

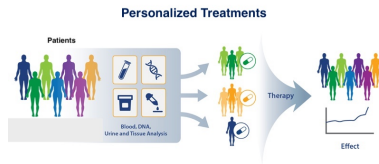
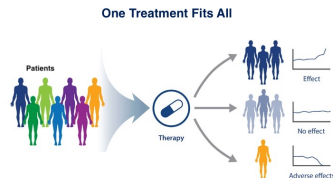
# Personalization with Latent Confounders

Leo Guelman

Head Statistician, DNA  
RBC Royal Bank

# Motivation for Personalization

- Personalization is founded on the premise that individuals have heterogeneous response to actions.
- Personalization algorithms aim to improve decision-making by identifying and exploiting this heterogeneity.

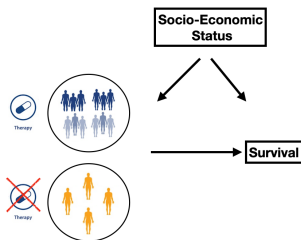


# Estimating Treatment Effects: **Non-Personalized Paradigm**

A/B Tests are 'gold standard' in the One-Treatment-Fits-All paradigm because they remove the influence of unobserved confounders (unmeasured variables that influence both the treatment and the outcome).

# Estimating Treatment Effects: **Non-Personalized Paradigm**

A/B Tests are 'gold standard' in the One-Treatment-Fits-All paradigm because they remove the influence of unobserved confounders (unmeasured variables that influence both the treatment and the outcome).



# Estimating Treatment Effects: **Non-Personalized Paradigm**

A/B Tests are 'gold standard' in the One-Treatment-Fits-All paradigm because they remove the influence of unobserved confounders (unmeasured variables that influence both the treatment and the outcome).

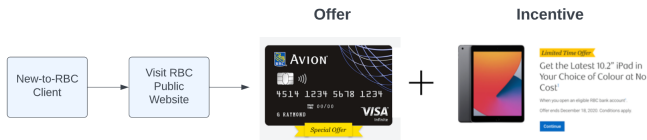


# Estimating Treatment Effects: **Personalized Paradigm**

- In the presence of unobserved confounders, experimental data is not 'gold standard' for estimating heterogeneous treatment effects (required for personalization)
- Experiments 'destroy' information that can be valuable to identify the values of unobserved confounders.
- Counterfactual-based decision making, which leads to a fusion of experimental and observational data, might be optimal for personalization.

# Motivating Example

- **Business objective:** Sell a credit card to new-to-RBC clients.
- **Current campaign: One-Treatment-Fits-All paradigm.** All new-to-RBC clients who visited the RBC public site get a credit card offer + iPad incentive.



- **Future campaign: Personalize the incentive.** Identify which new-to-RBC clients should receive an iPad incentive in the future to maximize the expected profitability of the campaign.

# Data Generating Process

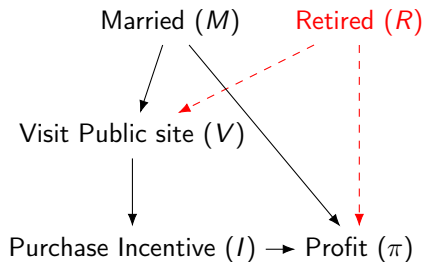


Figure: 'True' Causal Graph (current campaign).



# Data Generating Process

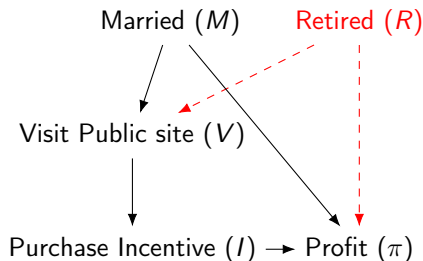


Figure: 'True' Causal Graph (current campaign).

$$P(R = 1) = 0.5, \quad P(M = 1) = 0.5$$

$$V := M \oplus R$$

$$I := V$$

	$R = 0$		$R = 1$	
	$M = 1$	$M = 0$	$M = 1$	$M = 0$
$I = 1$	0.25	0.50	0.45	0.05
$I = 0$	0.50	0.10	0.05	0.30

Table:  $E[\pi|M, R, I]$ . Highlighted cells reflect (new-to-RBC) client's 'natural' choice to visit the Public site or not.

# Four Approaches to Personalizing the Incentive

**Goal:** Identify which new-to-RBC clients should receive an iPad incentive in the future to maximize the expected profitability of the campaign.

Data Scientist 1



Data Scientist 2



Data Scientist 3



Data Scientist 4



# DS #1: Empirical Decision Criterion (EDC)

Associational Inference

$$\text{EDC} \rightarrow \operatorname{argmax}_{I \in 0,1} E[\pi | I, M]$$

# DS #1: Empirical Decision Criterion (EDC)

## Associational Inference

$$\text{EDC} \rightarrow \operatorname{argmax}_{I \in \{0,1\}} E[\pi | I, M]$$

$$E[\pi | I = 1, M = 1] = 0.25$$

$$E[\pi | I = 0, M = 1] = 0.05$$

$$E[\pi | I = 1, M = 0] = 0.05$$

$$E[\pi | I = 0, M = 0] = 0.10$$

	$R = 0$		$R = 1$	
	$M = 1$	$M = 0$	$M = 1$	$M = 0$
$I = 1$	0.25	0.50	0.45	0.05
$I = 0$	0.50	0.10	0.05	0.30

Table:  $E[\pi | M, R, I]$ .

# DS #1: Empirical Decision Criterion (EDC)

## Associational Inference

$$\text{EDC} \rightarrow \operatorname{argmax}_{I \in \{0,1\}} E[\pi|I, M]$$

$$E[\pi|I = 1, M = 1] = 0.25$$

$$E[\pi|I = 0, M = 1] = 0.05$$

$$E[\pi|I = 1, M = 0] = 0.05$$

$$E[\pi|I = 0, M = 0] = 0.10$$

	$R = 0$		$R = 1$	
	$M = 1$	$M = 0$	$M = 1$	$M = 0$
$I = 1$	0.25	0.50	0.45	0.05
$I = 0$	0.50	0.10	0.05	0.30

Table:  $E[\pi|M, R, I]$ .

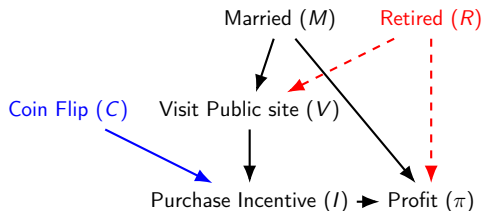
## Decision Rule:

- If Visit Site  $\wedge$  Married  $\rightarrow$  Purchase Incentive  $\rightarrow E[\pi] = 0.25$
- If Visit Site  $\wedge$  Not Married  $\rightarrow$  No Purchase Incentive  $\rightarrow E[\pi] = 0.30$

$$\text{Expected profit} = \boxed{0.275} = (0.25 + 0.30)/2.$$

# DS #2: Post-Visit Randomization (PVR)

Interventional Inference



$$V := M \oplus R$$

$$I := V \wedge C$$

Figure: Causal DAG with post-visit randomization.

# DS #2: Post-Visit Randomization (PVR) - cont'd

## Interventional Inference

$$\text{PVR} \rightarrow \operatorname{argmax}_{I \in \{0,1\}} E[\pi | \text{do}(I), M, V = 1]$$

## DS #2: Post-Visit Randomization (PVR) - cont'd

### Interventional Inference

$$\text{PVR} \rightarrow \operatorname{argmax}_{I \in \{0,1\}} E[\pi | \text{do}(I), M, V = 1]$$

$$E[\pi | \text{do}(I = 1), M = 1, V = 1] = 0.25$$

$$E[\pi | \text{do}(I = 0), M = 1, V = 1] = 0.50$$

$$E[\pi | \text{do}(I = 1), M = 0, V = 1] = 0.05$$

$$E[\pi | \text{do}(I = 0), M = 0, V = 1] = 0.30$$

	$R = 0$		$R = 1$	
	$M = 1$	$M = 0$	$M = 1$	$M = 0$
$I = 1$	0.25	0.50	0.45	0.05
$I = 0$	0.50	0.10	0.05	0.30

Table:  $E[\pi | M, R, I]$ .



# DS #2: Post-Visit Randomization (PVR) - cont'd

## Interventional Inference

$$\text{PVR} \rightarrow \operatorname{argmax}_{I \in \{0,1\}} E[\pi | do(I), M, V = 1]$$

$$E[\pi | do(I = 1), M = 1, V = 1] = 0.25$$

$$E[\pi | do(I = 0), M = 1, V = 1] = 0.50$$

$$E[\pi | do(I = 1), M = 0, V = 1] = 0.05$$

$$E[\pi | do(I = 0), M = 0, V = 1] = 0.30$$

	$R = 0$		$R = 1$	
	$M = 1$	$M = 0$	$M = 1$	$M = 0$
$I = 1$	0.25	0.50	0.45	0.05
$I = 0$	0.50	0.10	0.05	0.30

Table:  $E[\pi | M, R, I]$ .

## Decision Rule:

- If Visit Site  $\wedge$  Married  $\rightarrow$  No Purchase Incentive  $\rightarrow E[\pi] = \mathbf{0.50}$
- If Visit Site  $\wedge$  Not Married  $\rightarrow$  No Purchase Incentive  $\rightarrow E[\pi] = \mathbf{0.30}$

$$\text{Expected profit} = \boxed{0.40} = (0.50 + 0.30) / 2.$$

# DS #3: A/B Test on All New-to-RBC Clients

Interventional Inference

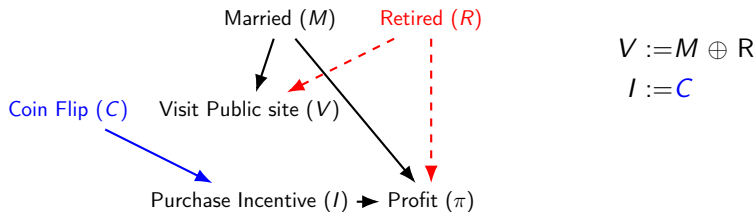


Figure: Causal DAG with A/B Test on all New-to-RBC clients.

# DS #3: A/B Test on All New-to-RBC Clients - cont'd

## Interventional Inference

$$\text{ABT} \rightarrow \operatorname{argmax}_{I \in \{0,1\}} E[\pi | \text{do}(I), M]$$

$$E[\pi | \text{do}(I = 1), M = 1] = 0.350 = (0.25 + 0.45)/2$$

$$E[\pi | \text{do}(I = 0), M = 1] = 0.275 = (0.50 + 0.05)/2$$

$$E[\pi | \text{do}(I = 1), M = 0] = 0.275 = (0.50 + 0.05)/2$$

$$E[\pi | \text{do}(I = 0), M = 0] = 0.200 = (0.10 + 0.30)/2$$

	$R = 0$		$R = 1$	
	$M = 1$	$M = 0$	$M = 1$	$M = 0$
$I = 1$	0.25	0.50	0.45	0.05
$I = 0$	0.50	0.10	0.05	0.30

Table:  $E[\pi | M, R, I]$ .

# DS #3: A/B Test on All New-to-RBC Clients - cont'd

## Interventional Inference

$$\text{ABT} \rightarrow \operatorname{argmax}_{I \in \{0,1\}} E[\pi | do(I), M]$$

$$E[\pi | do(I = 1), M = 1] = 0.350 = (0.25 + 0.45)/2$$

$$E[\pi | do(I = 0), M = 1] = 0.275 = (0.50 + 0.05)/2$$

$$E[\pi | do(I = 1), M = 0] = 0.275 = (0.50 + 0.05)/2$$

$$E[\pi | do(I = 0), M = 0] = 0.200 = (0.10 + 0.30)/2$$

	$R = 0$		$R = 1$	
	$M = 1$	$M = 0$	$M = 1$	$M = 0$
$I = 1$	0.25	0.50	0.45	0.05
$I = 0$	0.50	0.10	0.05	0.30

Table:  $E[\pi | M, R, I]$ .

## Decision Rule:

- If Married  $\rightarrow$  Purchase Incentive  $\rightarrow E[\pi] = 0.35$
- If Not Married  $\rightarrow$  Purchase Incentive  $\rightarrow E[\pi] = 0.275$

$$\text{Expected profit} = \boxed{0.315} = (0.35 + 0.275)/2.$$

# DS #4: Regret Decision Criterion (RDC)

Counterfactual Inference

$$\text{RDC} \rightarrow \operatorname{argmax}_{a' \in 0,1} E[\pi_{a'} | I = a, M]$$

# DS #4: Regret Decision Criterion (RDC)

## Counterfactual Inference

$$\text{RDC} \rightarrow \arg\max_{a' \in 0,1} E[\pi_{a'} | I = a, M]$$

$$\begin{aligned} P(\pi_{a'}, M) &= P(\pi_{a'}, M, a') + P(\pi_{a'}, M, a) \\ &= P(\pi_{a'} | M, a') P(M, a') + P(\pi_{a'} | M, a) P(M, a) \end{aligned}$$

$$\begin{aligned} P(\pi_{a'} | M) &= P(\pi_{a'} | M, a') P(a' | M) + P(\pi_{a'} | M, a) P(a | M) \\ &= P(\pi | M, a') P(a' | M) + P(\pi_{a'} | M, a) P(a | M) \quad (\text{Consistency}) \end{aligned}$$

$$P(\pi_{a'} | M, a) = \frac{1}{P(a | M)} \left[ P(\pi_{a'} | M) - P(\pi | M, a') P(a' | M) \right]$$

$$= \underbrace{\frac{1}{P(a | M)}}_{\text{observational}} \left[ \overbrace{P(\pi | M, do(a'))}^{\text{experimental}} - \underbrace{P(\pi | M, a') P(a' | M)}_{\text{observational}} \right]$$

# DS #4: Regret Decision Criterion (RDC) - cont'd

## Counterfactual Inference

$$E(\pi_{I=1}|M=1, I=0) =$$

$$\begin{aligned} & \frac{1}{P(I=0|M=1)} \left[ E(\pi|M=1, do(I=1)) - \right. \\ & \quad \left. E(\pi|M=1, I=1)P(I=1|M=1) \right] \\ &= \frac{1}{1/2} (0.350 - 0.25 \times 1/2) = 0.45 \\ &> 0.05 = E(\pi_{I=0}|M=1, I=0). \end{aligned}$$

$$E(\pi_{I=1}|M=0, I=0) = 0.50$$

$$E(\pi_{I=0}|M=1, I=1) = 0.50$$

$$E(\pi_{I=0}|M=0, I=1) = 0.30$$

	R = 0		R = 1	
	M = 1	M = 0	M = 1	M = 0
I = 1	0.25	0.50	0.45	0.05
I = 0	0.50	0.10	0.05	0.30

Table:  $E[\pi|M, R, I]$ .

# DS #4: Regret Decision Criterion (RDC) - cont'd

## Counterfactual Inference

$$E(\pi_{I=1}|M=1, I=0) =$$

$$\begin{aligned} & \frac{1}{P(I=0|M=1)} \left[ E(\pi|M=1, do(I=1)) - \right. \\ & \quad \left. E(\pi|M=1, I=1)P(I=1|M=1) \right] \\ &= \frac{1}{1/2} (0.350 - 0.25 \times 1/2) = 0.45 \\ &> 0.05 = E(\pi_{I=0}|M=1, I=0). \end{aligned}$$

	R = 0		R = 1	
	M = 1	M = 0	M = 1	M = 0
I = 1	0.25	0.50	0.45	0.05
I = 0	0.50	0.10	0.05	0.30

Table:  $E[\pi|M, R, I]$ .

$$E(\pi_{I=1}|M=0, I=0) = 0.50$$

$$E(\pi_{I=0}|M=1, I=1) = 0.50$$

$$E(\pi_{I=0}|M=0, I=1) = 0.30$$

### Decision Rule:

- If Not Visit Site  $\wedge$  Married  $\rightarrow$  Purchase Incentive  $\rightarrow E[\pi] = 0.45$
- If Not Visit Site  $\wedge$  Not Married  $\rightarrow$  Purchase Incentive  $\rightarrow E[\pi] = 0.50$
- If Visit Site  $\wedge$  Married  $\rightarrow$  No Purchase Incentive  $\rightarrow E[\pi] = 0.50$
- If Visit Site  $\wedge$  Not Married  $\rightarrow$  No Purchase Incentive  $\rightarrow E[\pi] = 0.30$

$$\text{Expected profit} = 0.4375 = (0.45 + 0.50 + 0.50 + 0.30)/4.$$



# Summary of Methods

Criterion	Decision Rule	$E[\pi]$
EDC	<ul style="list-style-type: none"><li>• If Visit Site <math>\wedge</math> Married <math>\rightarrow</math> <b>Purchase Incentive</b></li><li>• If Visit Site <math>\wedge</math> Not Married <math>\rightarrow</math> <b>No Purchase Incentive</b></li></ul>	.2750
PVR	<b>Never Purchase Incentive</b>	.4000
ABT	<b>Always Purchase Incentive</b>	.3150
RDC	<ul style="list-style-type: none"><li>• If Visit Site <math>\wedge</math> Married <math>\rightarrow</math> <b>No Purchase Incentive</b></li><li>• If Visit Site <math>\wedge</math> Not Married <math>\rightarrow</math> <b>No Purchase Incentive</b></li><li>• If Not Visit Site <math>\wedge</math> Married <math>\rightarrow</math> <b>Purchase Incentive</b></li><li>• If Not Visit Site <math>\wedge</math> Not Married <math>\rightarrow</math> <b>Purchase Incentive</b></li></ul>	.4375
Oracle		.4375

# Remarks

- If the goal is to learn personalized actions, experimental data alone is sub-optimal in the presence of unobserved confounders.
- Combining experimental and observational data under a Regret Decision Criterion (RDC) can provide information about the unobserved confounders, and hence outperform alternative optimization criteria.
- The expression derived from RDC works only in the binary treatment case. RDC-type randomization (Forney et al., 2017) was proposed to estimate counterfactual expressions empirically from an arbitrary number of treatments.

# References

- Elias Bareinboim, Andrew Forney, and Judea Pearl. 2015. Bandits with unobserved confounders: a causal approach. In Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1 (NIPS'15).
  - ▶ Implementation: <https://github.com/leoguelman/mabuc>
- Forney, A., Pearl, J. ; Bareinboim, E.. (2017). Counterfactual Data-Fusion for Online Reinforcement Learners. Proceedings of the 34th International Conference on Machine Learning, in Proceedings of Machine Learning Research 70:1156-1164