Estimating and Testing Causal Effects in the Pretest-Treatment-Posttest Design

Felix Flory

Friedrich Schiller University of Jena, Germany

Parameterization

Parameterization in orthogonal polynomials:

$$g_0(Z) = \beta_0 + \beta_1 Z + \beta_2 O_2 + \ldots + \beta_k O_k$$

$$g_1(Z) = \beta_{k+1} + \beta_{k+2} Z + \beta_{k+3} O_2 + \ldots + \beta_{2k} O_k.$$

$$E(Y|X, Z) = g_0(Z) + g_1(Z) \cdot X =$$

$$= \beta_0 + \beta_1 Z + \beta_2 O_2 + \ldots + \beta_k O_k + (\beta_{k+1} + \beta_{k+2} Z + \beta_{k+3} O_2 + \ldots + \beta_{2k} O_k.) \cdot X$$

$$= \beta_0 + \beta_1 Z + \beta_2 O_2 + \ldots + \beta_k O_k + \beta_{k+1} \cdot X + \beta_{k+2} ZX + \beta_{k+3} O_2 X + \ldots + \beta_{2k} O_k X$$

We see that the regression is linear in

$$(Z,\ O_2,\ \ldots, O_k,\ X,\ ZX,\ O_2 X,\ \ldots, O_k X\ ).$$

In Matrixnotation

$$E(Y \mid X, Z) = \begin{pmatrix} 1 & Z & O_2 & \dots & O_k & X & ZX & O_2X & \dots & O_kX \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \\ \beta_{k+1} \\ \beta_{k+2} \\ \beta_{k+3} \\ \vdots \\ \beta_{2k} \end{pmatrix}$$

Summarizing the beta coefficients up to a column-vector with dimension *2k* we write:

$$\beta' = \begin{pmatrix} \beta_0 & \beta_1 & \beta_2 & \dots & \beta_k & \beta_{k+1} & \beta_{k+2} & \beta_{k+3} & \dots & \beta_{2k} \end{pmatrix}$$

Matrixnotation

The Equations then simplify to:

$$g_0(Z) = \begin{pmatrix}1 & Z & O_2 & \ldots & O_k & 0 & 0 & 0 & \ldots & 0\end{pmatrix}\beta$$

$$g_1(Z) = \begin{pmatrix}0 & 0 & 0 & \ldots & 0 & 1 & Z & O_2 & \ldots & O_k\end{pmatrix}\beta$$

The *ACE* is then computed by

$$ACE = E[g_1(Z)] = E[\,\beta_{k+1} + \beta_{k+1} Z + \beta_{k+1} O_2 + \ldots + \beta_{2k} O_k] =$$

$$= \beta_{k+1} + \beta_{k+1} E(Z) + \beta_{k+1} E(O_2) + \ldots + \beta_{2k} E(O_k)$$

$$= \begin{pmatrix}0 & 0 & 0 & \ldots & 0 & 1 & E(Z) & E(O_2) & \ldots & E(O_k)\end{pmatrix}\beta$$

Matrixnotation

$$ACE = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 1 & E(Z) & E(O_2) & \dots & E(O_k) \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \\ \beta_{k+1} \\ \beta_{k+2} \\ \beta_{k+3} \\ \vdots \\ \beta_{2k} \end{pmatrix}$$

Estimation of Regression Coefficients

To estimate the parameters of β we specify the design matrix:

$$\boldsymbol{X} = \begin{pmatrix} 1 & z_1 & o_{2\,1} & \dots & o_{k\,1} & x_1 & z_1 x_1 & o_{2\,1} x_1 & \dots & o_{k\,1} x_1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & z_N & o_{2\,N} & \dots & o_{k\,N} & x_N & z_1 x_N & o_{2\,N} x_N & \dots & o_{k\,N} x_N \end{pmatrix}$$

Let $\boldsymbol{y}' = (y_1 \ \dots \ y_N)$ be the column vector of the observed outcome values

of each person. To estimate the parameters summarized in

$$\hat{\beta}' = \begin{pmatrix} \hat{\beta}_0 & \hat{\beta}_1 & \hat{\beta}_2 & \dots & \hat{\beta}_k & \hat{\beta}_{k+1} & \hat{\beta}_{k+2} & \hat{\beta}_{k+3} & \dots & \hat{\beta}_{2k} \end{pmatrix}$$

we use the well known formula:

$$\hat{\beta} = (\boldsymbol{X}' \boldsymbol{X})^{-1} \boldsymbol{X}' \boldsymbol{y}$$

Estimation of the Causal Effects

The estimated conditional average causal effect of the treatment for a given covariate value z:

$$A\hat{C}E_{Z=z} = \hat{g}_1(z) = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 1 & z & o_2 & \dots & o_k \end{pmatrix} \hat{\beta}$$

The (unconditional) average causal effect $ACE = \mathbf{A}\beta$ is estimated by:

$$A\hat{C}E = \mathbf{A}\hat{\beta}$$

with $\mathbf{A} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 1 & E(Z) & E(O_2) & \dots & E(O_k) \end{pmatrix}$.

But these parameters are mostly unknown. Most often you will only have estimates for **A**

$$\hat{\mathbf{A}} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 1 & \hat{E}(Z) & \hat{E}(O_2) & \dots & \hat{E}(O_k) \end{pmatrix}$$

Significance Test of the *ACE*

For the hypothesis $\mathbf{A}\,\beta - \delta = 0$ the test statistic

$$F = \frac{(\mathbf{A}\hat{\beta} - \mathbf{A}\beta)'\,[\mathbf{A}\,(\boldsymbol{X}'\boldsymbol{X})^{-1}\,\mathbf{A}']^{-1}\,(\mathbf{A}\hat{\beta} - \mathbf{A}\beta)}{\hat{Q}_e \,/\, (N - 2k - 1)}$$

is F- distributed with $df_1 = 1$ and $df_2 = N - 2k - 1$.

Defining $ACE_{Sample} = \hat{\boldsymbol{A}}\,\beta$

and for the hypothesis $\hat{\boldsymbol{A}}\,\beta - \delta = 0$ the test statistic

$$F_{Sample} = \frac{(\hat{\boldsymbol{A}}\hat{\beta} - \hat{\boldsymbol{A}}\beta)'\,[\hat{\boldsymbol{A}}\,(\boldsymbol{X}'\boldsymbol{X})^{-1}\,\hat{\boldsymbol{A}}']^{-1}\,(\hat{\boldsymbol{A}}\hat{\beta} - \hat{\boldsymbol{A}}\beta)}{\hat{Q}_e \,/\, (N - 2k - 1)}$$

is F- distributed with $df_1 = 1$ and $df_2 = N - 2k - 1$

Unfortunately F_{Sample} is not the appropriate statistic for the hypothesis

$\mathbf{A}\,\beta - \delta = 0$

The corresponding t-Test for the ACE

The following test statistics are student-t-distributed with

df = N – rank($\boldsymbol{X}$):

$$t = \frac{\mathbf{A}\hat{\beta} - \mathbf{A}\beta}{\sigma_\psi} \qquad \text{with} \qquad \sigma_\psi = \sqrt{\frac{SSE}{N - rank(X)} \cdot \mathbf{A}(\boldsymbol{X}'\boldsymbol{X})^{-1}\mathbf{A}'}$$

$$t_{Sample} = \frac{\hat{\mathbf{A}}\hat{\beta} - \hat{\mathbf{A}}\beta}{\sigma_{\psi Sample}} \qquad \text{with} \qquad \sigma_{\psi Sample} = \sqrt{\frac{SSE}{N - rank(X)} \cdot \hat{\mathbf{A}}(\boldsymbol{X}'\boldsymbol{X})^{-1}\hat{\mathbf{A}}'}$$

However we need the test statistic for the Hypothesis $\hat{\mathbf{A}}\hat{\beta} - \mathbf{A}\beta = 0$

Simulation Generating Model

$$Z \sim N(0,1)$$

$$X = \begin{cases} 1 & \text{for } U \leq P(X = 1 \mid Z = z) \\ 0 & \text{otherwise} \end{cases}$$

$$U \sim [0,1] \qquad \text{and} \qquad P(X = 1 \mid Z = z) = \frac{e^{0+0.5 \cdot z}}{1 + e^{0+0.5 \cdot z}}$$

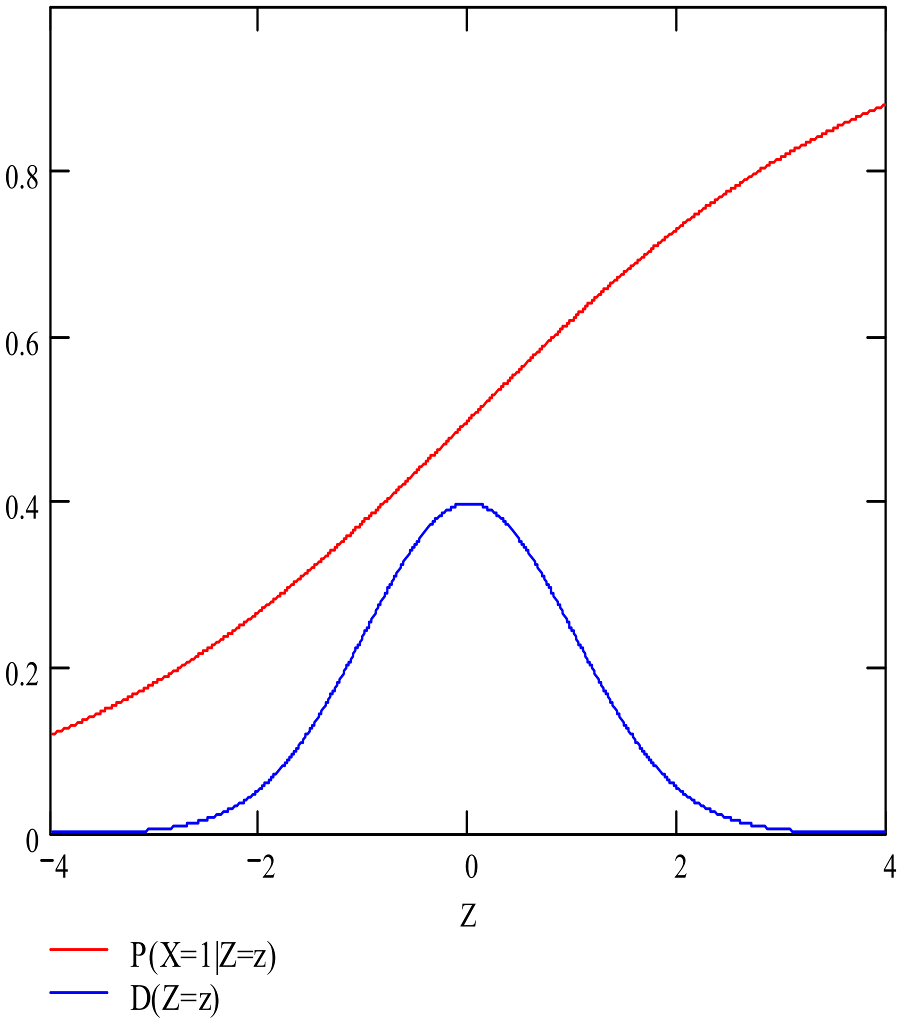
$$Y = E(Y|X, Z) + \varepsilon \qquad \text{with } \varepsilon \sim N(0, 2)$$

$$E(Y|X, Z) = g_0(Z) + g_1(Z) \cdot X$$

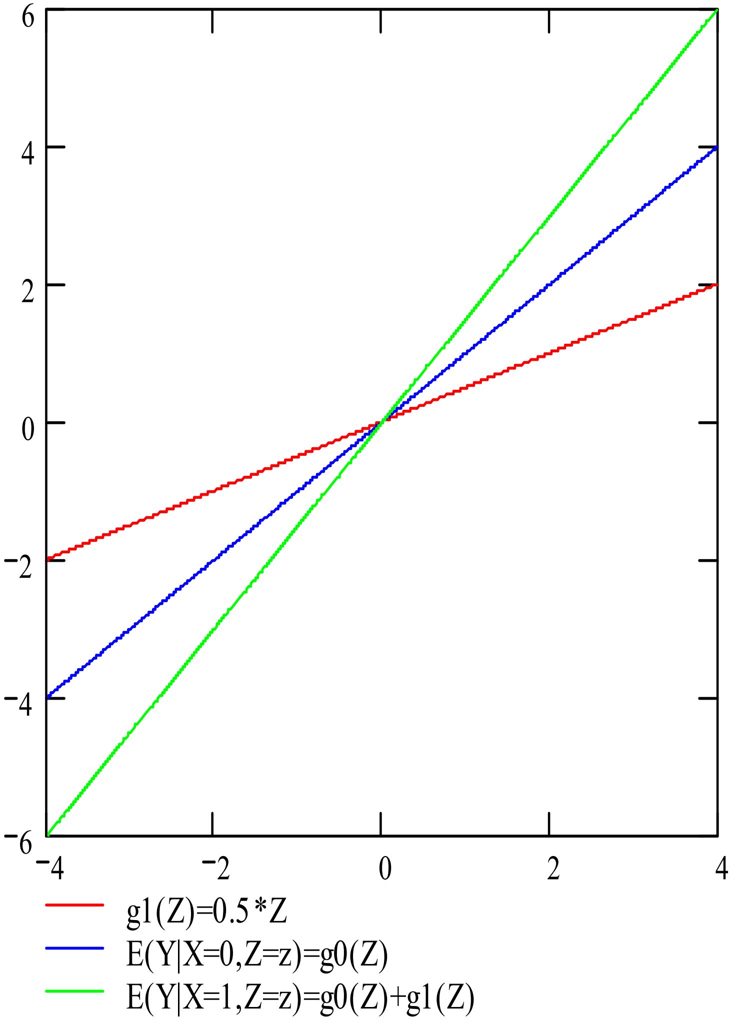
$$g_0(Z) = Z$$

$$g_1(Z) = DE \cdot Z \qquad \rightarrow \qquad E(Y|X, Z) = Z + ZX$$

Diagrams



0.8
0.6
0.4
0.2
0
−4
−2
0
2
4
Z
P(X=1|Z=z)
D(Z=z)



6
4
2
0
−2
−4
−6
−4
−2
0
2
4
g1(Z)=0.5*Z
E(Y|X=0,Z=z)=g0(Z)
E(Y|X=1,Z=z)=g0(Z)+g1(Z)

Simulation Outcomes

Sample Size = 500 Number of Loops = 4000

DE	Var(g1)	Freq(sig)	mean(ACEhat)	mean($\sigma\psi$Sample)	std(ACEhat)
0	0	0.05010	-0.00039	0.18425	0.18386
0.5	0.25	0.04625	0.00346	0.18457	0.18481
1	1	0.05975	-0.00187	0.18427	0.19280
1.5	2.25	0.06500	0.00072	0.18444	0.19812
2	4	0.08475	-0.00125	0.18440	0.20793
2.5	6.25	0.09700	0.00652	0.18436	0.21653
3	9	0.11525	-0.00520	0.18453	0.22676
3.5	12.25	0.12775	0.00205	0.18442	0.23932
4	16	0.16300	0.00298	0.18463	0.25925
5	25	0.22375	-0.00312	0.18455	0.29318
6	36	0.25350	0.00480	0.18454	0.31920
8	64	0.37400	-0.00609	0.18458	0.41039
10	100	0.46225	0.00183	0.18438	0.48734
20	400	0.69175	0.00571	0.18451	0.91222