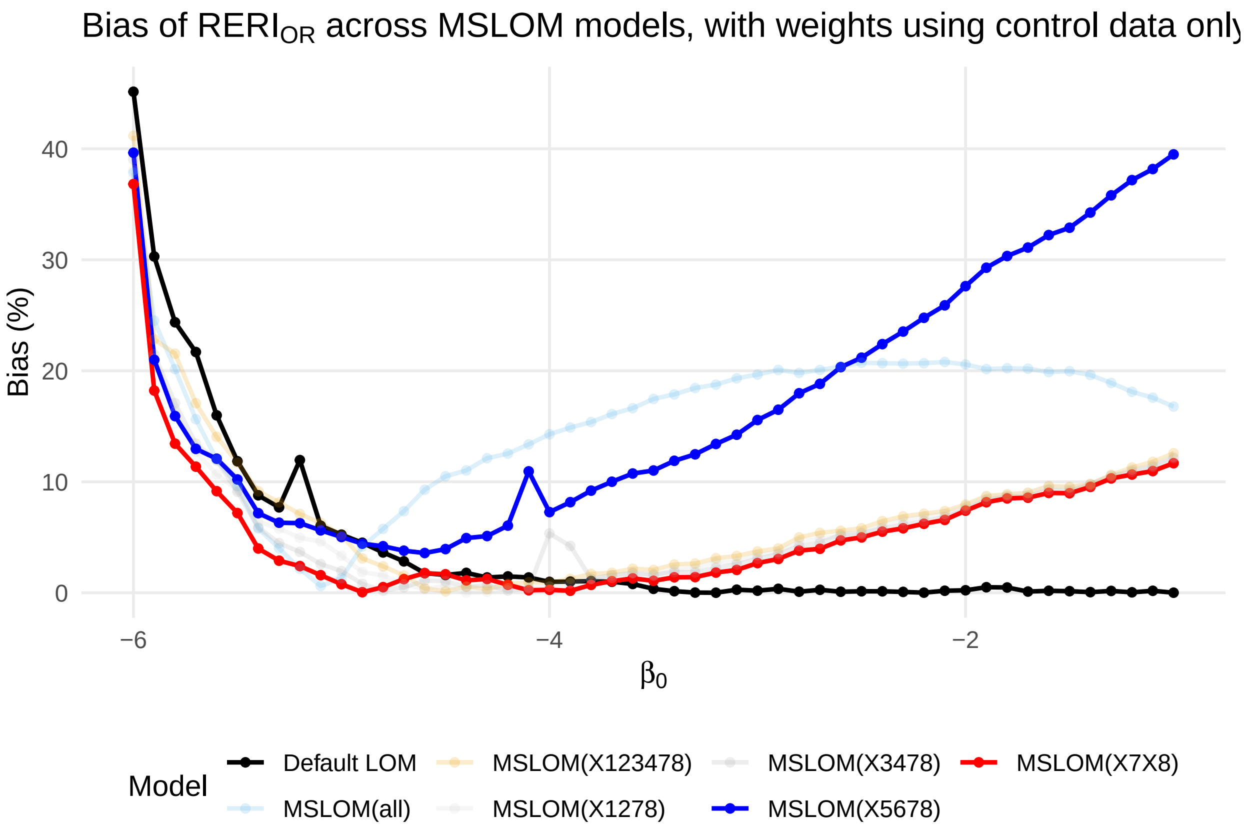
Graphs

Bias of $RERI_{OR}$ across models, with weights using control data only

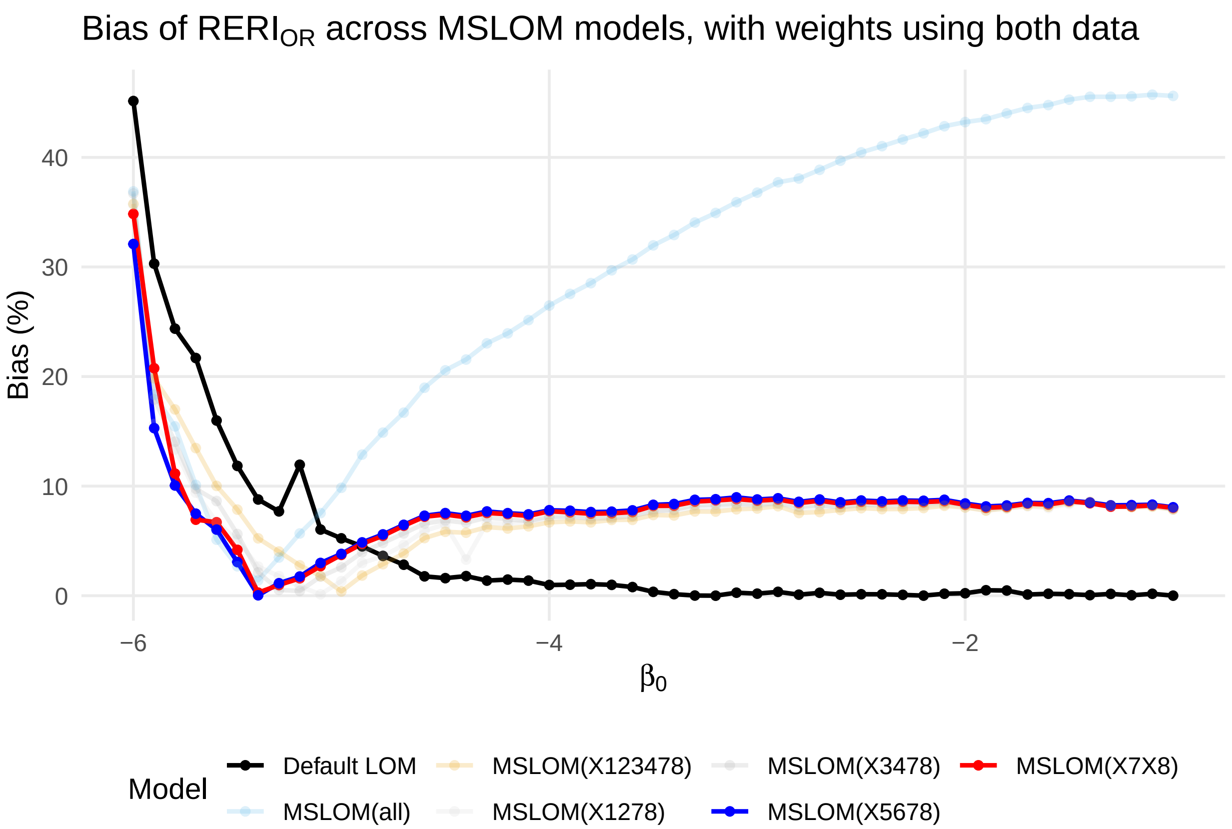


Bias of $RERI_{OR}$ across models, with weights using both data





Bias of $RERI_{OR}$ across MSLOM models, with weights using control data only
Bias (%)
0
10
20
30
40
−6
−4
−2
β_0
Model
Default LOM
MSLOM(all)
MSLOM(X123478)
MSLOM(X1278)
MSLOM(X3478)
MSLOM(X5678)
MSLOM(X7X8)



Bias of $RERI_{OR}$ across MSLOM models, with weights using both data
Bias (%)
0
10
20
30
40
−6
−4
−2
β_0
Model
Default LOM
MSLOM(all)
MSLOM(X123478)
MSLOM(X1278)
MSLOM(X3478)
MSLOM(X5678)
MSLOM(X7X8)