I reject your RCT and substitute my own

A workshop on inverse propensity weighting

Damian Pavlyshyn[1]

[1]Disease Elimination Program, Burnet Institute, Melbourne

7 July, 2023

Burnet

Acknowledgement of country

I recognise the Boon Wurrung people of the Kulin Nations as the traditional owners of the land on which we meet today, and acknowledge that they never ceded sovereignty over it.

I pay my respect to elders past and present, and extend that respect to all first nations people.

Burnet

How to perform an IPW analysis

Make Your Own Eggbox Jellyfish

Ask an adult to help you with scissors and cutting out

Burnet

How to perform an IPW analysis

Make Your Own

Causal analysis

Ask a statistician to help you with

inverse propensity weighting

Burnet

Case study: An observational drug trial

A disease, indicated by low values of a biomarker has an existing, well-established treatment B.
An experimental treatment A has been given to some patients.
Question: Is A more effective than B at increasing this biomarker?
Target trial:

1. Assign each participant to a treatment and control group at random, giving the control group treatment B and the treatment group treatment A.
2. For each participant i, take a baseline and post-treatment measurement of the biomarker. The outcome is the difference between these measurements, which we call Y_i.
3. Compare the average of the outcomes Y_i in the treatment group with the Y_i in the control group.

Burnet

The data

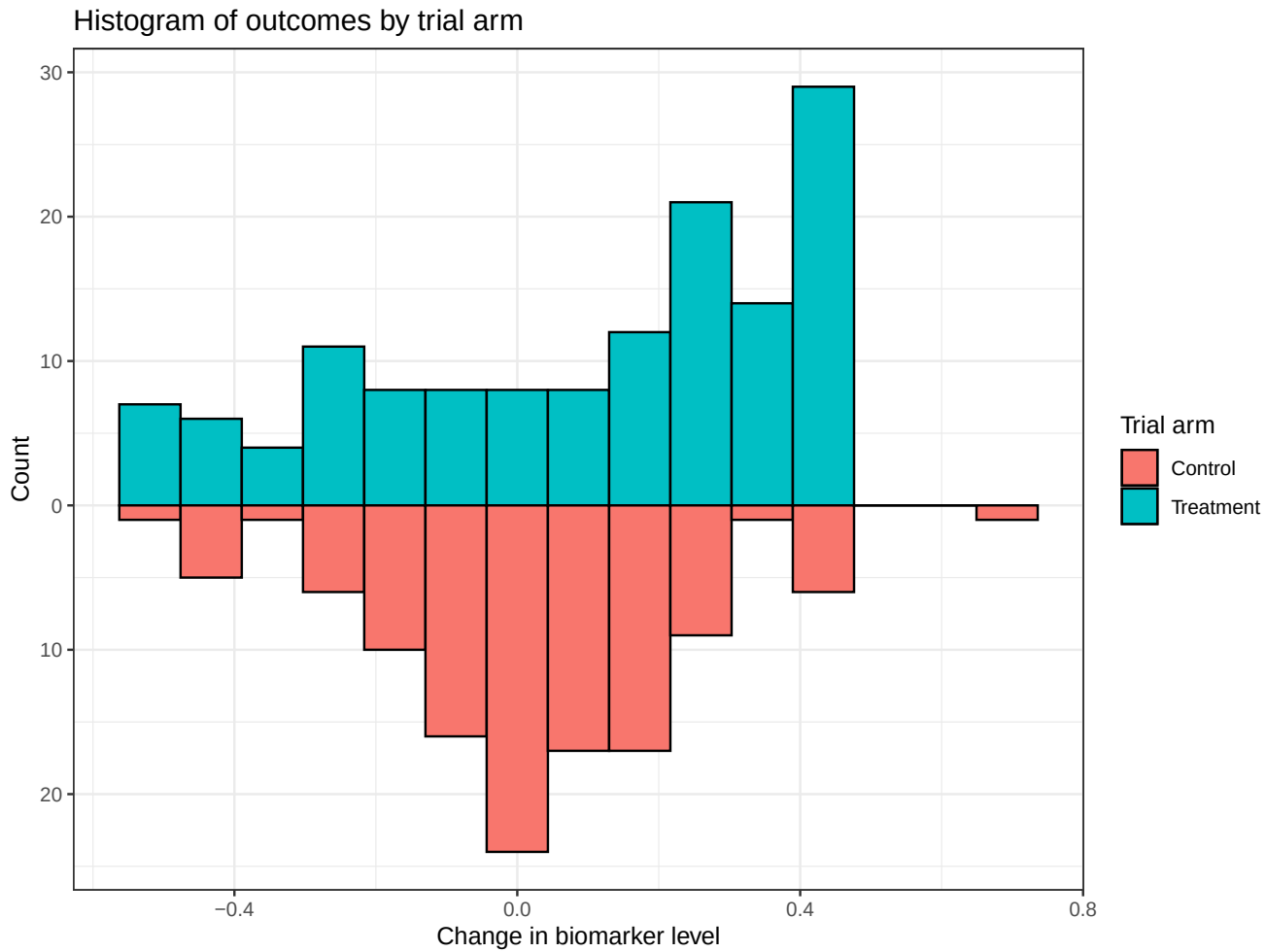
First few rows of the measurements:

PID	Arm	Baseline	Follow-up	Outcome
1	Treatment	1.42	1.64	0.21
2	Control	1.42	1.66	0.24
3	Treatment	1.86	1.76	-0.10
4	Treatment	2.84	2.57	-0.26
5	Control	0.49	0.96	0.46
6	Treatment	1.73	1.21	-0.52
⋮	⋮	⋮	⋮	⋮

Summary table of the outcomes by trial arm:

Arm	Count	Average outcome
Control	112	0.02
Treatment	136	0.10

Burnet



Estimated ATE	Standard error	95% CI
0.073	0.032	[0.0095, 0.14]

Burnet

Wait ... what are we even estimating?

The potential outcomes model: Each individual i, has a "treatment outcome" and a "control outcome". The treatment effect TE_i is the difference between outcomes for this individual.

The average treatment effect ATE is the average of these treatment effects:

$$\text{ATE} = \frac{1}{n}\sum_{i=1}^{n} \text{TE}_i.$$

Interpretation: "If all participants were assigned to the treatment arm, how much would the outcome change compared to if all participants were assigned to the control arm?"

Burnet

Ideal data

We can imagine a table with both potential outcomes for each individual

PID	Arm	Control outcome	Treatment outcome	Treatment effect
1	Treatment	-0.15	0.21	0.36
2	Control	0.24	0.26	0.01
3	Treatment	-0.38	-0.10	0.28
4	Control	0.16	-0.26	-0.43
5	Control	0.46	0.22	-0.24
6	Treatment	-0.09	-0.52	-0.44
⋮	⋮	⋮	⋮	⋮

Recall our summary table from a few slides ago:

True ATE	Estimated ATE	Standard error	95% CI
-0.025	0.073	0.032	[0.0095, 0.14]

Burnet

Observed data

But we can never see all of it

PID	Arm	Control outcome	Treatment outcome	Treatment effect
1	Treatment	?	0.21	?
2	Control	0.24	?	?
3	Treatment	?	-0.10	?
4	Control	0.16	?	?
5	Control	0.46	?	?
6	Treatment	?	-0.52	?
⋮	⋮	⋮	⋮	⋮

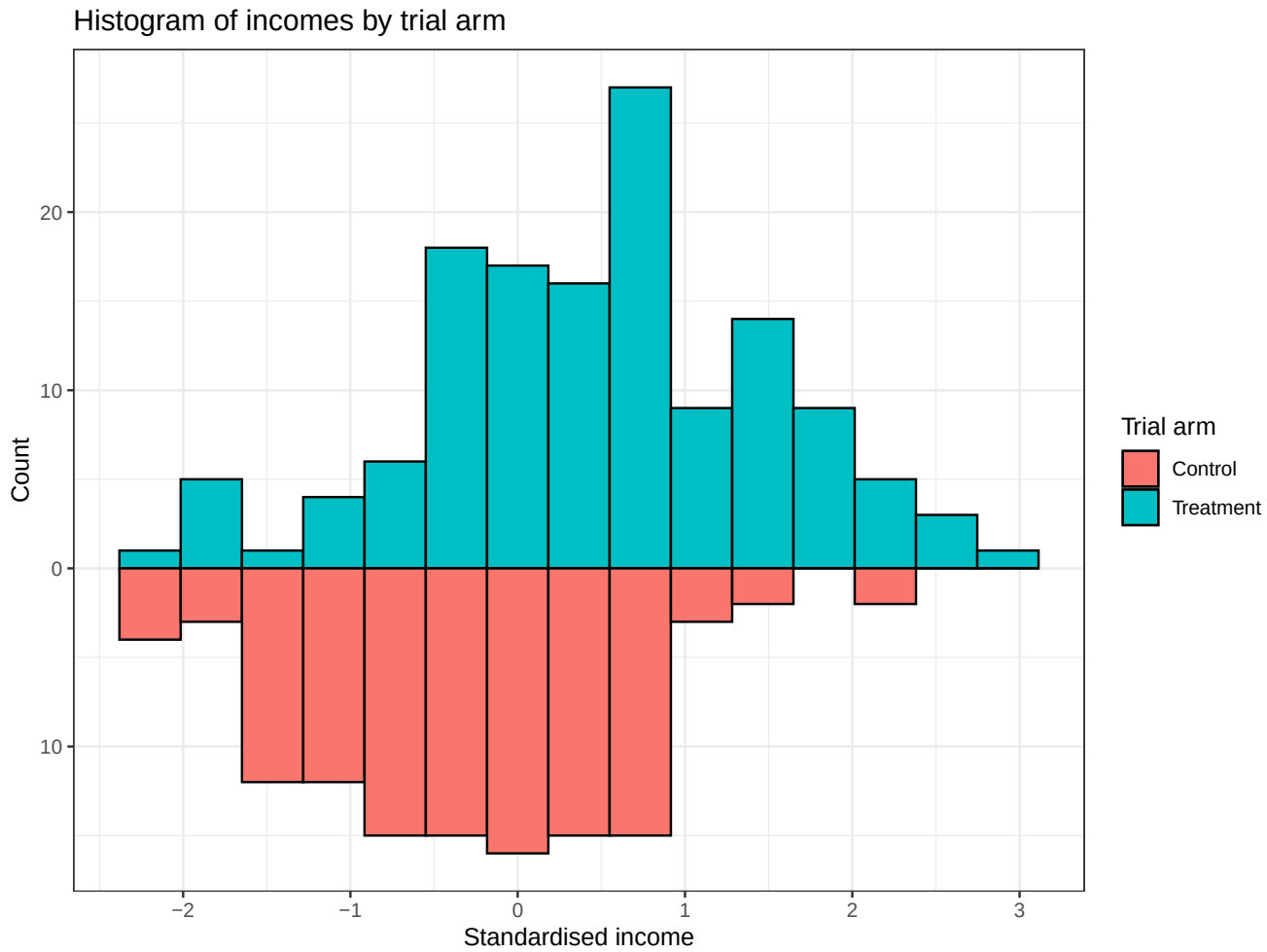
Recall our summary table from a few slides ago:

True ATE	Estimated ATE	Standard error	95% CI
-0.025	0.073	0.032	[0.0095, 0.14]

Burnet

We did an RCT analysis on something that's not an RCT

Treatment assignment is confounded by income



In an RCT these distributions would be balanced.

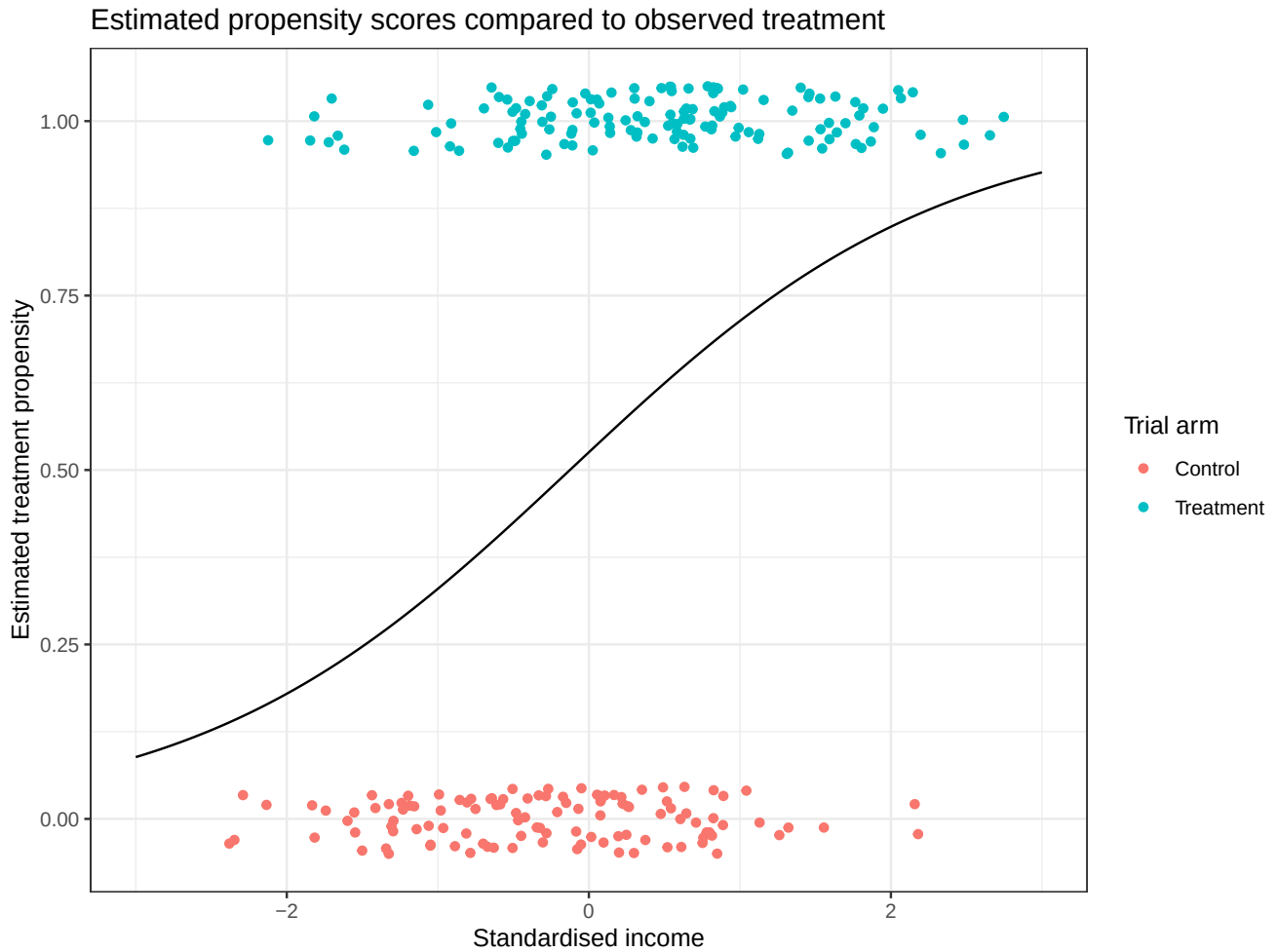
Burnet

Treatment propensity

The probability $e(X_i)$ that an individual with covariates X_i is assigned to treatment.

In an RCT: $e(X_i) = 0.5$

In our case: $e(X_i)$ varies with income.



Burnet

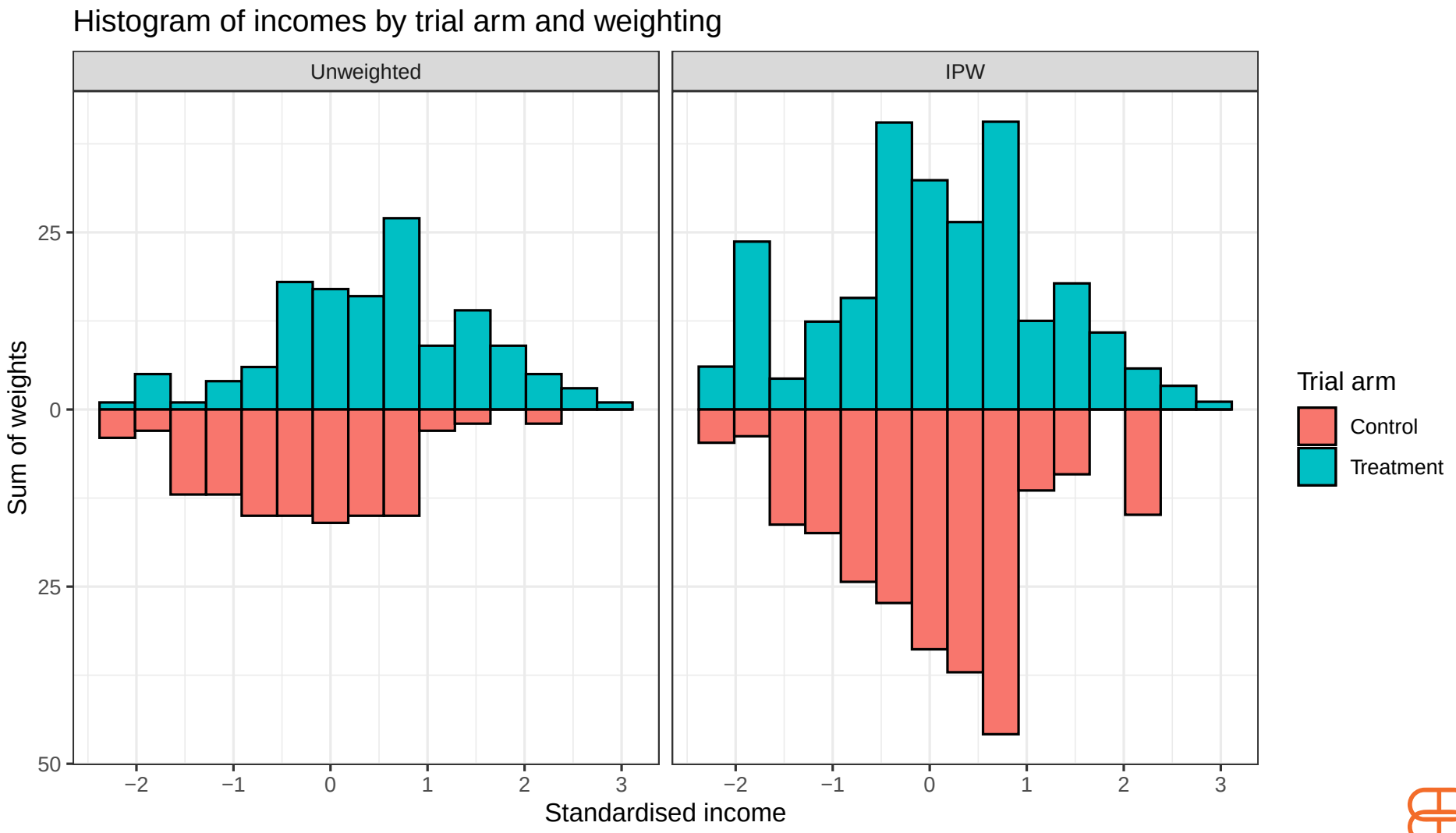
Inverse propensity weights

Idea: Balance the incomes of the treatment and control arms by emphasising and de-emphasising certain points

$$\text{IPW}(X_i) = \begin{cases} \frac{1}{\hat{e}(X_i)} & \text{if } i \text{ in treatment arm} \\ \frac{1}{1-\hat{e}(X_i)} & \text{if } i \text{ in control arm} \end{cases}$$

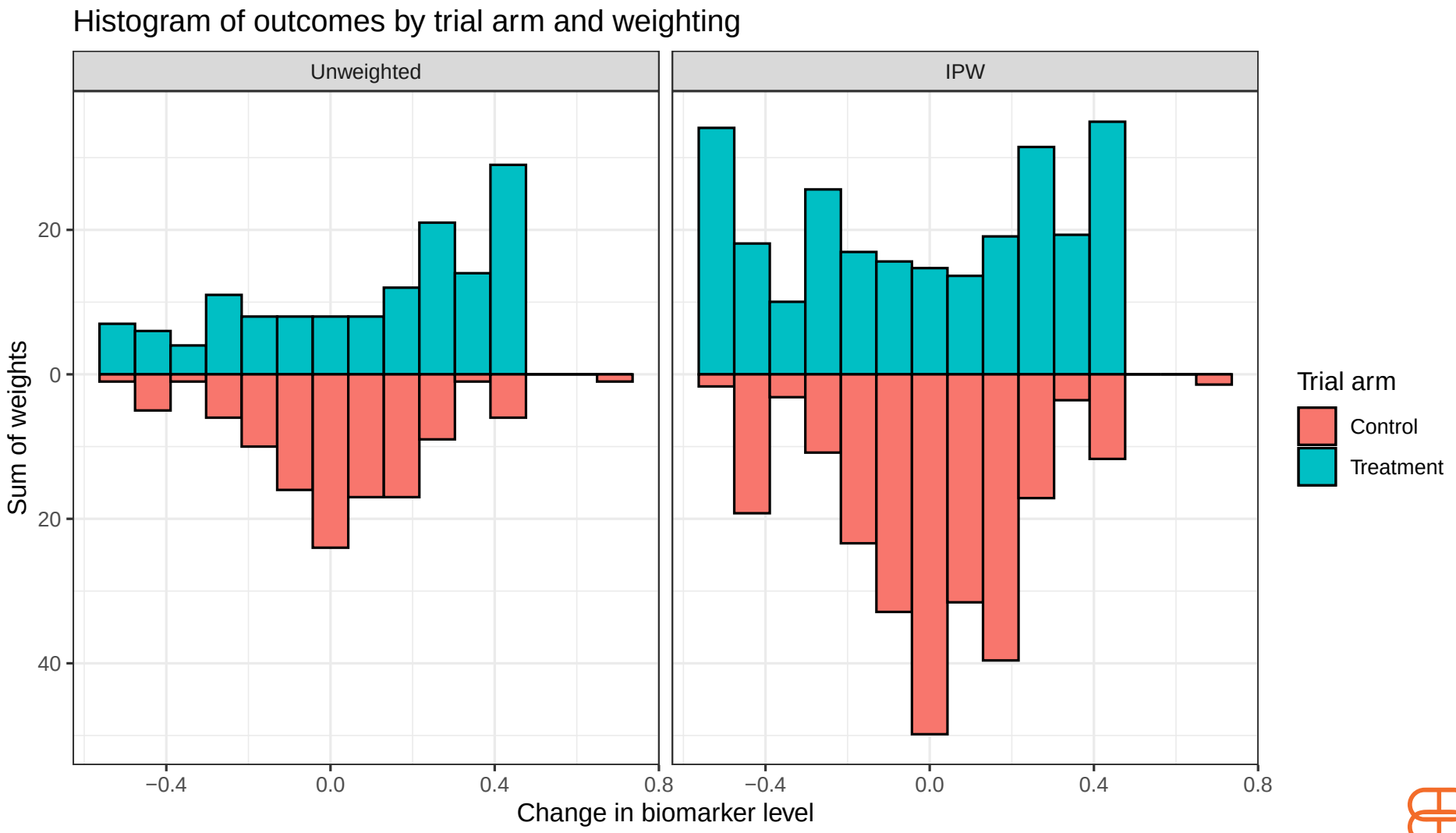
Burnet

Inverse propensity weighted data



Burnet

Inverse propensity weighted data



Burnet

What does this mean for our dataset?

Recall that the true ATE was -0.025.
Let's see how our IPW estimate does compared to the naive RCT estimate:

Estimate type	Estimated ATE	95% CI
RCT	0.073	[0.0095, 0.14]
IPW	-0.028	[-0.099, 0.049]

Burnet

Some warnings

Why many statisticians call all causal inference bullshit

This analysis relies heavily on the following assumptions:

1. we estimated our propensity scores accurately; and
2. we included all possible confounding variables.

This is *fundamentally impossible* to verify!!!

The price of accounting for causality:

1. confidence intervals will almost always be wider than for RCT estimates; and
2. there *must* be some randomness that's independent of the treatment assignment — the less extrinsic randomness, the wider the confidence intervals.

Burnet

What have we learned?

The average treatment effect: Measures the effect of a treatment over a population, even when different individuals respond differently to treatment.

The potential outcomes model: A model for causality that imagines the treatment *and* control outcomes for each individual, even though we only ever observe one of these.

Inverse propensity weighting: A method for modifying observational data to mimic experimental data. Works by balancing the covariate distributions of the treatment and control arms.

Burnet

Thank You

The wonderful world of causal inference

CATE, LATE, and treatment heterogeneity: Estimating different effects of a treatment on different parts of the population.

Matching methods: Create experimental data by finding similar individuals in treatment and control arms. The actually correct way to do causal inference.

Regression discontinuity designs: Evaluating all-of-population interventions and why we should distribute scholarship at random.

Instrumental variables: Dealing with unmeasured confounding and other black magic.

G-methods: Avoid at all costs.

Burnet

The IPW estimate of the average treatment effect

The RCT etimate:

$$\widehat{\text{ATE}} = \frac{1}{n_1} \sum_{i \text{ treated}} Y_i - \frac{1}{n_0} \sum_{i \text{ control}} Y_i$$
$$= \frac{1}{n} \Big(\sum_{i \text{ treated}} \frac{1}{n_1/n} Y_i - \sum_{i \text{ control}} \frac{1}{n_0/n} Y_i \Big)$$

The IPW estimate:

$$\widehat{\text{ATE}} = \frac{1}{n} \Big(\sum_{i \text{ treated}} \frac{1}{\hat{e}(X_i)} Y_i - \sum_{i \text{ control}} \frac{1}{1 - \hat{e}(X_i)} Y_i \Big)$$

Burnet