Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Causally Learning an Optimal Rework Policy

Oliver Schacht[+°], Sven Klaassen[+°], Philipp Schwarz[*°], Martin Spindler[+°], Sebastian Imhof[*], Daniel Grünbaum[*]

[*]ams Osram, [+]Economic AI, [°]University of Hamburg

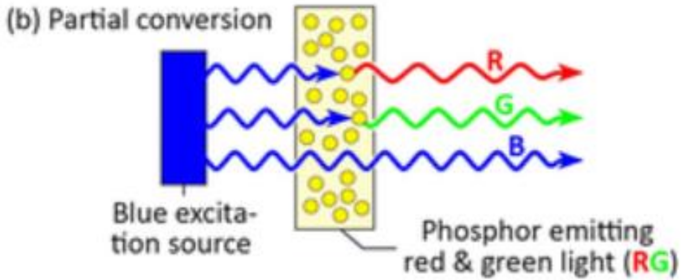
2023 Causal Data Science Meeting, online
November 8, 2023

Outline

- Rework process in opto-electronic semiconductor manufacturing
- Technical background about double/debiased machine learning
- Application to rework-policy learning and results

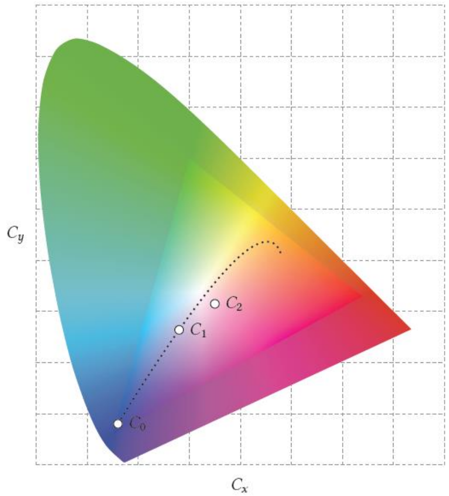
Phosphorous Conversion Process

- Light-emitting diods (LEDs) are opto-electronic semiconductors naturally emitting a blue wavelength
- Customers however demand a broad range of desired target color points, i.e. emitting a white light.
- The process of phosphorous **coating** applies multiple layers to the chips, shifting the wavelength towards white light in the color spectrum.



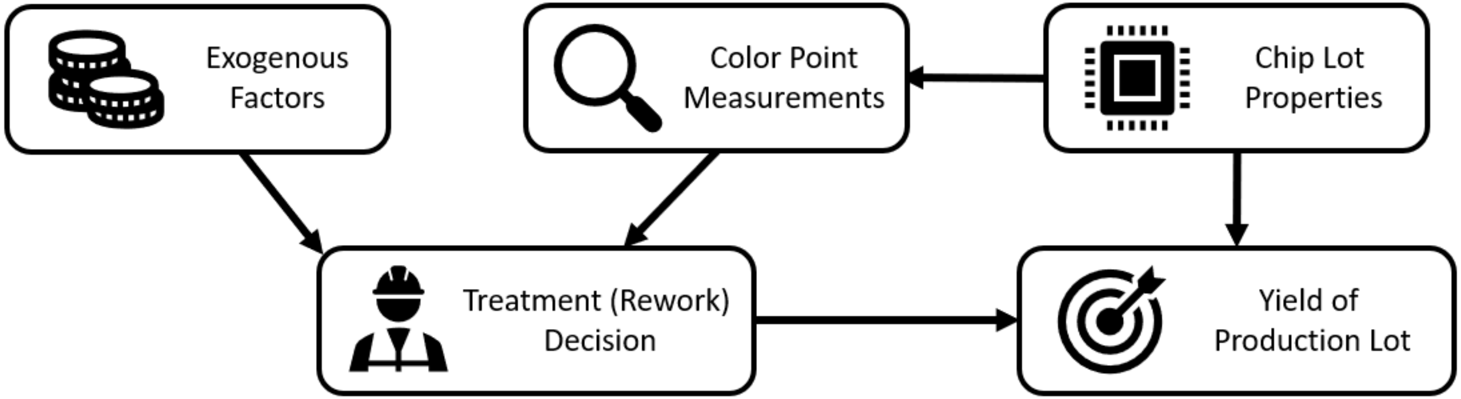
Source: Jaehee Cho, Jun Hyuk Park, Jong Kyu Kim, and E. Fred Schubert. White light-emitting diodes: History, progress, and future. Laser & Photonics Reviews, 11, 2017.

Rework Process



- Rework is a production step which **repeats a previous production stage** with adjustments
- **In the observed data, the rework decision is purely made by visual inspection of the point C_1**

Causal Model



Double Machine Learning

- **Double/debiased machine learning (DML)** is a framework based on machine learning tools for causal inference and estimation of treatment effects introduced by Chernozhukov et. al (2018).
- Resulting estimator has good properties ($\sqrt{N}$-consistency, approx. Gaussian)
- Growing number of estimators and models available in python and R package DoubleML [Bach et. al (2022)].

DoubleML

The Key Ingredients of DML

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

1. Neyman Orthogonality

   Inference is based on a method-of-moments estimator that obeys the **Neyman orthogonality condition.**

2. High-Quality Machine Learning Estimators

   The nuisance parameters are estimated with high-quality (fast-enough converging) machine learning methods.

3. Sample Splitting

   To avoid the biases arising from overfitting, a form of **sample splitting** is used.

Data

- $n = 32{,}669$ observations of independent chip lots
- We use 72 measurements as covariates X per chip lot
- A being the Action / Treatment Variable $A_i \in \{0,1\}$
- Y being the observed outcome (% of chips usable from lot i at the end of the process)

Interactive Regression Model

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

$$
\begin{aligned}
Y &= g_0(A, X) + U, \qquad \mathbb{E}[\,U \mid X, A\,] = 0 \\
A &= m_0(X) + V, \qquad \mathbb{E}[\,V \mid X\,] = 0
\end{aligned}
$$

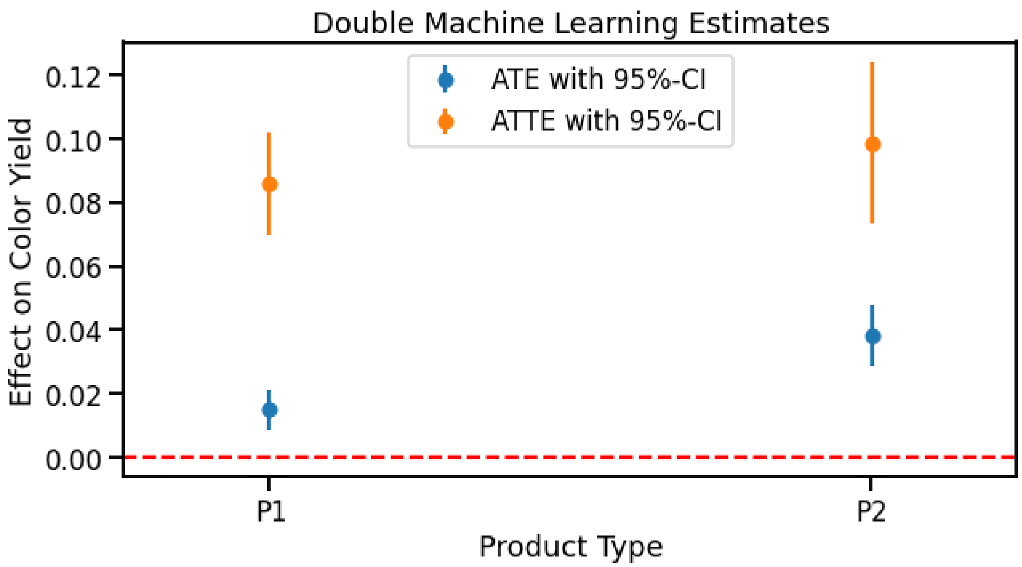
We estimate (with ML)

$$
\begin{aligned}
\hat{g}(A, X) &= \mathbb{E}[\,Y \mid A, X\,] \\
\widehat{m}(X) &= \mathbb{E}[A \mid X]
\end{aligned}
$$

To form a score (for ATE)

$$
\begin{aligned}
&\psi(W_i; \theta, \eta) \coloneqq \psi_a(W_i, \eta)\theta + \psi_b(W_i, \eta) \\
&= -\theta + g(1, X_i) - g(0, X_i) + \frac{A_i\big(Y_i - g(1, X_i)\big)}{m(X_i)} - \frac{(1 - A_i)(Y_i - g(0, X_i))}{1 - m(X_i)}
\end{aligned}
$$

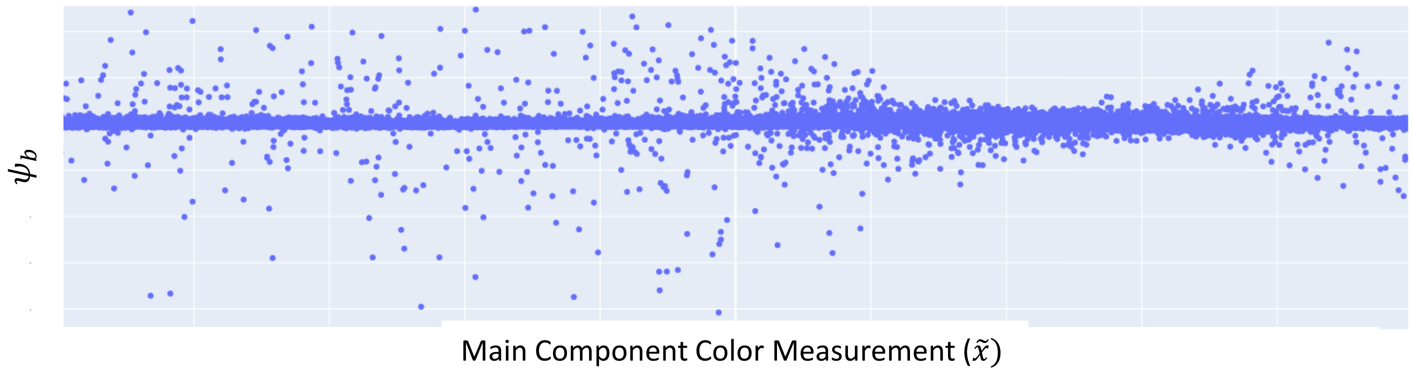
ATE and ATTE



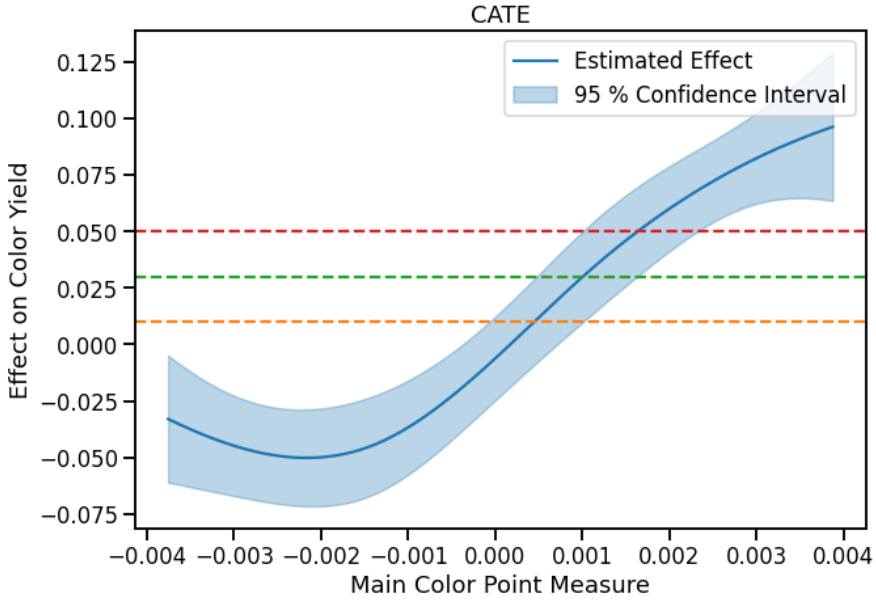
→ Effect on treated larger than on average – How can we derive a policy?

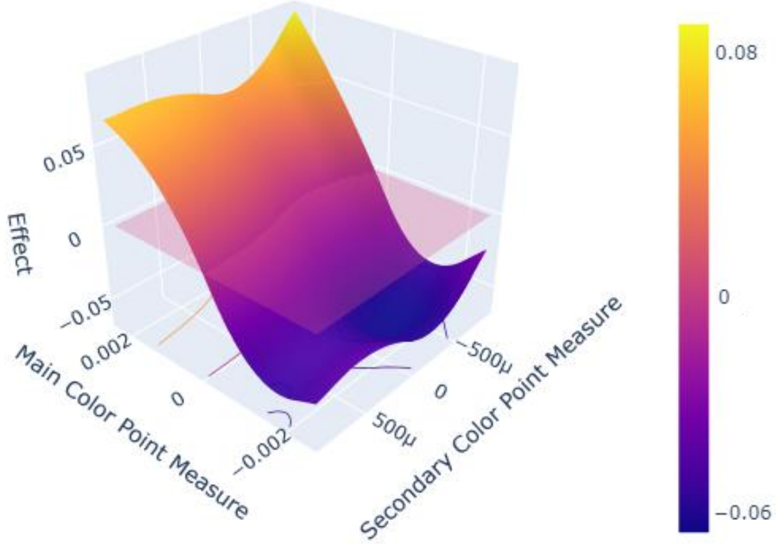
CATE Estimation with Double Machine Learning

- Conditional average treatment effects can be derived from ψ_b
- Projection onto a predefined basis vector $b(\tilde{x})$ enables us to approximate a conditional average treatment effect (CATE) $\theta_0(\tilde{x}) = b(\tilde{x})^{\mathrm{T}}\beta$.



Policies based on the CATE

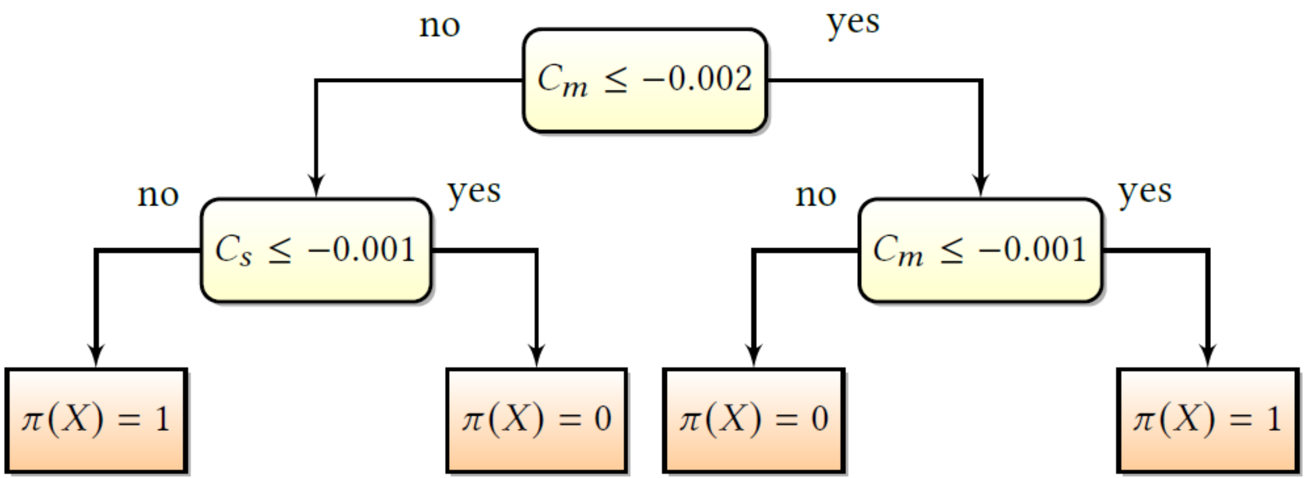




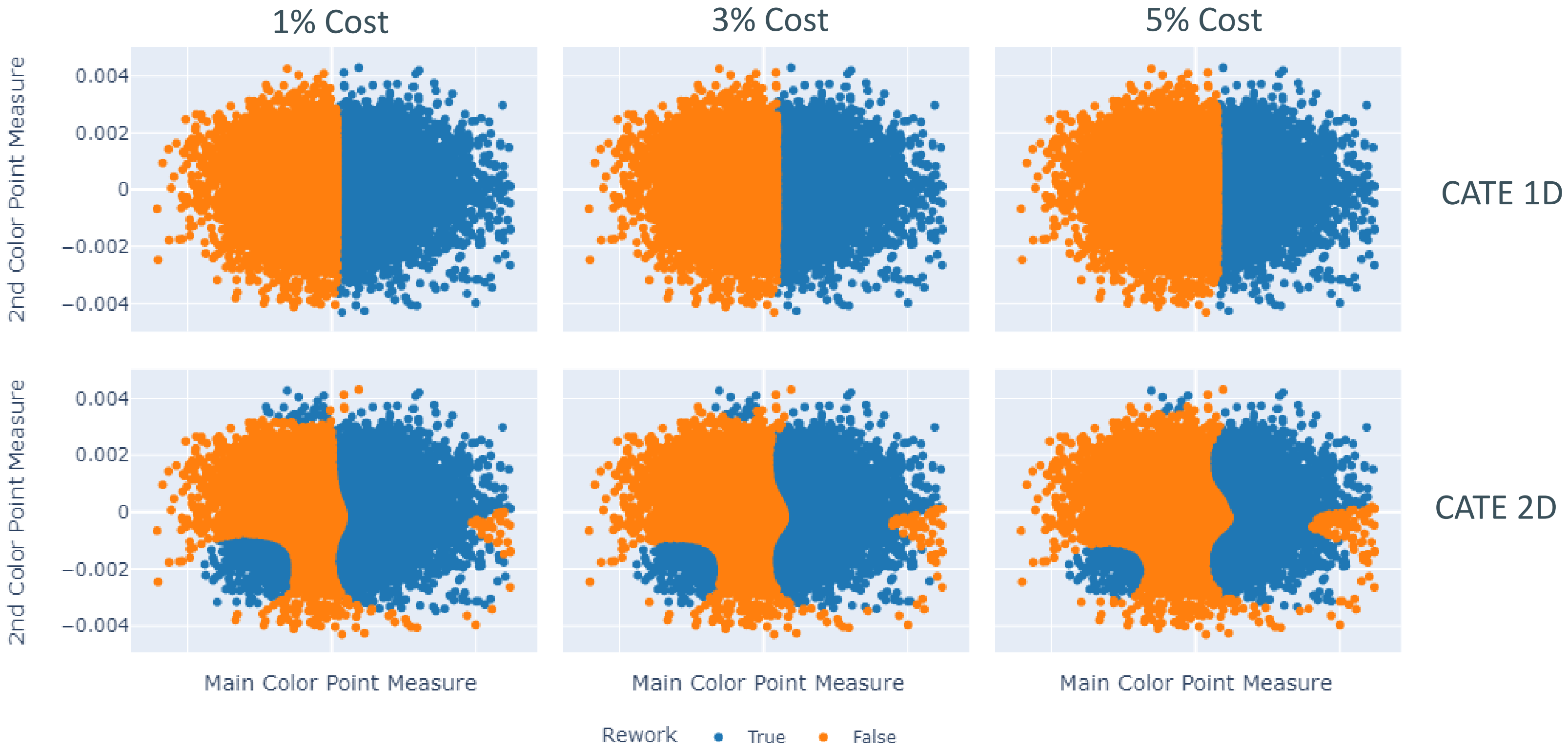
Policy Learning by classification

- Given ψ_b we can directly estimate the optimal policy over a policy class Π
$$\hat{\pi} = \arg\max_{\pi \in \Pi} \frac{1}{n} \sum_{i=1}^{n} (2\pi(X_i) - 1)\psi_b(W_i, \hat{\eta})$$
- This is equivalent to a weighted classification problem with target $\operatorname{sign}\left(\psi_b(W_i, \hat{\eta})\right)$ and sample weights $|\psi_b(W_i, \hat{\eta})|$.
- Athey and Wager (2021) propose an exact tree search here, which is computationally demanding ($\mathcal{O}(n^{depth})$).

Policytree

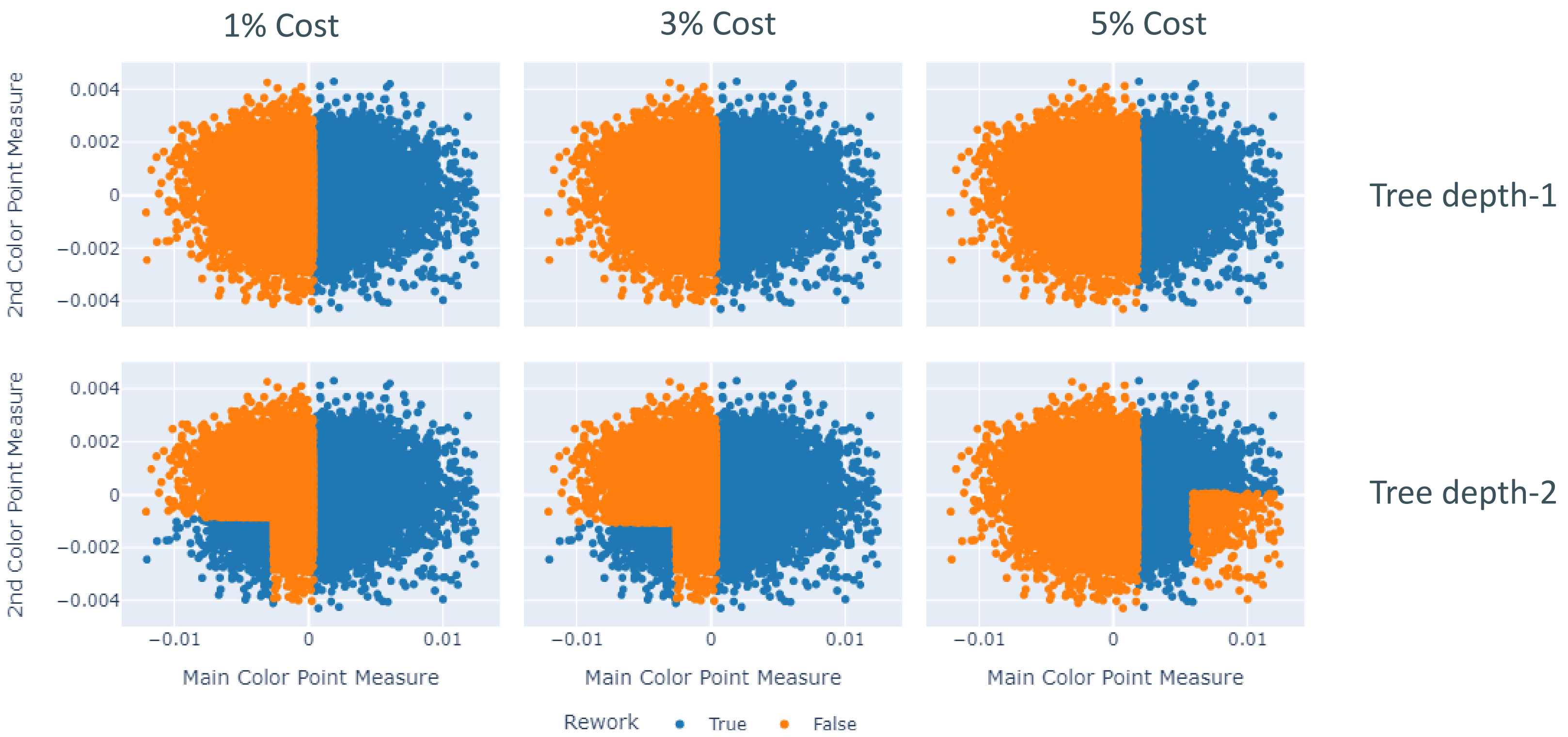

no
yes
$C_m \leq -0.002$
no
yes
$C_s \leq -0.001$
no
yes
$C_m \leq -0.001$
$\pi(X) = 1$
$\pi(X) = 0$
$\pi(X) = 0$
$\pi(X) = 1$

Results (1)



1% Cost
3% Cost
5% Cost
CATE 1D
CATE 2D
2nd Color Point Measure
0.004
0.002
0
−0.002
−0.004
Main Color Point Measure
Rework
True
False

Results (2)


1% Cost
3% Cost
5% Cost
Tree depth-1
Tree depth-2
2nd Color Point Measure
0.004
0.002
0
−0.002
−0.004
−0.01
0
0.01
Main Color Point Measure
Rework
True
False

Results (3)

Method	0	1	2
CATE 1D	0.0289	0.0265	0.0247
CATE 2D	0.0300	0.0292	0.0258
Depth-1 Tree	0.0301	0.0290	0.0242
Depth-2 Tree	0.0323	0.0312	0.0246

Table 3: Policy Value (Observed policy: 0.0178)

Method	Policy 0	Policy 1	Policy 2
CATE 1D	0.3864	0.3352	0.2806
CATE 2D	0.4179	0.3599	0.2973
Depth-1 Tree	0.4309	0.3842	0.2555
Depth-2 Tree	0.4708	0.4222	0.2393

Table 1: Share of reworked panels (Observed policy: 0.2101)

Take aways and future research

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

- Splitting along a “main component” is consistently estimated by multiple methods
- Secondary splits in areas with little overlap → large confidence bands

→ 0.62% more yield could be achieved by reworking 6.62% additional chip lots.

References

Susan Athey and Stefan Wager. **Policy learning with observational data**. *Econometrica*, 89(1):133–161, 2021.

Philipp Bach, Victor Chernozhukov, Malte S Kurz, and Martin Spindler. **DoubleML - An object-oriented implementation of double machine learning in python**. *JMLR*, 23:53–1, 2022.

Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. **Double/debiased machine learning for treatment and structural parameters**. *The Econometrics Journal*, 21(1):C1–C68, 01 2018

Oliver Schacht, Sven Klaassen, Philipp Schwarz, Martin Spindler, Daniel Grunbaum, and Sebastian Imhof. **Causally learning an optimal rework policy**. *Proceedings of The KDD'23 Workshop on Causal Discovery, Prediction and Decision, PMLR*, 2023.

Vira Semenova and Victor Chernozhukov. **Debiased machine learning of conditional average treatment effects and other causal functions**. *The Econometrics Journal*, 24(2):264–289, 2021.

Thank you.

For more info about DoubleML visit https://docs.doubleml.org/

We are open for your questions ☺

oliver.schacht@uni-hamburg.de

www.ams-osram.com

www.economicai.com