Calibration demo

2023-08-09

Introduction

This document presents a demonstration of calibration methods for treatment effect predictors. We generate synthetic data with covariates, treatment assignments, and outcomes. We explore two calibration methods: the Best Linear Predictor (BLP) and causal isotonic calibration.

The BLP method provides a linear calibration gaurantees by learning an optimal linear transformation of the original predictor, so that the linearly calibrated predictor cannot be improved by applying any linear transformation (i.e., scaling and shifting).

Isotonic calibration offers non-parametric (distribution-free) calibration gaurantees by (1) an optimal monotone transformation of the original predictor and (2) providing a calibrated predictor that cannot be improved by applying any transformation (linear or nonlinear).

Data Generation

We begin by generating synthetic data for the demonstration. We set the seed for reproducibility and create variables for covariates, treatment assignments, potential outcomes, observed outcomes, and the conditional average treatment effect (CATE).

```r
# Set random seed for reproducibility
set.seed(12345)
n <- 2000

# Generate covariate W from a uniform distribution between -1 and 1
W <- runif(n, -1, 1)

# Calculate treatment assignment probabilities using logistic function
pi <- plogis(0.5 * W)
A <- rbinom(n, size = 1, pi)

# Define outcome regression functions and CATE
mu0 <- plogis(W)
mu1 <- plogis(1 + 2 * W)
cate <- mu1 - mu0

# Generate potential outcomes based on treatment assignment
Y0 <- rbinom(n, size = 1, mu0)
Y1 <- rbinom(n, size = 1, plogis(1 + 2 * W))

# Create observed outcomes based on treatment assignment
Y <- ifelse(A == 1, Y1, Y0)
```

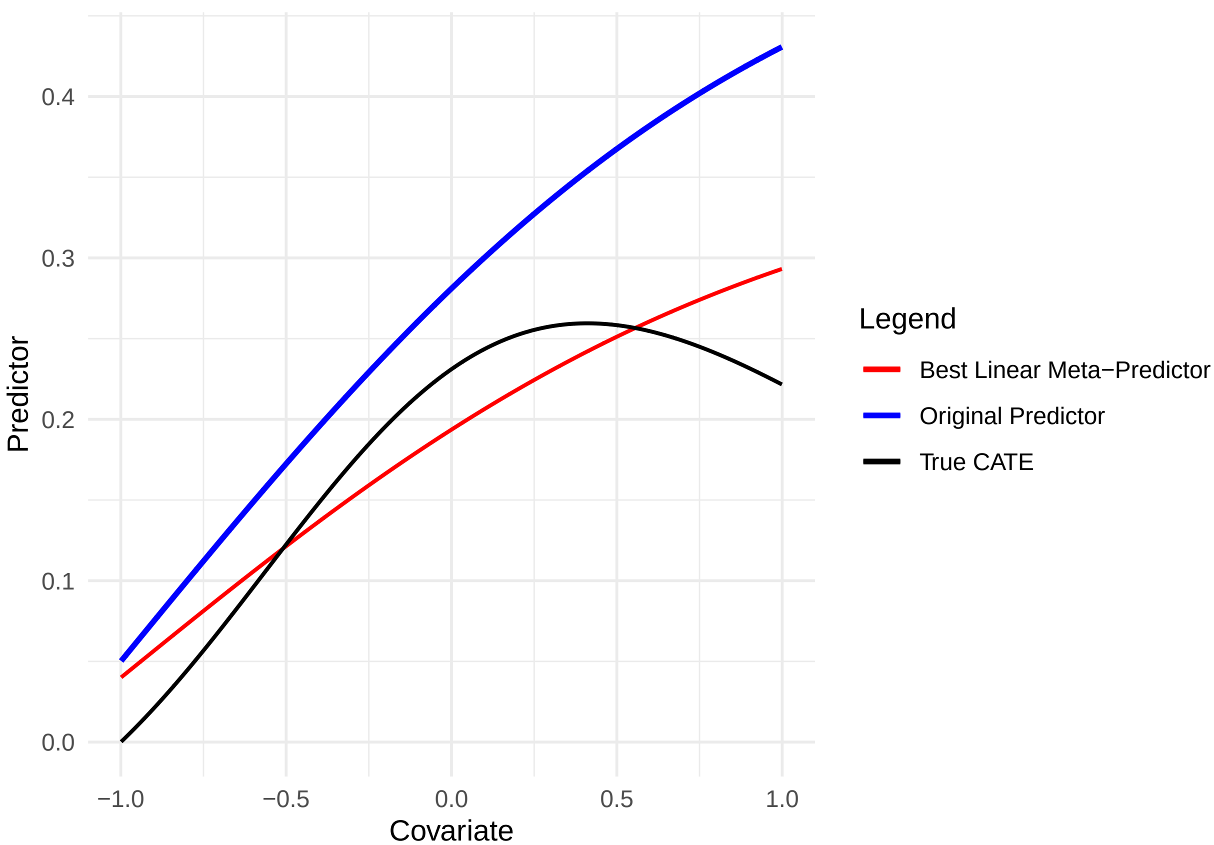
Initial predictor and calibration plot

We first create an initial predictor of the Individual Treatment Effect (ITE), denoted as tau.hat, which is a fixed function for simplicity.

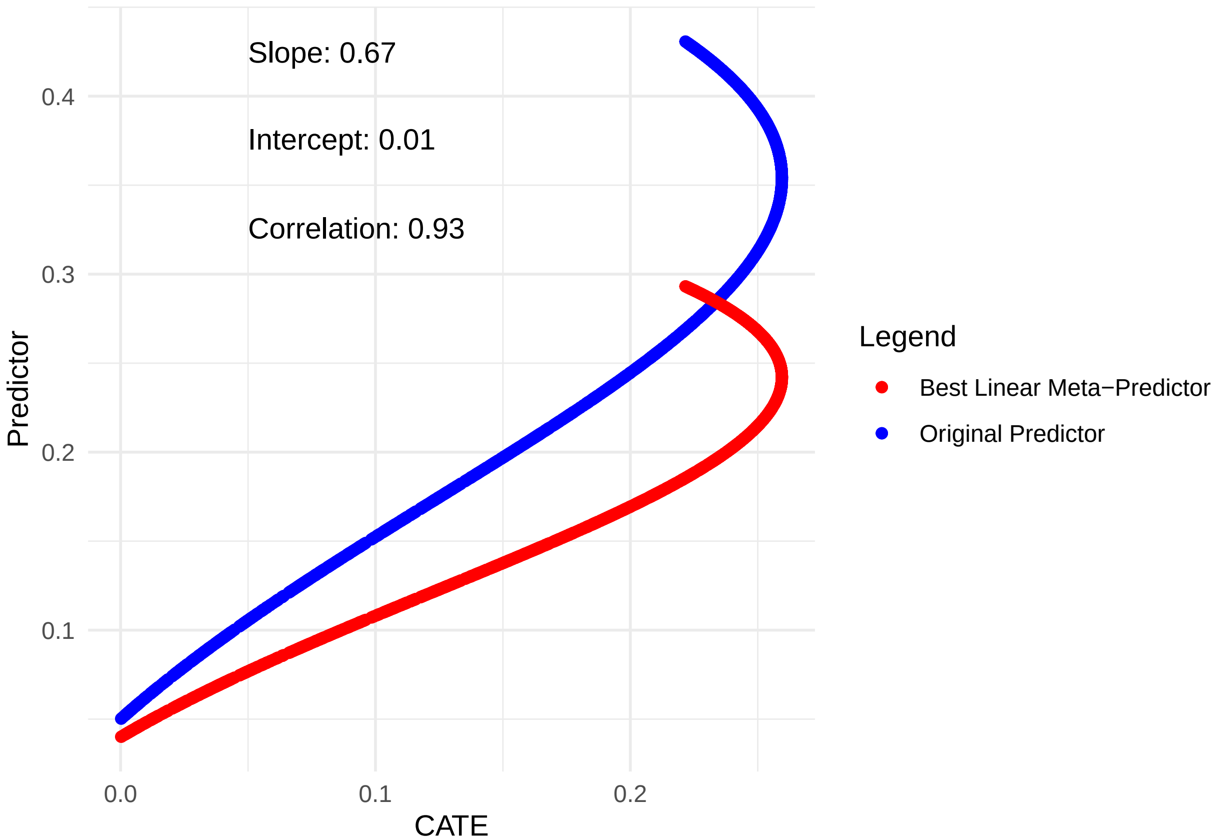
```
# use machine learning to obtain initial predictor of ITE Y_1 - Y_0
# for simplicity, we define our predictor tau.hat as a fixed function.

tau.hat <- plogis(1 + W  ) - 0.45
```

We visualize the initial predictor, the Best Linear Predictor (BLP), and the true CATE as functions of the covariate.

```
## Warning: Using ‘size‘ aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use ‘linewidth‘ instead.
## This warning is displayed once every 8 hours.
## Call ‘lifecycle::last_lifecycle_warnings()‘ to see where this warning was
## generated.
```



Now, lets make a calibrate plot. This is a scatter plot of the true CATE values vs the predicted CATE values.



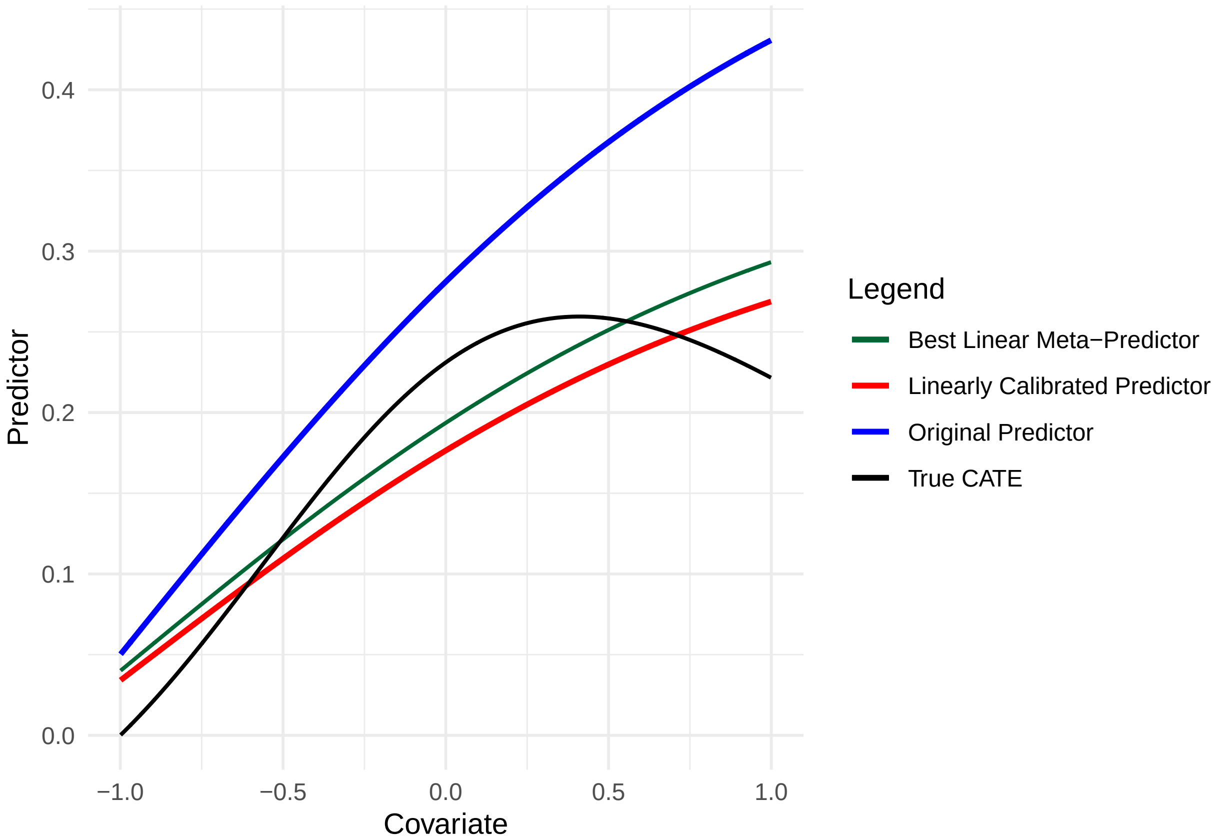
Linear calibration with Victor's BLP method

We apply the Best Linear Predictor (BLP) method to linearly calibrate the original predictor.

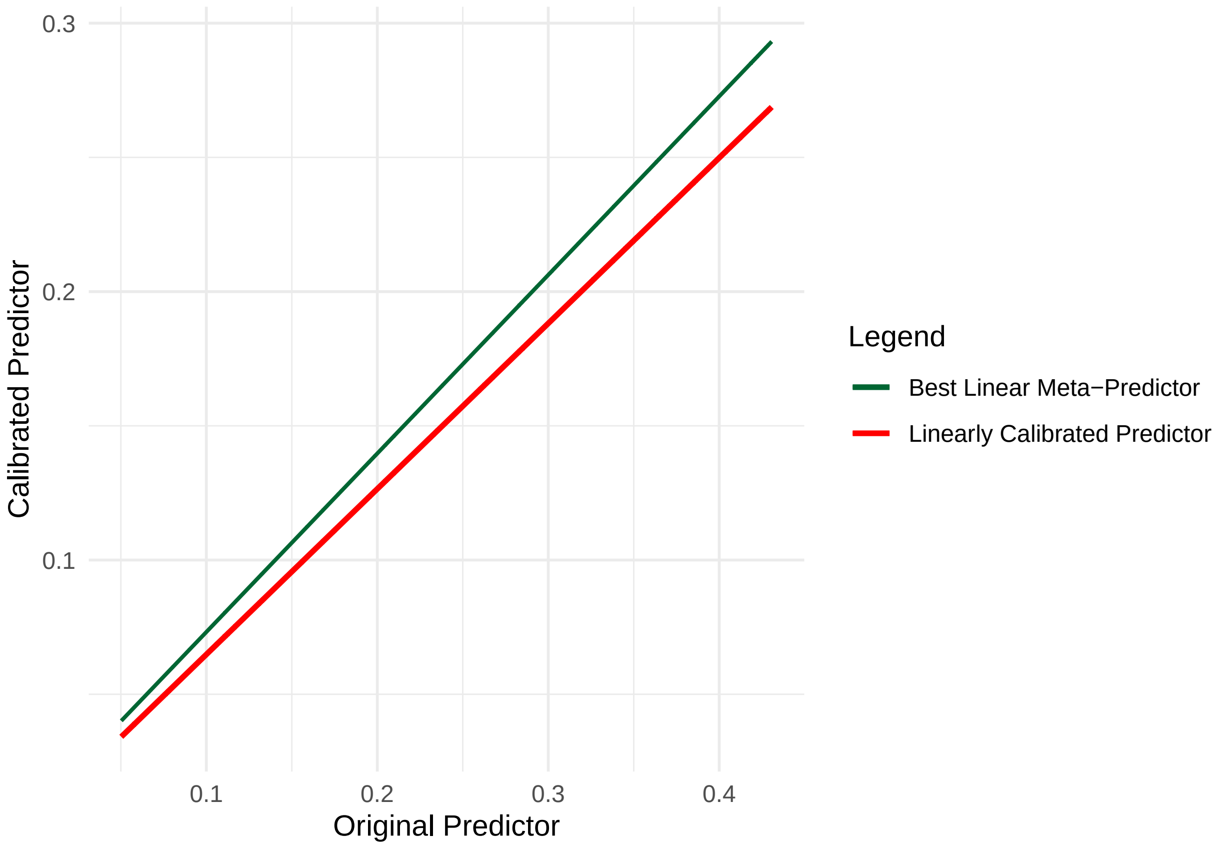
```
# unbiased surrogate outcome for CATE/ITE
pseudo_outcome <- cate + (A/pi) * (Y - mu1) - ((1-A)/(1-pi)) * (Y - mu0)

# fit best linear predictor of tau.hat of the surrogate outcome
# provides estimat of BLP of ITE/CATE
fit <- lm(pseudo_outcome ~ tau.hat, data = data.frame(tau.hat, pseudo_outcome))
intercept <- coef(fit)[1]
slope <- coef(fit)[2]

# get linear calibrated predictor, i.e. BLP given tau.hat
tau.BLP.hat <- intercept + slope * tau.hat





cor_tau <-  cor(tau.BLP.hat, cate)
# Calculate the regression coefficients
fit <- lm(cate ~ tau.hat, data = data.frame(tau.hat, cate))
intercept <- coef(fit)[1]
slope <- coef(fit)[2]
```

```
tau.BLP.oracle <- intercept + tau.hat * slope
```

We visualize the linearly calibrated predictor as a function of the covariate.



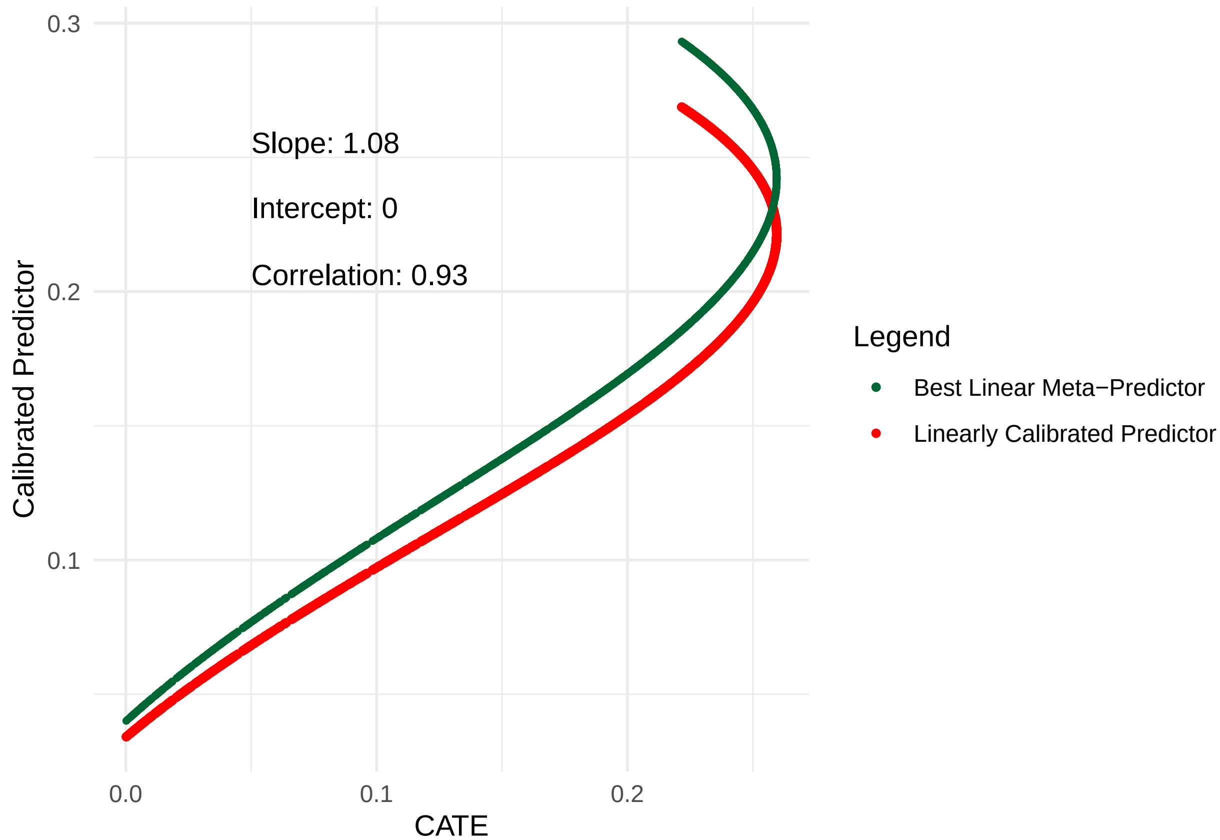
We see that the linearly calibrated predictor is a linear transformation of the original predictor.



Check calibration of BLP-corrected predictor

Lets check the new calibration plot of the linearly calibrated predictor.

The following plot shows how the treatment effect predictions assigned to individuals (y-axis) varies as a function of the actual conditional average treatment effect of the individuals.



We see now that the linearly calibrated predictor is well calibrated in a linear sense. The fit cannot be improved by fitting a linear model on top of the predictor. We have correctly estimated the BLP.

However, we see that the BLP approach is unable to correct for the poor calibration (flattening) at the end regions. This is because BLP is parametric and can only calibrate by applying linear transformations to the predictor. A more nonparametric approach can correct for nonlinear transformation.

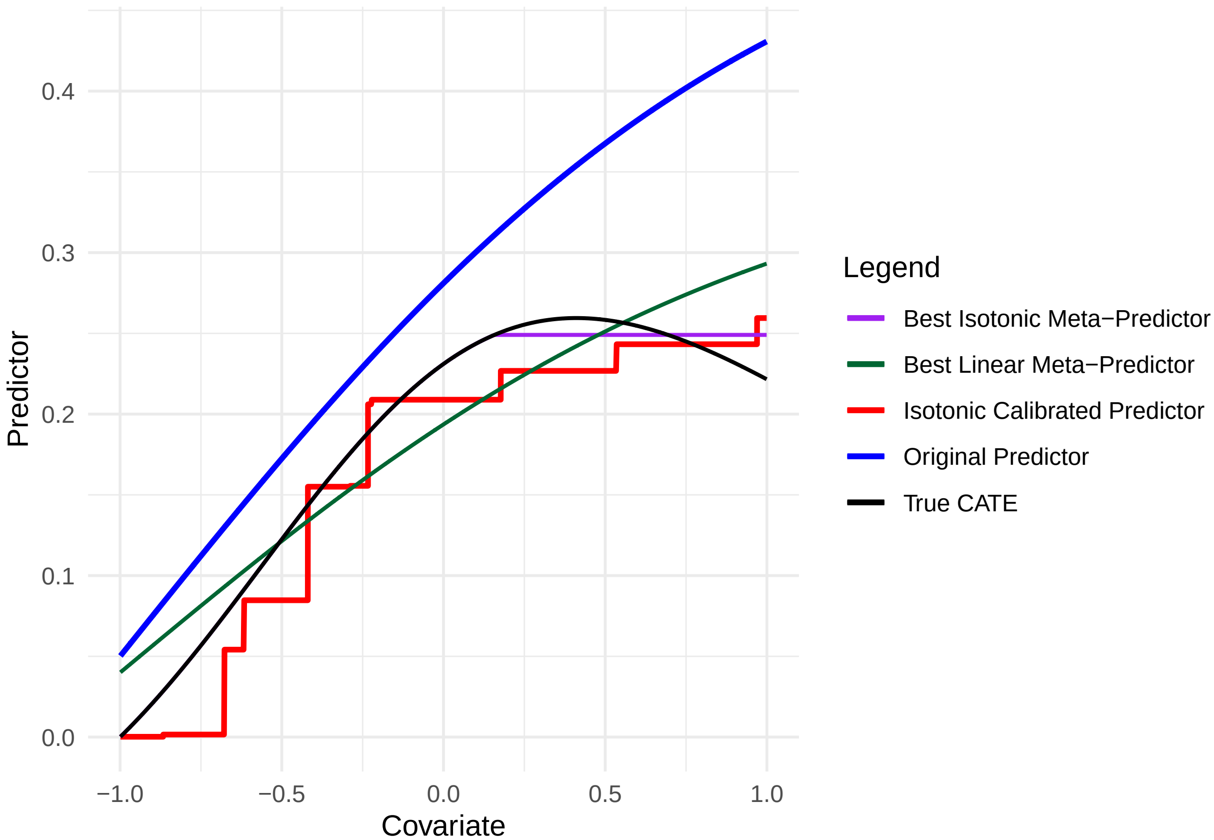
Causal isotonic calibration

Next, we introduce causal isotonic calibration, a nonparametric method that can correct for any monotone transformations of the original predictor and provides distribution-free calibration gaurantees.

Since linear transformations (with positive slope) are monotone, this method is strictly more powerful than the BLP approach.

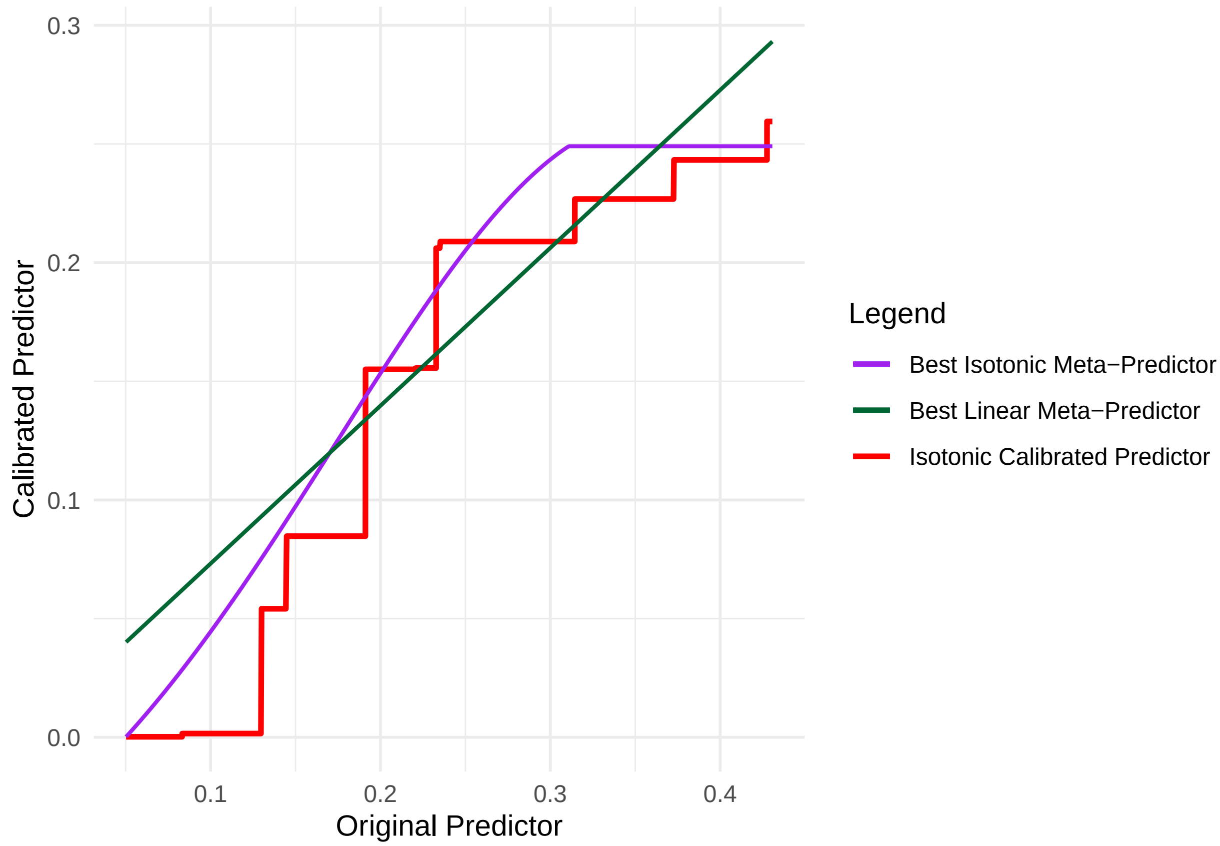
```
# unbiased surrogate outcome for CATE/ITE
pseudo_outcome <- cate + (A/pi) * (Y - mu1) - ((1-A)/(1-pi)) * (Y - mu0)
# fit isotonic regression of pseudo outcome
tau.iso.hat <- as.stepfun(isoreg(tau.hat, pseudo_outcome))(tau.hat)
# note isotonic regression overfits at very end boundaries of tau.hat
# tmp fix: bound predictions into range of cate.
tau.iso.hat <- pmin(pmax(tau.iso.hat, min(cate)), max(cate))
```

We visualize the isotonic calibrated predictor as a function of the covariate.



A benefit of isotonic calibration is that it automatically does data-driven heterogeneous treatment effect subgroup identification. The piece-wise constant values of the calibrated predictor defines subgroups (or bins) of individuals. Individuals with a given bin have conditional treatment effects that are quantitatively similar (given only information provided by the original predictor). Moreover, the conditional treatment effect of a given subgroup is meaningfully different form those of other subgroups.

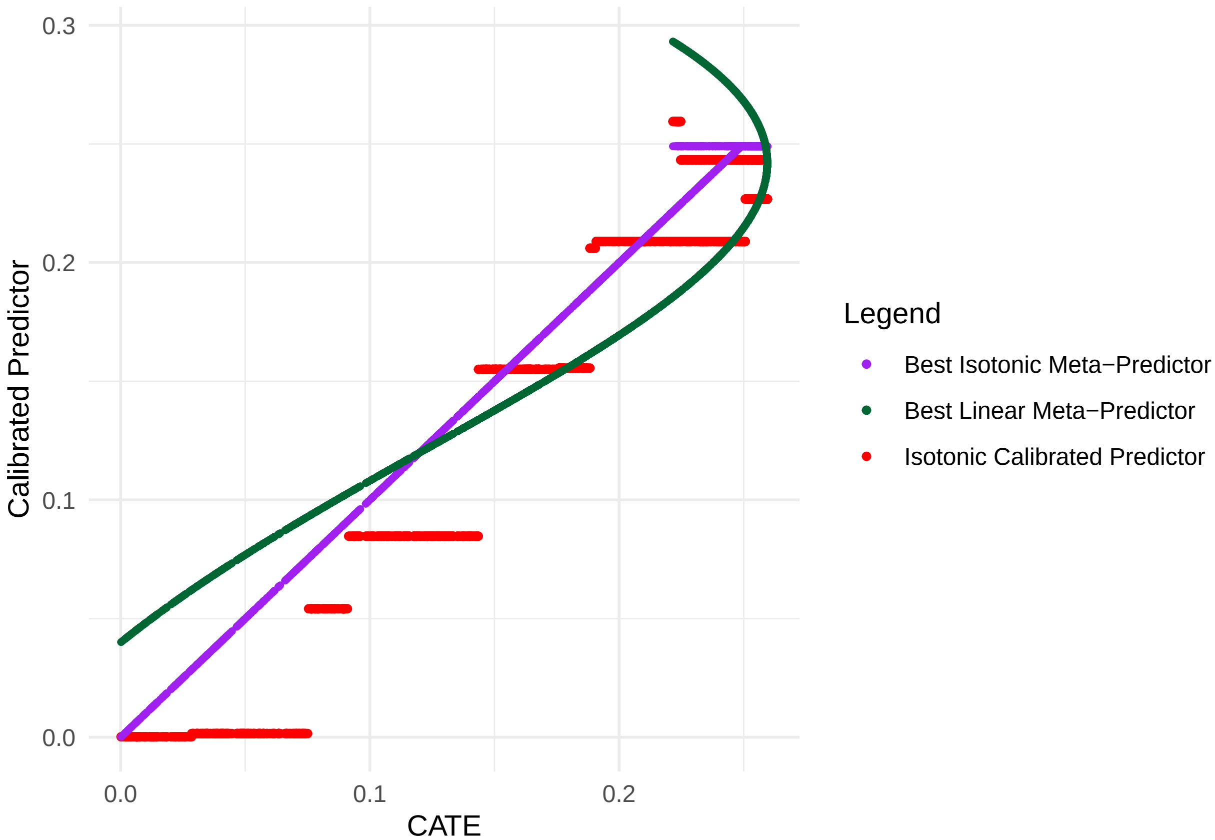
We see that the isotonic calibrated predictor is a monotone piece-wise constant transformation of the original predictor.



Check calibration of isotonic-corrected predictor

Lets check the new calibration plot of the isotonic calibrated predictor.

The following plot shows how the treatment effect predictions assigned to individuals (y-axis) varies as a function of the actual conditional average treatment effect of the individuals.



We see now that the linearly calibrated predictor is well calibrated in a linear sense. The fit cannot be improved by fitting a linear model on top of the predictor. We have correctly estimated the BLP.

However, we see that the BLP approach is unable to correct for the poor calibration (flattening) at the end regions. This is because BLP is parametric and can only calibrate by applying linear transformations to the predictor. A more nonparametric approach can correct for nonlinear transformation.