PROPENSITY SCORE REGRESSION WITH CONTINUOUS EXPOSURE

EXAMPLE: In this example, we have a data generating process with $k = 10$ predictors, joint Normally distributed $X \sim Normal(\boldsymbol{\mu}, \Sigma)$, with treatment model given by

$$Z|X = x \sim Normal(\mathbf{x}_\alpha \alpha, \sigma_Z^2)$$

and outcome model

$$Y|X = x, Z = z \sim Normal(\mu(x, z), \sigma_Y^2)$$

with

$$\mu(x, z) = \mathbf{x}_0\beta + \psi_0 z + \psi_1 z^2 = \beta_0 + \sum_{l=1}^{k} \beta_l x_l + \psi_0 z + \psi_1 z^2.$$

In the model, all predictors are predictors of Z, but only first three components are confounders. In the analysis

$$\alpha = (4, 1, 1, 1, 1, 1, 1, -2, -2, 2)^\top \qquad \beta = (2, -2, 2.2, 3.6, 0, 0, 0, 0, 0, 0)^\top$$

with treatment effect parameters $\psi_0 = 1$, $\psi_1 = 0.2$.

```
#Set-up
set.seed(865177)
library(mvnfast)
n<-10000
k<-10
Mu<-sample(-5:5,size=k,rep=T)/2
Sigma<-0.1*0.8^abs(outer(1:k,1:k,'-'))
al<-c(4,rep(1,6),rep(-2,4))
be<-rep(0,k+1)
be[1:5]<-c(2,-2.0,2.2,3.6)
sigZ<-5;sigY<-1

#Exploratory sample
X<-rmvn(n,mu=Mu,Sigma)
Xal<-cbind(1,X)
ps.true<-Xal %*% al
Z<-rnorm(n,ps.true,sigZ)

#Find a plotting grid
Z0<-round(mean(Z),0)
zgrid<-seq(Z0-5,Z0+5,by=0.1)
nZ<-length(zgrid)

#Intercept
intercept.val<-be[1]+sum(Mu*be[-c(1:2)])
psi<-c(1,0.2)
Xm<-cbind(1,X)
muval<-as.vector(Xm %*% be) + Z*psi[1] + Z^2*psi[2]
Y<-rnorm(n,muval,sigY)
summary(Z)

+    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
+ -10.934   4.886   8.509   8.540  12.132  28.798
```

The quantities ψ_0 and ψ_1 measure the unconfounded effect of treatment, as, marginally,

$$\mathbb{E}[Y(z)] = \psi_0 z + \psi_1 z^2 + \mu_0\beta.$$

where $\mu_0 = (1, \mu^\top)$ so that the intercept is $\mu_0\beta = 0.8$. Here, $\sigma_Z = 5$ and $\sigma_Y = 1$. We have that the ATE curve is

$$\mathbb{E}[Y(z) - Y(0)] = \psi_0 z + \psi_1 z^2.$$

Propensity score regression can be carried out using balancing scores that block the backdoor paths. For up to quadratic terms, we may need to model

$$b_1(x) = \mathbb{E}_{Z|X}[Z|X = x]$$

and

$$b_2(x) = \mathbb{E}_{Z|X}[Z^2|X = x]$$

and then fit a model such as

$$m(x, z) = \beta_0 + \psi_0 z + \psi_1 z^2 + \phi_0 b_1(x) + \phi_1 b_2(x)$$

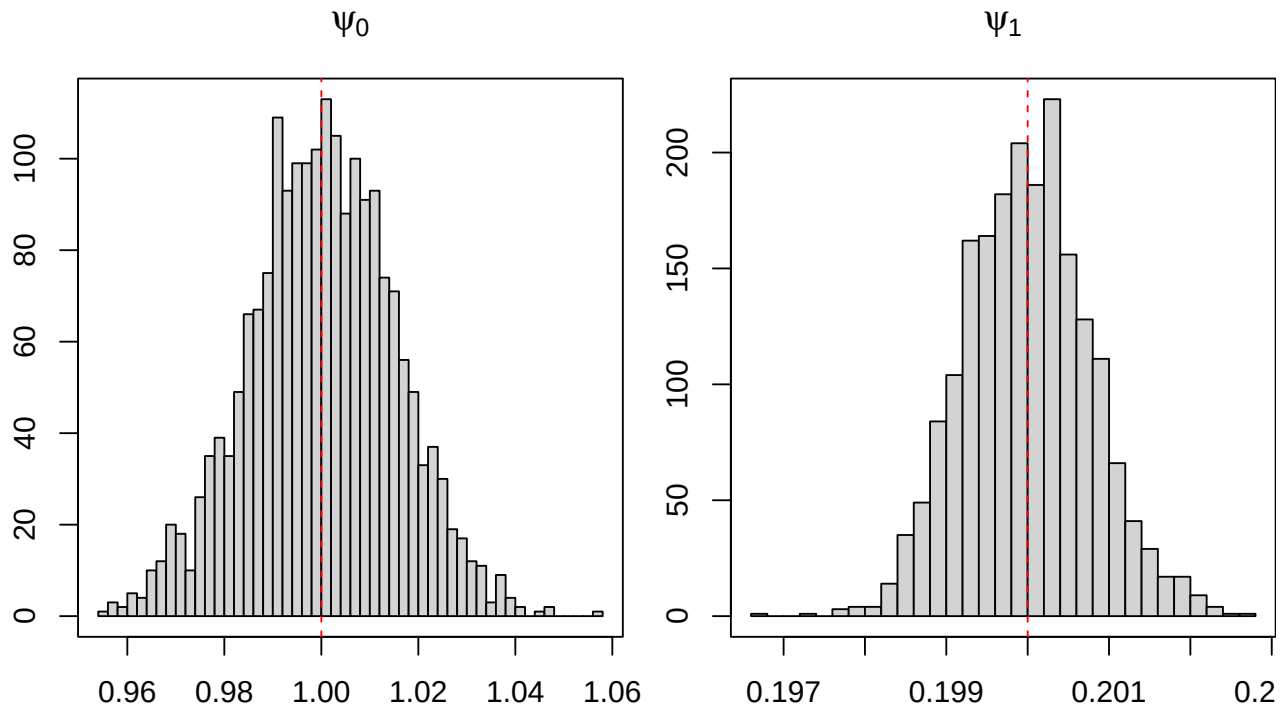
and so on.

We carry out a simulation study of 2000 replicates with $n = 1000$.

```
#Monte Carlo study
n<-1000
nreps<-2000
gest.samp<-matrix(0,ncol=2,nrow=nreps)
for(irep in 1:nreps){
    X<-rmvn(n,mu=Mu,Sigma)
    Xal<-cbind(1,X)
    ps.true<-Xal %*% al
    Z<-rnorm(n,ps.true,sigZ)
    Xm<-cbind(1,X)
    muval<-as.vector(Xm %*% be)+Z*psi[1] + Z^2*psi[2]
    Y<-rnorm(n,muval,sigY)

    ps1<-fitted(lm(Z~X))
    ps2<-fitted(lm(Z^2~X))

    fit1<-lm(Y~Z+I(Z^2)+ps1+ps2)
    gest.samp[irep,]<-coef(fit1)[2:3]
}
```



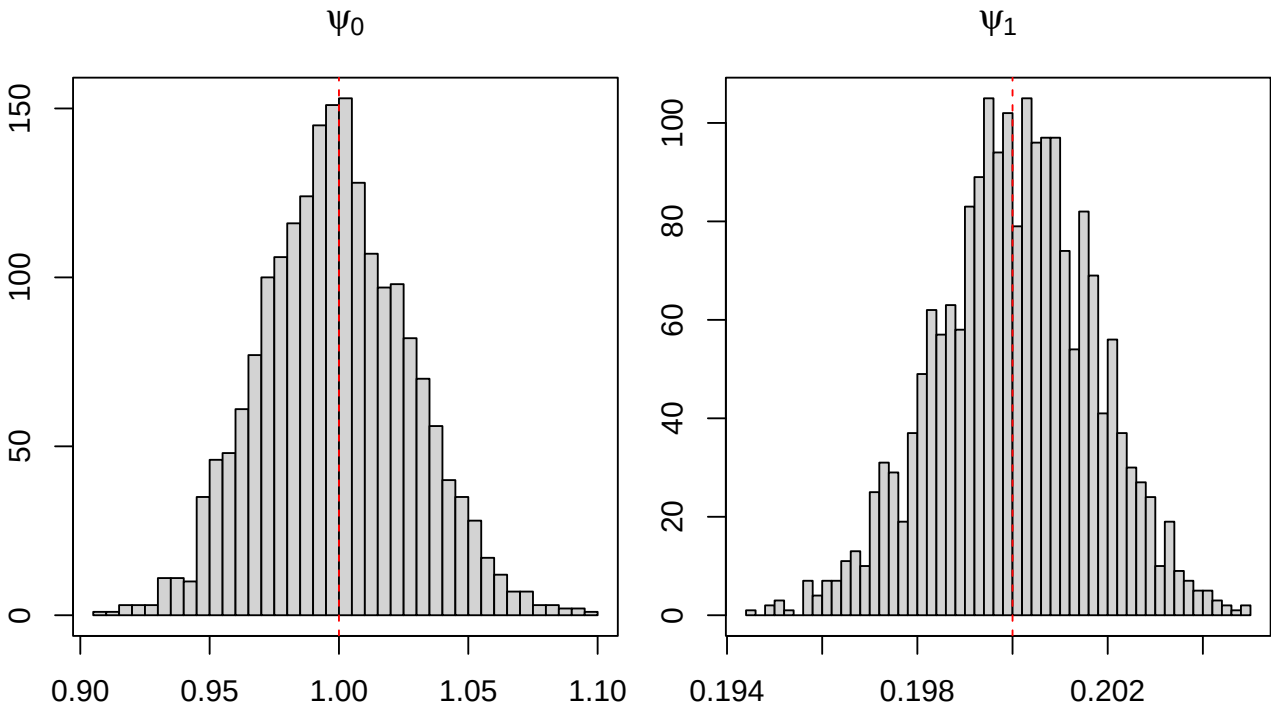
Now with only one term in the propensity score:

$$m(x,z) = \beta_0 + \psi_0 z + \psi_1 z^2 + \phi_0 b_1(x)$$

```
#Monte Carlo study
n<-1000
nreps<-2000
gest2.samp<-matrix(0,ncol=2,nrow=nreps)
for(irep in 1:nreps){
    X<-rmvn(n,mu=Mu,Sigma)
    Xal<-cbind(1,X)
    ps.true<-Xal %*% al
    Z<-rnorm(n,ps.true,sigZ)
    Xm<-cbind(1,X)
    muval<-as.vector(Xm %*% be)+Z*psi[1] + Z^2*psi[2]
    Y<-rnorm(n,muval,sigY)

    ps1<-fitted(lm(Z~X))

    fit1<-lm(Y~Z+I(Z^2)+ps1)
    gest2.samp[irep,]<-coef(fit1)[2:3]
}
```



That is, we still get unbiased estimation of ψ_0 and ψ_1 even if we only use the balancing score $b_1(x)$.

EXAMPLE: We now consider the model with interactions with the first confounder:

$$\mu(x,z) = \psi_0 z + \psi_1 z^2 + \psi_2 z x_1 + \psi_3 z^2 x_1 + \mathbf{x}_0\beta.$$

We first fit the propensity score model with augmenting term

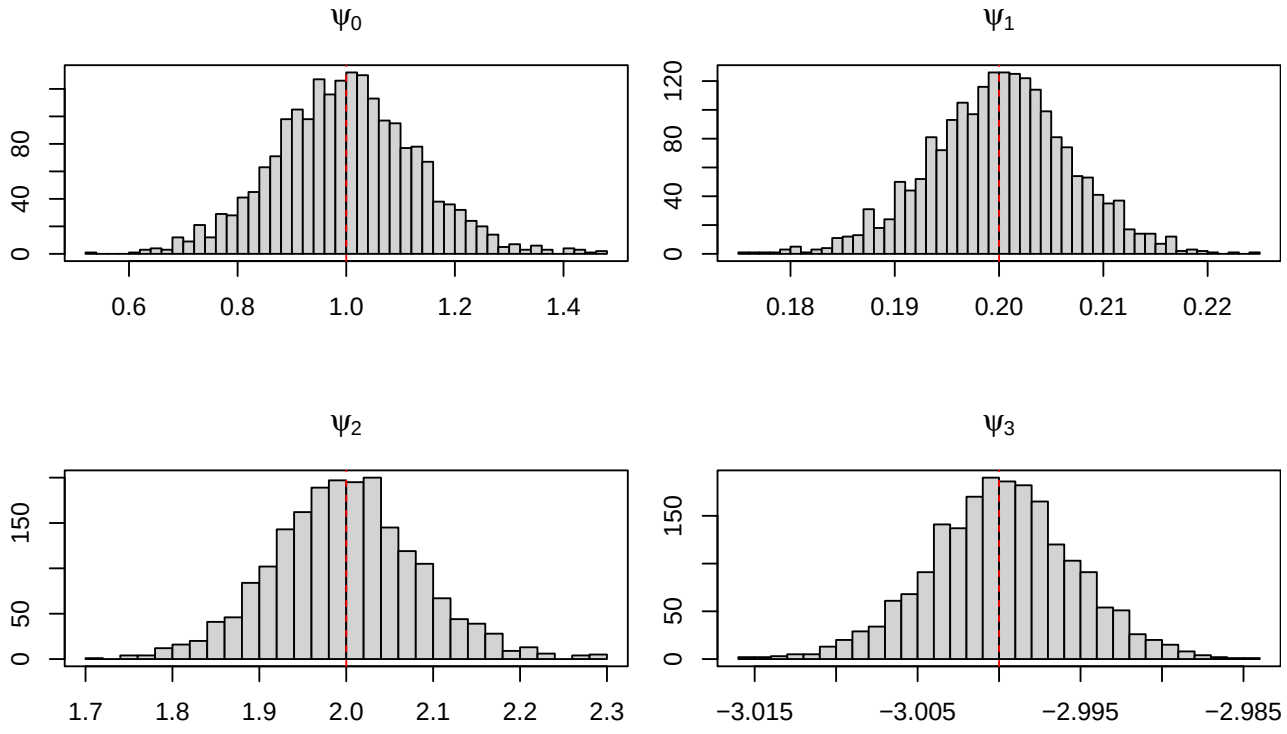
$$\phi_0 b_1(x) + \phi_1 b_1(x) x_1 + \phi_2 b_2(x) + \phi_3 b_2(x) x_1$$

```
#Monte Carlo study
n<-1000
nreps<-2000
gest3.samp<-matrix(0,ncol=4,nrow=nreps)
psi<-c(1,0.2,2,-3)

for(irep in 1:nreps){
    X<-rmvn(n,mu=Mu,Sigma)
    Xal<-cbind(1,X)
    ps.true<-Xal %*% al
    Z<-rnorm(n,ps.true,sigZ)
    Zsq<-Z^2
    Xm<-cbind(1,X)
    X1<-X[,1]
    muval<-as.vector(Xm %*% be)+Z*(psi[1]+psi[3]*X1) + Z^2*(psi[2]+psi[4]*X1)
    Y<-rnorm(n,muval,sigY)

    ps1<-fitted(lm(Z~X))
    ps2<-fitted(lm(Z^2~X))

    fit1<-lm(Y~Z+Zsq+Z:X1+Zsq:X1+ps1+ps1:X1+ps2+ps2:X1)
    gest3.samp[irep,]<-coef(fit1)[c(2:3,6:7)]
}
```



We now fit the propensity score model with augmenting term

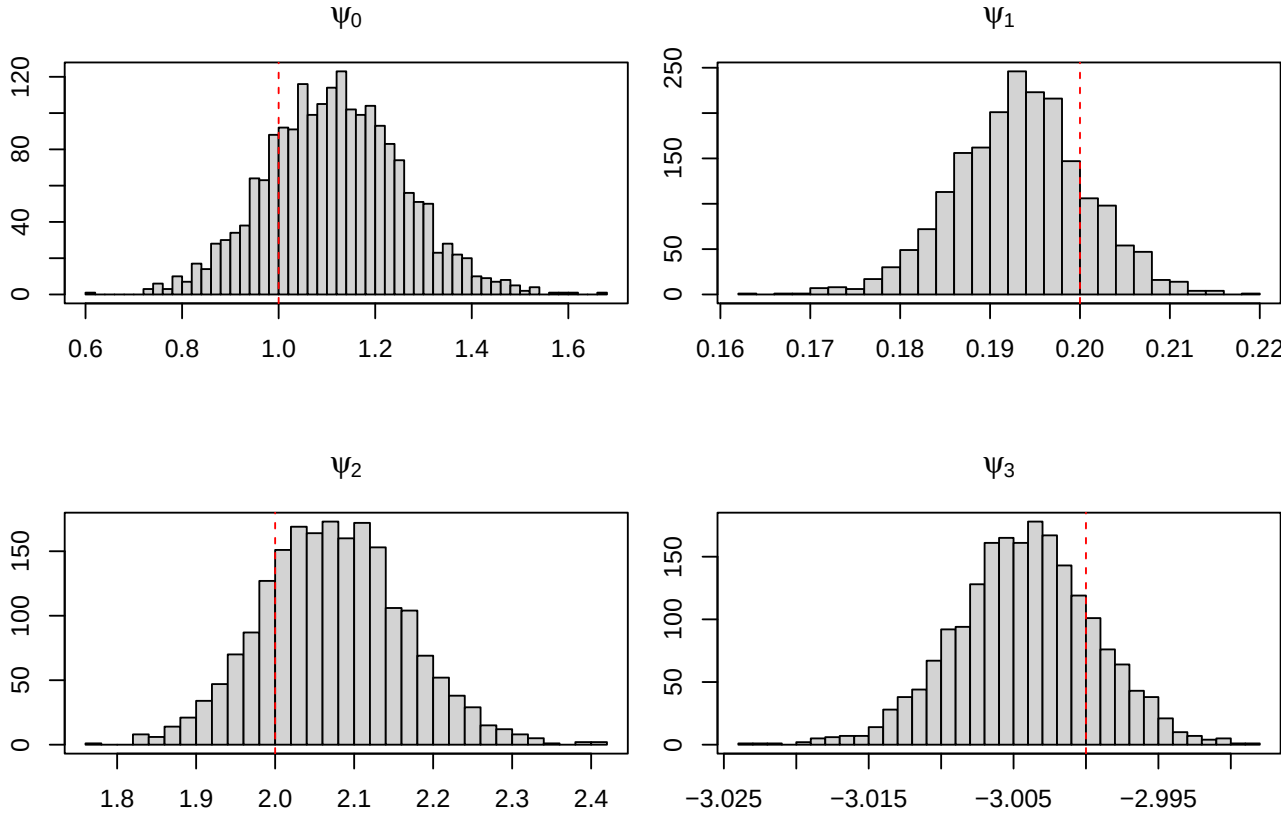
$$\phi_0 b_1(x) + \phi_1 b_1(x) x_1$$

```
#Monte Carlo study
n<-1000
nreps<-2000
gest4.samp<-matrix(0,ncol=4,nrow=nreps)
psi<-c(1,0.2,2,-3)

for(irep in 1:nreps){
    X<-rmvn(n,mu=Mu,Sigma)
    Xal<-cbind(1,X)
    ps.true<-Xal %*% al
    Z<-rnorm(n,ps.true,sigZ)
    Zsq<-Z^2
    Xm<-cbind(1,X)
    X1<-X[,1]
    muval<-as.vector(Xm %*% be)+Z*(psi[1]+psi[3]*X1) + Z^2*(psi[2]+psi[4]*X1)
    Y<-rnorm(n,muval,sigY)

    ps1<-fitted(lm(Z~X))
    ps2<-fitted(lm(Z^2~X))

    fit1<-lm(Y~Z+Zsq+Z:X1+Zsq:X1+ps1+ps1:X1)
    gest4.samp[irep,]<-coef(fit1)[c(2:3,5:6)]
}
```



In this case we do get bias as the path via the $Z^2 X_1$ interaction is not blocked by conditioning only on $b_1(X)$.