

Web Material
Propensity Score Matching: Should We Use It in Designing Observational Studies?

Table of Contents	Page No.
Web Appendix 1	2-3
Web Appendix 2	4-6
Web Appendix 3	7
Web Appendix 4	8-9
Web Appendix 5	10-11

Web Appendix 1

The true outcome model is denoted by Equation (1) as:

$$Y_i = \beta_0 + \beta_1 A + \beta_2^T \mathbf{X}_i + \epsilon_i.$$

For simplicity, we assume that the $p \times 1$ vector of confounders $\mathbf{X} = \langle X_1, X_2, \dots, X_p \rangle^T$ are mutually independent, i.e., $X_k \perp\!\!\!\perp X_j$ for all $k \neq j$. Additionally, we assume $\mathbf{E}(X_k) = 0$. When we replace $\mathbf{X}$ with the logit PS, $Z = \text{logit}(e(\mathbf{X}))$, in equation (1), we have (Wan(2016,2019)):

$$Y_i = \beta_0 + \beta_1 A + \gamma Z + \nu_i + \epsilon_i,$$

where $\gamma = \frac{\|\beta_2\|}{\|\alpha_1\|} \cos(\theta)$ and $\nu_i = \beta_2^T \sin(\theta) \mathbf{X}_i$, θ is the angle between two coefficient vectors β_2 and α_1 .

(1) Covariate matching design

If we perform a 1:1 exact covariate matching design on confounders $\mathbf{X}$ and formed M matched pairs on $\mathbf{X}_m$, $m = 1, 2, \dots, M$. This pair-matched design mimics a randomized block design. Within the m th pair, we have the outcomes of treated and untreated subjects denoted as follows:

$$\begin{aligned} Y_{1m} &= \beta_0 + \beta_1 + \beta_2^T \mathbf{X}_m + \epsilon_{1m}, \\ Y_{0m} &= \beta_0 + \beta_2^T \mathbf{X}_m + \epsilon_{0m} \end{aligned}$$

We let $\Sigma_{\mathbf{X}}$ represent the $p \times p$ variance-covariance matrix of $\mathbf{X}$ in the matched population under the exact covariate matching design. The variances of the sample means $\bar{Y}_{i.} = \frac{\sum_{m=1}^M Y_{im}}{M}$, $i = 0, 1$ and mean difference $\bar{Y}_{1.} - \bar{Y}_{0.}$ in an exactly matched covariate matching design are

$$\text{Var}(\bar{Y}_{1.}) = \text{Var}(\bar{Y}_{0.}) = \frac{\beta_2^T \Sigma_{\mathbf{X}} \beta_2 + \sigma_{\epsilon}^2}{M},$$

and

$$\text{Var}(\bar{Y}_{1.} - \bar{Y}_{0.}) = \frac{2\sigma_{\epsilon}^2}{M}$$

The intraclass correlation within pair is

$$\begin{aligned} \rho_1 &= \frac{\beta_2^T \Sigma_{\mathbf{X}} \beta_2}{\beta_2^T \Sigma_{\mathbf{X}} \beta_2 + \sigma_{\epsilon}^2} \\ &= \frac{1}{1 + \sigma_{\epsilon}^2 / \beta_2^T \Sigma_{\mathbf{X}} \beta_2} \end{aligned}$$

(2) PSM design

Similarly, if we perform a 1:1 matching design on exact logit PS. In each matched pair formed on the logit PS Z_n , $n = 1, 2, \dots, N$, the covariate values of the treated and untreated subjects are not necessary the same (i.e., $\mathbf{X}_{1n} \neq \mathbf{X}_{0n}$) but their propensity scores are same (i.e., $e(\mathbf{X}_{1n}) = e(\mathbf{X}_{0n})$). This pair-matched design also mimics a randomized block design. Within the n th pair, we have the outcomes of treated and untreated subjects denoted as follows:

$$\begin{aligned} Y_{1n} &= \beta_0 + \beta_1 + \gamma Z_n + \nu_{1n} + \epsilon_{1n}, \\ Y_{0n} &= \beta_0 + \gamma Z_n + \nu_{0n} + \epsilon_{0n}, \end{aligned}$$

The variance of the sample mean difference $\bar{Y}_{1.} - \bar{Y}_{0.}$ in an exactly matched PSM design is

$$\text{Var}(\bar{Y}_{1.} - \bar{Y}_{0.}) = \frac{2(\sigma_{\nu}^2 + \sigma_{\epsilon}^2)}{N}$$

where $\sigma_\nu^2 = \sin(\theta)^2 \beta_2^T \tilde{\Sigma}_X \beta_2$ and $\tilde{\Sigma}_X$ represent the $p \times p$ variance-covariance matrix of $\mathbf{X}$ in the matched population from PSM. Since $Z = \alpha_0 + \alpha_1^T \mathbf{X}$, the intraclass correlation within pair is

$$\begin{aligned} \rho_2 &= \frac{\gamma^2 \alpha_1^T \tilde{\Sigma}_X \alpha_1}{\gamma^2 \alpha_1^T \tilde{\Sigma}_X \alpha_1 + \sigma_\nu^2 + \sigma_\epsilon^2} \\ &= \frac{1}{1 + (\sin(\theta)^2 \beta_2^T \tilde{\Sigma}_X \beta_2 + \sigma_\epsilon^2) / ((\frac{\|\beta_2\|}{\|\alpha_1\|} \cos(\theta))^2 \alpha_1^T \tilde{\Sigma}_X \alpha_1)} \end{aligned}$$

When $\theta = 0$ or $\theta = \pi$, α_1 and β_2 are parallel and we have the maximum ρ_2

$$\rho_{2,max} = \frac{1}{1 + \sigma_\epsilon^2 / ((\frac{\|\beta_2\|}{\|\alpha_1\|})^2 \alpha_1^T \tilde{\Sigma}_X \alpha_1)}$$

When $\theta = \pi/2$ and α_1 and β_2 are perpendicular, we have the minimum ρ_2

$$\rho_{2,min} = \frac{1}{1 + (\beta_2^T \tilde{\Sigma}_X \beta_2 + \sigma_\epsilon^2)}$$

When $M \cong N$, the sample mean difference $\bar{Y}_1 - \bar{Y}_0$ generally has larger variance in a exact PSM design than in a exact covariate matching design. Thus, covariate matching design is more efficient given the same number of matched pairs. When the two coefficient sets β_2 and α_1 are parallel, $\theta = 0$ and $\sin(\theta) = 0$. Then the variances of $\bar{Y}_1 - \bar{Y}_0$ are same in both designs.

Web Appendix 2

```

library(MatchIt)
library("tableone")
library(smd)

library("lmtest")
library("sandwich")
library("gee")

nsim<-5000

calsize.vec<-c(20,1,0.2,0.02,0.002,0.0002)

est.ps.nox<-corr<-n.sim<-smd.u2<-man.dist<-numeric(nsim)
robuststd.ps.nox<-naivestd.ps.nox<-numeric(nsim)

est.ps.nox.sum<-std.ps.nox.sum<-n.sim.sum<-smd.u2.sum<-largeP.u2.sum<- man.dist.sum<-numeric(length(calsize.vec))
std.ps.nox.mod.sum<-std.hc.ps.nox.mod.sum<-corr.sum<-numeric(length(calsize.vec))

n = 1000000

trt<- sample(rep(0:1, each = n/2), n, FALSE)
X1 <- ifelse(trt==1, runif(n, 1,6), runif(n, 0,5))
X2 <- ifelse(trt==1, runif(n, 1,6), runif(n, 0,5))
Y<-2*trt+X1+X2+rnorm(n)

glm(trt~X1+X2,family=binomial(link='logit'))

b<-c(1,1)
a<-c(0.47,0.47)

norm.a<-norm(a,type="2")
norm.b<-norm(b,type="2")

p1<-(norm.b/norm.a)^2*(0.47^2*1.8*2)

1/(1+1/(p1))

var(matchData$X1)

for (i in 1: length(calsize.vec)){

calsize<-calsize.vec[i]

for (j in 1:nsim){

n = 10000

trt<- sample(rep(0:1, each = n/2), n, FALSE)
X1 <- ifelse(trt==1, runif(n, 1,6), runif(n, 0,5))
X2 <- ifelse(trt==1, runif(n, 1,6), runif(n, 0,5))
Y<-2*trt+X1+X2+rnorm(n)

```

```

dkn <- data.frame(trt, X1, X2, Y, data_method = "kn")

m.out2<- matchit(trt ~ X1 + X2,
distance = "logit", method = "nearest",
ratio = 1, discard="both", replace = FALSE, caliper =calsize,
std.caliper=T, m.order = "random", data=dkn)

matchData<-match.data(m.out2)

matchData.order<-matchData[order(matchData$subclass),]

### GEE analysis ###

gee.ps.nox <- gee(Y~ trt, data=matchData.order,
id = subclass,
family =gaussian,
corstr ="exchangeable")

corr[j]<- gee.ps.nox$working.correlation[1,2]
est.ps.nox[j]<-coef(gee.ps.nox)[2]
robuststd.ps.nox[j]<-coef(summary(gee.ps.nox))[2,4]
naivestd.ps.nox[j]<-summary(gee.ps.nox)$coefficients[2,2]
n.sim[j]<-dim(matchData)[1]

###balance diagnosis ###;

smd.u2[j]<-as.numeric(smd(x= matchData$X1,g= matchData$trt)[1,2])

mean.grp.1<-apply(matchData[matchData$trt==1,c("X1","X2")],2, mean)
mean.grp.0<-apply(matchData[matchData$trt==0,c("X1","X2")],2, mean)

S<-var(matchData[,c("X1","X2")])

###pairwise-mahalanobis-distance-in-r##
man.dist[j]<-t(mean.grp.1-mean.grp.0)%*%solve(S)%*%(mean.grp.1-mean.grp.0)

### Compute the imbalance ###

} ###end of j

est.ps.nox.sum[i]<-mean(est.ps.nox)
std.ps.nox.sum[i]<-sqrt(var(est.ps.nox))
std.ps.nox.mod.sum[i]<-mean(naivestd.ps.nox,na.rm=T)
std.hc.ps.nox.mod.sum[i]<-mean(robuststd.ps.nox)
corr.sum[i]<-mean(corr)

n.sim.sum[i]<-mean(n.sim/2)
smd.u2.sum[i]<-mean(smd.u2)

largeP.u2.sum[i]<-mean(abs(smd.u2)>0.1)
man.dist.sum[i]<-mean(man.dist)
print(i)
}

est.ps.nox.sum
std.ps.nox.sum
std.ps.nox.mod.sum

```

std.hc.ps.nox.mod.sum
corr.sum
n.sim.sum
smd.u2.sum
largeP.u2.sum
man.dist.sum

Web Appendix 3

Coefficients used in the simulation section “Imbalance and Bias”

TABLE I: Coefficients of Confounders

$sine(\theta)$	The Outcome Model					The Treatment Model				
	β_{21}	β_{22}	β_{23}	β_{24}	β_{25}	α_{11}	α_{12}	α_{13}	α_{14}	α_{15}
[0,0.2]	0.075	0.523	-0.299	-0.598	-0.523	-0.070	-0.491	0.421	0.632	0.421
(0.8,1]	-0.762	-0.305	0.152	0.457	0.305	0.586	-0.260	-0.456	0.586	-0.195

Coefficients used in the simulation section “Model Dependence”

$$\text{logit}(P(A = 1|\mathbf{X})) = \alpha_0 + \alpha_{11}X_1 + \alpha_{12}X_2 + \alpha_{13}X_3 + \alpha_{14}X_4 + \alpha_{15}X_5$$

where $\alpha_0 = -0.9$, $\alpha_{11} = -0.070$, $\alpha_{12} = -0.491$, $\alpha_{13} = 0.421$, $\alpha_{14} = 0.632$, $\alpha_{15} = 0.421$.

$$Y = \beta_0 + \beta_1 A + k * (\beta_{21}X_1 + \beta_{22}X_2 + \beta_{23}X_3 + \beta_{24}X_4 + \beta_{25}X_5) \\ + l * (\beta_{31}X_1^2 + \beta_{32}X_2^2 + \beta_{33}X_3^2 + \beta_{34}X_3X_4 + \beta_{35}X_2X_5)$$

where $\beta_0 = -0.8$, $\beta_1 = 1$, $\beta_{21} = 0.075$, $\beta_{22} = 0.523$, $\beta_{23} = -0.299$, $\beta_{24} = -0.598$, $\beta_{25} = -0.523$, $\beta_{31} = 0.299$, $\beta_{32} = 0.598$, $\beta_{33} = -0.224$, $\beta_{34} = 0.673$, $\beta_{35} = -0.224$, $k = 1.2$, $l = 0.5$

Web Appendix 4

TABLE II: Estimates for Figure 1

Outcome	Caliper Size					
	20	2	0.2	0.02	0.002	0.0002
Estimate from $\mathcal{M}(A)$	0.152	0.416	0.494	0.497	0.502	0.499
Estimate from $\mathcal{M}(A, X_4, X_5)$	-0.400	0.280	0.488	0.500	0.505	0.500
Estimate from $\mathcal{M}(A, X_1, X_2, X_3, X_4, X_5)$	0.500	0.497	0.500	0.500	0.501	0.500
Empirical <i>s.e.</i> , $\mathcal{M}(A)$	0.124	0.121	0.137	0.146	0.241	0.657
Model-based <i>s.e.</i> , $\mathcal{M}(A)$	0.159	0.198	0.217	0.224	0.292	0.681
Robust <i>s.e.</i> , $\mathcal{M}(A)$	0.158	0.197	0.216	0.222	0.289	0.659
Empirical <i>s.e.</i> , $\mathcal{M}(A, X_4, X_5)$	0.157	0.102	0.111	0.117	0.179	0.492
Model-based <i>s.e.</i> , $\mathcal{M}(A, X_4, X_5)$	0.127	0.153	0.168	0.173	0.231	0.555
Robust <i>s.e.</i> , $\mathcal{M}(A, X_4, X_5)$	0.123	0.142	0.154	0.159	0.207	0.475
Empirical <i>s.e.</i> , $\mathcal{M}(A, X_1, X_2, X_3, X_4, X_5)$	0.071	0.077	0.085	0.089	0.114	0.278
Model-based <i>s.e.</i> , $\mathcal{M}(A, X_1, X_2, X_3, X_4, X_5)$	0.072	0.078	0.085	0.088	0.114	0.276
Proportion of SMD of X_3 larger than 0.1	0.927	0.068	0.026	0.041	0.253	0.695
Mahalanobis distance	0.988	0.089	0.009	0.010	0.032	0.275
SMD	0.169	0.040	0.002	-0.001	0.001	0.000

TABLE III: Estimates for Figure 2

Outcome	Caliper Size					
	20	2	0.2	0.02	0.002	0.0002
Estimate from $\mathcal{M}(A)$	-0.989	0.143	0.479	0.500	0.501	0.498
Estimate from $\mathcal{M}(A, X_4, X_5)$	-0.474	0.259	0.487	0.501	0.503	0.503
Estimate from $\mathcal{M}(A, X_1, X_2, X_3, X_4, X_5)$	0.499	0.499	0.500	0.499	0.501	0.498
Empirical <i>s.e.</i> , $\mathcal{M}(A)$	0.238	0.095	0.088	0.089	0.120	0.295
Model-based <i>s.e.</i> , $\mathcal{M}(A)$	0.105	0.127	0.138	0.144	0.202	0.508
Robust <i>s.e.</i> , $\mathcal{M}(A)$	0.092	0.088	0.092	0.095	0.123	0.281
Empirical <i>s.e.</i> , $\mathcal{M}(A, X_4, X_5)$	0.147	0.088	0.089	0.091	0.128	0.343
Model-based <i>s.e.</i> , $\mathcal{M}(A, X_4, X_5)$	0.095	0.111	0.121	0.125	0.171	0.418
Robust <i>s.e.</i> , $\mathcal{M}(A, X_4, X_5)$	0.089	0.088	0.092	0.095	0.126	0.288
Empirical <i>s.e.</i> , $\mathcal{M}(A, X_1, X_2, X_3, X_4, X_5)$	0.071	0.078	0.085	0.087	0.115	0.281
Model-based <i>s.e.</i> , $\mathcal{M}(A, X_1, X_2, X_3, X_4, X_5)$	0.072	0.078	0.085	0.087	0.115	0.276
Proportion of SMD of X_3 larger than 0.1	1.000	0.313	0.019	0.033	0.237	0.685
Mahalanobis distance	0.983	0.089	0.008	0.010	0.032	0.276
SMD	0.339	0.081	0.004	-0.001	-0.001	-0.005

TABLE IV: Estimates for Figure 3

Outcome	Caliper Size					
	20	2	0.2	0.02	0.002	0.0002
Estimate from $\mathcal{M}(A)$	-0.074	0.631	0.983	0.999	1.001	0.998
Estimate from $\mathcal{M}(A, X_1, X_2, X_3, X_4, X_5)$	1.045	1.053	1.008	1.000	1.001	1.002
Empirical <i>s.e.</i> , $\mathcal{M}(A)$	0.218	0.186	0.201	0.208	0.266	0.649
Model-based <i>s.e.</i> , $\mathcal{M}(A)$	0.164	0.193	0.223	0.229	0.295	0.702
Robust <i>s.e.</i> , $\mathcal{M}(A)$	0.157	0.179	0.203	0.208	0.267	0.625
Empirical <i>s.e.</i> , $\mathcal{M}(A, X_1, X_2, X_3, X_4, X_5)$	0.173	0.181	0.200	0.206	0.263	0.644
Model-based <i>s.e.</i> , $\mathcal{M}(A, X_1, X_2, X_3, X_4, X_5)$	0.172	0.179	0.200	0.206	0.264	0.639
Proportion of SMD of X_3 larger than 0.1	1.000	0.672	0.018	0.029	0.222	0.696
Mahalanobis distance	0.983	0.183	0.008	0.009	0.029	0.283
SMD	0.339	0.117	0.006	-0.000	-0.001	-0.004

Web Appendix 5

We extended the same simulation settings using the coefficients listed in the first row in Table (I) and under “Model Dependence”, but for correlated variables. We generated 3 correlated continuous variables $\sim N(0, 1)$ with pairwise correlation coefficient 0.5, 2 correlated binary variables $\sim \text{Binom}(0.5)$ with correlation coefficient 0.3.

The results in the supplemental Figures 1 and 2 and are consistent with the corresponding scenarios with independent $\mathbf{X}$.

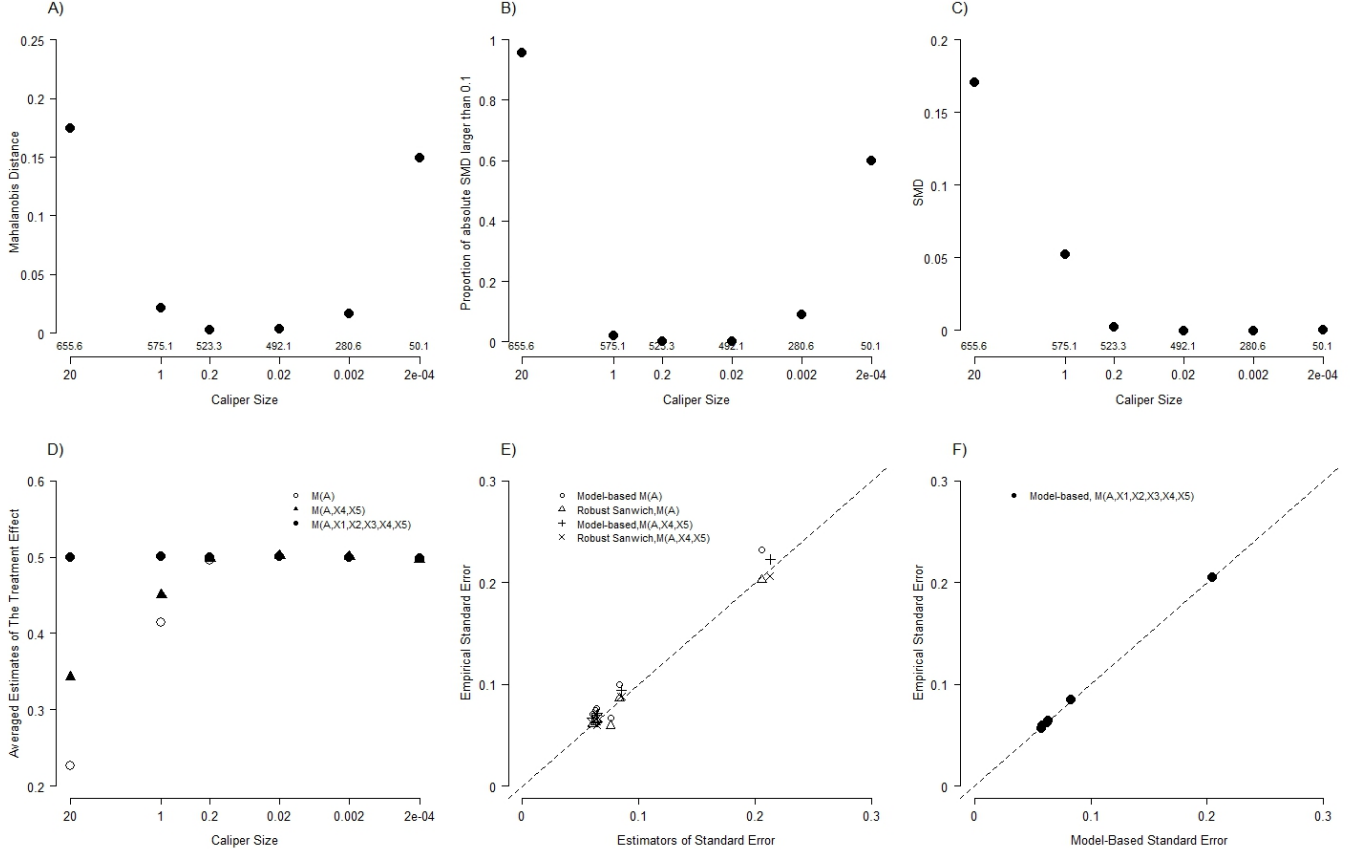


FIG. 1: Imbalance and Bias for the *sine* distance > 0.8 . A) The trend of the mahalanobis distance with shrinking caliper size; B) The trend of the proportion of absolute SMD of X_3 larger than 0.1; C) The trend of SMD of X_3 ; D) The trend of estimators of PATE; E) The concordance between empirical standard error and model-based/Robust Sandwich estimators for $\mathcal{M}(A)$ and $\mathcal{M}(A, X_4, X_5)$; F) The concordance between empirical standard error and model-based estimators for $\mathcal{M}(A, X_1, X_2, X_3, X_4, X_5)$

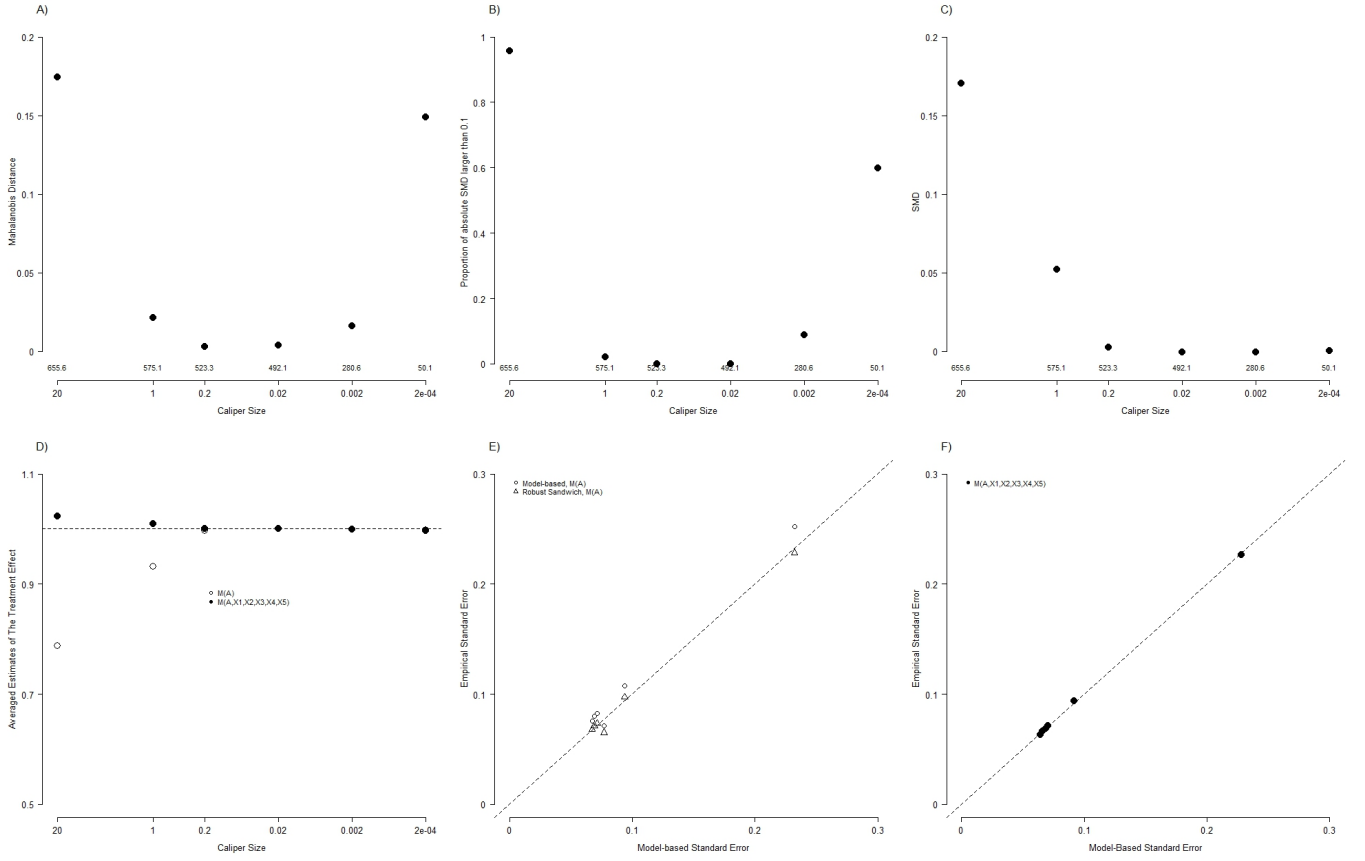


FIG. 2: Sensitivity to Model Dependence. A) The trend of the mahalanobis distance with shrinking caliper size; B) The trend of the proportion of absolute SMD of X_3 larger than 0.1; C) The trend of SMD of X_3 ; D) The trend of estimators of PATE; E) The concordance between empirical standard error and model-based estimators for $\mathcal{M}(A)$; F) The concordance between empirical standard error and model-based estimators for $\mathcal{M}(A, X_1, X_2, X_3, X_4, X_5)$