	R function	Covariate distribution	θ distribution
Uncorr.	`uniformW`	Uniform$(a, a+r)$	$a \sim$ U$(-10, 0)$, $r \sim$ U$(0.5, 10)$
	`bernoulliW`	Bernoulli(p)	$p \sim$ U$(0.1, 0.9)$
	`binomialW`	Binomial(n, p)	$n \sim$ U$(2, 10)$, $p \sim$ U$(0.1, 0.9)$
	`normalW`	Normal(μ,σ^2)	$\mu \sim$ U$(-2, 2)$, $\sigma \sim$ U$(0.25, 2)$
	`gammaW`	Gamma(a,b)	$a \sim$ U$(0.5, 2.5), b \sim$ U$(0.5, 2.5)$
Corr.	`normalWCor`	Normal(w, σ^2)	$\sigma \sim$ U$(0.25, 2)$
	`bernoulliWCor`	Bernoulli(expit$(w/$r))	$r \sim$ U$(0.5, 2)$
	`uniformWCor`	$w+$ Uniform$(a, a+r)$	$a \sim$ U$(-2, 0), r \sim$ U$(0.25, 5)$

Table 1: Possible distributions for covariates W

A random data generating distribution is generated as follows. First, choose D_{max}, the maximum dimension of the covariate vector W, and n the number of observations. Next simulate W as follows:

- Randomly sample D, the number of covariates, from a Uniform(1,D_{max}) distribution.
- Randomly sample $\mathcal{D}_1$, the distribution for first covariate, from the uncorrelated covariate distributions listed in the top half of Table 1.
- Randomly select parameters for $\mathcal{D}_1$ from the corresponding distributions listed in the third column of Table 1.
- Generate W_1 by sampling n i.i.d. copies from $\mathcal{D}_1$ with the selected parameter values.
- For $i = 2, \ldots, D$
  - Choose $\mathcal{D}_i$, the distribution for the i-th covariate from the uncorrelated *and* correlated covariate distributions in Table 1.
  - Randomly select parameters for $\mathcal{D}_i$ from the corresponding distributions listed in the third column Table 1.
  - Generate W_i by sampling n copies from $\mathcal{D}_i$ with the selected parameters and possibly inducing correlation with W_{i-1} for distributions in bottom half of Table 1.

Next, simulate A as follows:

- Draw M_1^g, the number of main terms for the propensity score, from Uniform(1,D_{max}).
- For $i = 1, \ldots, M_1^g$:
  - Randomly sample $f_{1,i}$ from the univariate functions in Table 2.
  - Randomly sample $\theta_{1,i}$, the parameters for $f_{1,i}$, from the distributions listed in the third column of Table 2.
  - Let $h_{1,i}(W) = f_i(W_i; \theta_{1,i})$

R function	$f(w;\theta)$	θ distribution
linUni	cw	$c \sim \mathrm{U}(-2,2)$
polyUni	$\sum_{i=1}^{m} c_i w^{b_i}$	$m \sim \mathrm{DU}(1,3), c_i \sim \mathrm{U}(-1,1), b_i \sim \mathrm{DU}(1,3)$
sinUni	$a\ \sin(bw)$	$a \sim \mathrm{U}(-1,1), b \sim \mathrm{U}(-1,1)$
jumpUni	$\sum_{i=1}^{m} a_i I(w \in (b_i, b_{i+1}])$	$m \sim \mathrm{DU}(1,5), a_i \sim U(-2,2)$, $b_i = i/m$-th quantile w
qGammaUni	c qgamma(expit(w); a, b)	$a \sim \mathrm{U}(0.5,5), b \sim \mathrm{U}(0.5,5), c \sim \mathrm{U}(-2,2)$
dNormUni	$\frac{c}{2\pi\sigma}\exp\{-(w-\mu)^2/(2\sigma^2)\}$	$c \sim \mathrm{U}(-4,4), \mu \sim \mathrm{U}(-5,5), \sigma \sim \mathrm{U}(0.25,3)$
pLogisUni	$\frac{c}{1+\exp\{(aw-\mu)/\sigma\}}$	$c \sim \mathrm{U}(-2,2), a \sim \mathrm{U}(-2,2), \mu \sim \mathrm{U}(-2,2)$, $\sigma \sim \mathrm{U}(0.25,2)$
dNormMixUni	$\frac{c_1}{2\pi\sigma}\exp\{-(w-\mu_1)^2/(2\sigma_1^2)\}+$ $\frac{c_2}{2\pi\sigma}\exp\{-(w-\mu_2)^2/(2\sigma_2^2)\}$	$c_1 \sim \mathrm{U}(-2,2), \mu_1 \sim \mathrm{U}(-5,5), \sigma_1 \sim \mathrm{U}(0.5,2)$ $c_2 \sim \mathrm{U}(-2,2), \mu_2 \sim \mathrm{U}(-5,5), \sigma_2 \sim \mathrm{U}(0.5,2)$

Table 2: Possible univariate functions for main terms

- If $D > 1$, sample M_2^g, the number of two-way interactions, from Uniform(1,$M_1^g - 1$).
- For $j = 1, \ldots, M_2^g$:
  - Randomly sample $f_{2,j}$ from the bivariate functions in Table 3.
  - Randomly sample $\theta_{2,j}$, the parameters for $f_{2,i}$, from the distributions listed in the third column of Table 3.
  - Randomly sample two numbers without replacement from $\{1, \ldots, D\}$. Call these numbers a, b.
  - Let $h_{2,j}(W) = f_{2,i}(W_a, W_b; \theta_{2,j})$
- If $D > 2$, sample M_3^g, the number of three-way interactions, from Uniform(1,M_2^g-1).
- For $k = 1, \ldots, M_3^g$:
  - Randomly sample $f_{3,k}$ from the trivariate functions in Table 4.
  - Randomly sample $\theta_{3,k}$, the parameters for $f_{3,k}$, from the distributions listed in the third column of Table 4.
  - Randomly sample three numbers without replacement, say a, b, and c, from $\{1, \ldots, D\}$.
  - Let $h_{3,k}(W) = f_{2,i}(W_a, W_b, W_c; \theta_{3,k})$
- Let $\mathrm{logit}\{g_0(1|W)\} = \sum_{i=1}^{M_1^g} h_i(W) + \sum_{j=1}^{M_2^g} h_j(W) + \sum_{k=1}^{M_3^g} h_k(W)$.
- Draw n independent copies of A from a Bernoulli distribution with conditional probability that $A = 1$ given by $g_0(1|W)$.

Finally, simulate Y as follows:

- Draw M_1^Q, a random number of main terms for outcome regression, from Uniform(2,D_{max})

R function	$f(w_1, w_2; \theta)$	θ distribution
`linBiv`	cw_1w_2	$c \sim \mathrm{U}(-2, 2)$
`polyBiv`	$cw_1^{b_1}w_2^{b_2}$	$c \sim \mathrm{U}(-0.25, 0.25), b_1 \sim \mathrm{DU}(1, 3)$, $b_2 \sim \mathrm{DU}(1, 3)$
`sinBiv`	$a\ \sin(bw_1w_2)$	$a \sim \mathrm{U}(-1, 1), b \sim \mathrm{U}(-1, 1)$
`jumpBiv`	$\sum_{i=1}^{m} \{a_i I(w_1 \in (b_{1,i}, b_{1,i+1}]) \; I(w_2 \in (b_{2,i}, b_{2,i+1}])\}$	$m \sim \mathrm{DU}(1, 5), a_i \sim U(-2, 2)$, $b_{1,i} = i/m$-th quantile w_1, $b_{2,i} = i/m$-th quantile w_2
`dNormAddBiv`	$\frac{c}{2\pi\sigma}\exp\{-(aw_1 + bw_2 - \mu)^2/(2\sigma^2)\}$	$c \sim \mathrm{U}(-5, 5), a \sim \mathrm{U}(-1, 1)$ $b \sim \mathrm{U}(-1, 1), \mu \sim \mathrm{U}(-5, 5)$, $\sigma \sim \mathrm{U}(0.25, 3)$
`dNormMultBiv`	$\frac{c}{2\pi\sigma}\exp\{-(aw_1w_2 - \mu)^2/(2\sigma^2)\}$	$c \sim \mathrm{U}(-5, 5), a \sim \mathrm{U}(-1, 1)$ $\mu \sim \mathrm{U}(-5, 5), \sigma \sim \mathrm{U}(0.25, 3)$
`pLogisUni`	$\frac{c}{1+\exp\{(aw_1+bw_2-\mu)/\sigma\}}$	$c \sim \mathrm{U}(-4, 4), a \sim \mathrm{U}(-1, 1)$, $b \sim \mathrm{U}(-1, 1), \mu \sim \mathrm{U}(-2, 2)$, $\sigma \sim \mathrm{U}(0.25, 2)$

Table 3: Possible bivariate functions for two-way interactions

R function	$f(w_1, w_2, w_3; \theta)$	θ distribution
`linBiv`	$cw_1w_2w_3$	$c \sim \mathrm{U}(-2, 2)$
`polyBiv`	$cw_1^{b_1}w_2^{b_2}w_3^{b_3}$	$c \sim \mathrm{U}(-0.25, 0.25), b_1 \sim \mathrm{DU}(1, 3)$, $b_2 \sim \mathrm{DU}(1, 3), b_3 \sim \mathrm{DU}(1, 3)$
`sinBiv`	$a\ \sin(bw_1w_2w_3)$	$a \sim \mathrm{U}(-1, 1), b \sim \mathrm{U}(-1, 1)$
`jumpBiv`	$\sum_{i=1}^{m} \{a_i I(w_1 \in (b_{1,i}, b_{1,i+1}]) \; I(w_2 \in (b_{2,i}, b_{2,i+1}]) \; I(w_3 \in (b_{3,m-i}, b_{3,m-i+1}])\}$	$m \sim \mathrm{DU}(1, 5), a_i \sim U(-2, 2)$, $b_{1,i} = i/m$-th quantile w_1, $b_{2,i} = i/m$-th quantile w_2 $b_{3,i} = i/m$-th quantile w_3

Table 4: Possible trivariate functions for three-way interactions

- Randomly sample a function f_A from the univariate functions listed in Table 2.
- Randomly sample θ_A, the parameters for f_A, from the distributions listed in the third column of Table 2.
- Let $h_A(A) = f_A(A; \theta_A)$
- For $i = 1, \ldots, M_1^Q - 1$
  - Randomly sample a function $f_{1,i}$ from the univariate functions listed in Table 2.
  - Randomly sample $\theta_{1,i}$, the parameters for $f_{1,i}$ from the distributions listed in the third column of Table 2.
  - Let $h_{1,i}(W) = f_i(W_i; \theta_{1,i})$
- If $D > 1$, sample M_2^Q, the number of two-way interactions from Uniform(1,$M_1^Q - 1$).
- For $j = 1, \ldots, M_2^Q$
  - Randomly sample a function $f_{2,j}$ from the bivariate functions listed in Table 3.
  - Randomly sample $\theta_{2,j}$, the parameters for $f_{2,i}$ from the distributions listed in the third column of Table 3.
  - Randomly sample two numbers without replacement, say a, b, from $\{1, \ldots, D\}$.
  - Let $h_{2,j}(W) = f_{2,i}(W_a, W_b; \theta_{2,j})$
- If $D > 2$, sample M_3^Q, the number of three-way interactions from Uniform(1,M_2^Q-1).
- For $k = 1, \ldots, M_3^Q$
  - Randomly sample a function $f_{3,k}$ from the possible trivariate functions listed in Table 4.
  - Randomly sample $\theta_{3,k}$, the parameters for $f_{3,k}$ from the distributions listed in the third column of Table 4.
  - Randomly sample three numbers without replacement, say a, b, and c, from $\{1, \ldots, D\}$.
  - Let $h_{3,k}(W) = f_{2,i}(W_a, W_b, W_c; \theta_{3,k})$
- Let $\bar{Q}_0(A, W) = h_A(A) + \sum_{i=1}^{M_1^g - 1} h_i(W) + \sum_{j=1}^{M_2^g} h_j(W) + \sum_{k=1}^{M_3^g} h_k(W)$.
- Randomly sample $\mathcal{E}$, a distribution for errors, from the possible distributions listed in Table 5.
- Randomly sample $\theta_{\mathcal{E}}$, the parameters for $\mathcal{E}$ from the distributions listed in Table 5.
- Generate ϵ, by drawing n independent errors from $\mathcal{E}(W; \theta_{\mathcal{E}})$, possibly inducing correlation between ϵ and W (as in the distributions in the bottom of Table 5).
- Let $Y_i = \bar{Q}_0(A_i, W_i) + \epsilon_i$ for $i = 1, \ldots, n$.

	R function	Error distribution	θ distribution
Uncorr.	`normalErr`	$\text{Normal}(0, \sigma^2)$	$\sigma \sim \text{U}(0.5, 7)$
	`uniformErr`	$\text{Uniform}(-r/2, r/2)$	$r \sim \text{U}(1, 25)$
	`gammaErr`	$\text{ShiftedGamma}(a, b)$	$a \sim \text{U}(0.5, 7.5)$
			$b \sim \text{U}(0.5, 7.5)$
Uncorr.	`normalErrW` (binary W_1)	$\text{Normal}(0, \sigma^2\{1 + I(w_1 = 1)\})$	$\sigma \sim \text{U}(0.5, 7)$
	`normalErrW` (cont. W_1)	$\text{Normal}(0, (\text{expit}(w_1)\sigma)^2)$	$\sigma \sim \text{U}(0.5, 7)$
	`uniformErrW` (binary W_1)	$\text{Uniform}(-r - I(w_1 = 1), r + I(w_1 = 1))$	$r \sim \text{U}(1, 25)$
	`uniformErrW` (cont. W_1)	$\text{Uniform}(-r\text{expit}(w_1), r\text{expit}(w_1))$	$\sigma \sim \text{U}(0.5, 7)$

Table 5: Possible error distributions. ShiftedGamma denotes a Gamma distribution shifted to have mean zero.