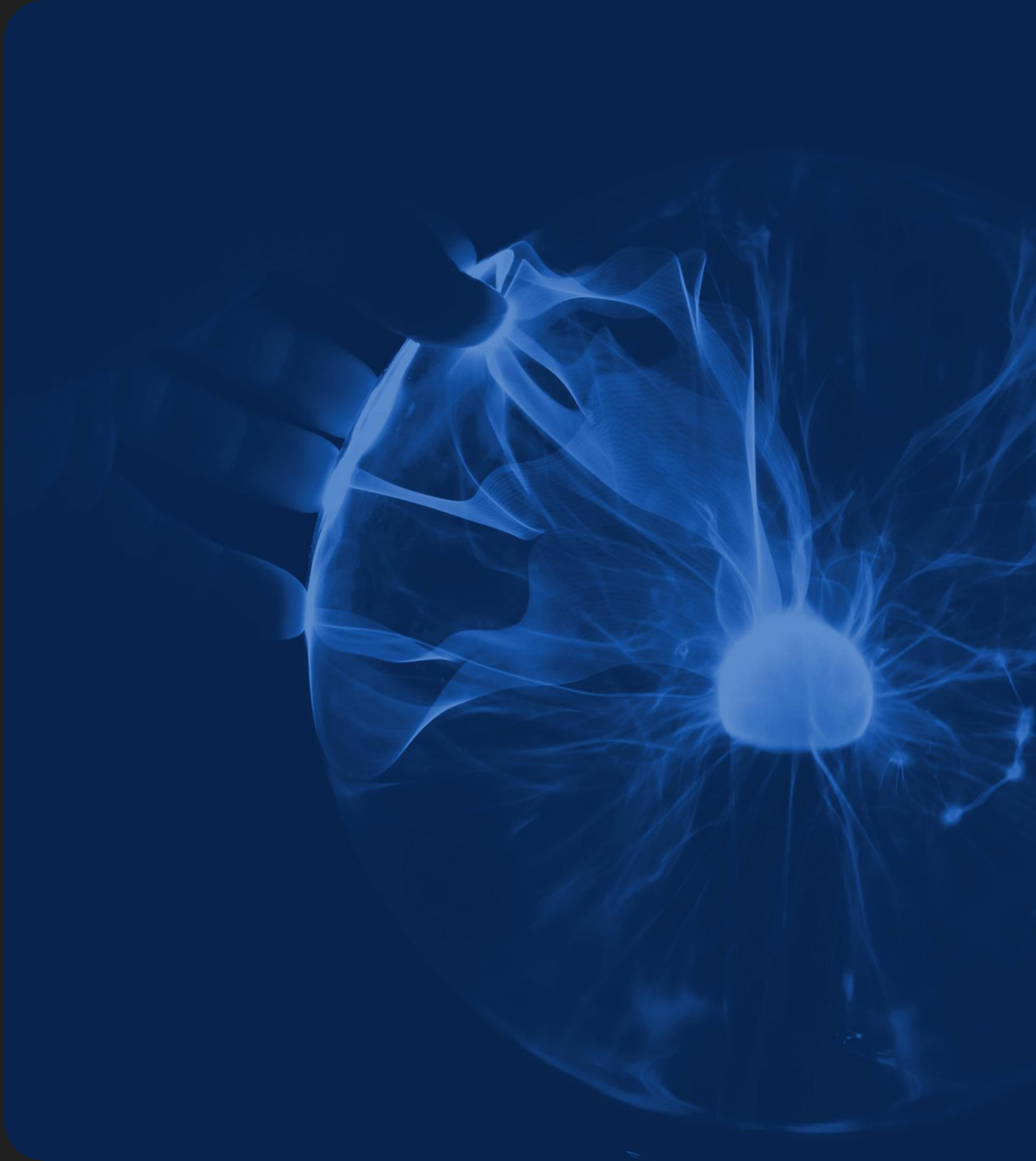
Bayesian Parametric G-Formula

As a Sensitivity Analysis



Causal Inference

Why it's Different

- Association implies bidirectionality
- Explicitly asking the question "Does X ***cause*** Y?"

Popular Methods

There are a lot of methods available for causal inference. Some popular ones include

Directed acyclic graphs (DAGs)

- Useful for variable selection
- Explicitly showing what relationships are assumed to be true

Propensity Score Methods

- Models the treatment
- Includes matching, weighting and stratification

Parametric G-Formula

- Models the outcome

Confounding

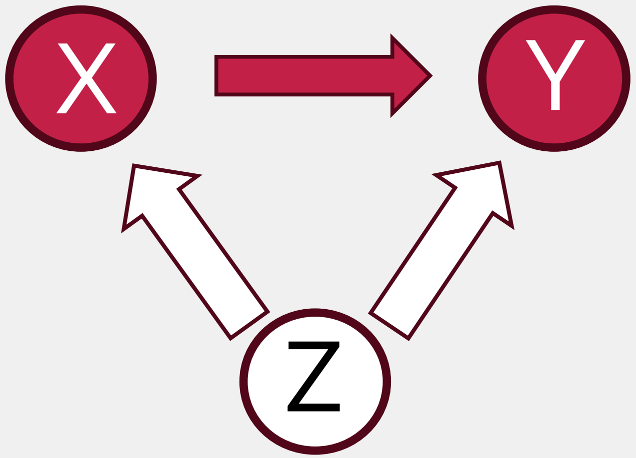
phastar

Confounding

- Confounding occurs when the variable is associated with both the outcome and the exposure

- This can cause a spurious relationship between two variables. We need to adjust for confounders to prevent this.

- There are two main types of confounding: measured and unmeasured



- In this directed acyclic graph (DAG), we have three variables: X, Y and Z.

- X is the treatment; Y is the outcome and Z is the confounder.

- Red arrow indicates what we're interested in estimating.

phastar

Measured vs Unmeasured Confounding

Measured Confounding

- Variable is available in dataset
- Able to adjust for measured confounding

Unmeasured Confounding

- Variable is not available in dataset
- Need to use sensitivity analyses to assess impact

Methods for Confounding

phastar

Popular Methods for Measured Confounding

Propensity Score-Based Methods

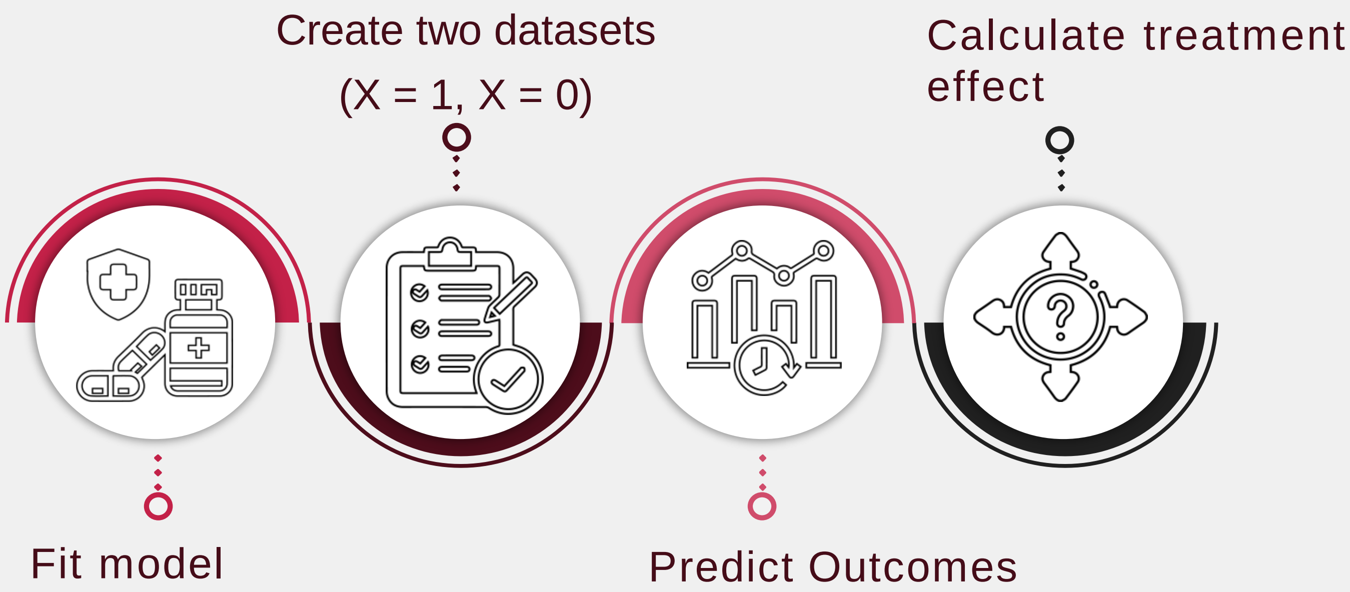
- Propensity score is probability of treatment given covariates
- Focus is on the treatment-covariate relationship

Parametric G-Formula

- Parametric g-formula uses a model for the outcome
- Focus is on the outcome-covariate relationship

Parametric G-Formula

How it works



The parametric g-formula is part of a broader class of methods called the g-methods. It is a plug-in estimator, so when these values are derived from a parametric method it's called the parametric g-formula.

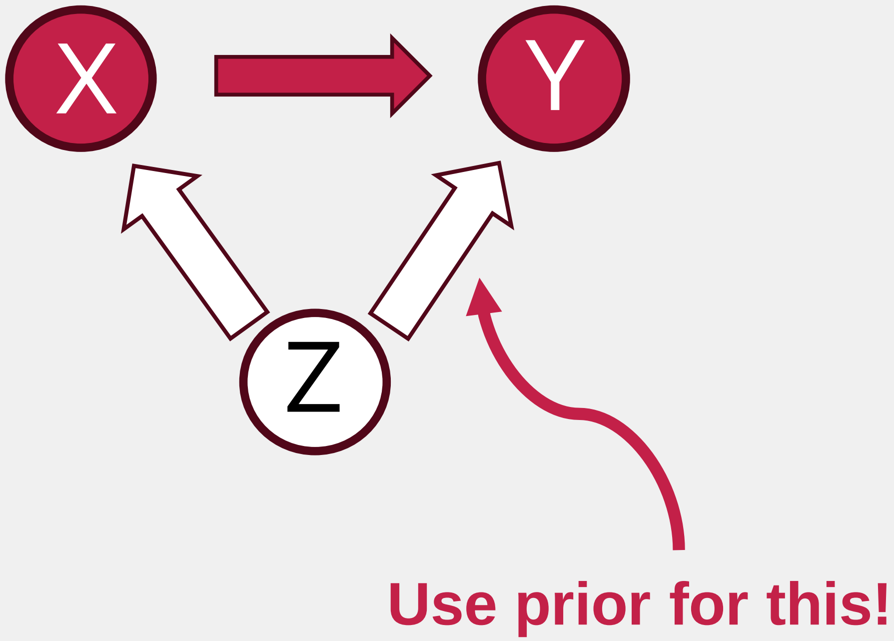
phastar

Bayesian Parametric G-Formula

phastar

So Why Bayesianly?

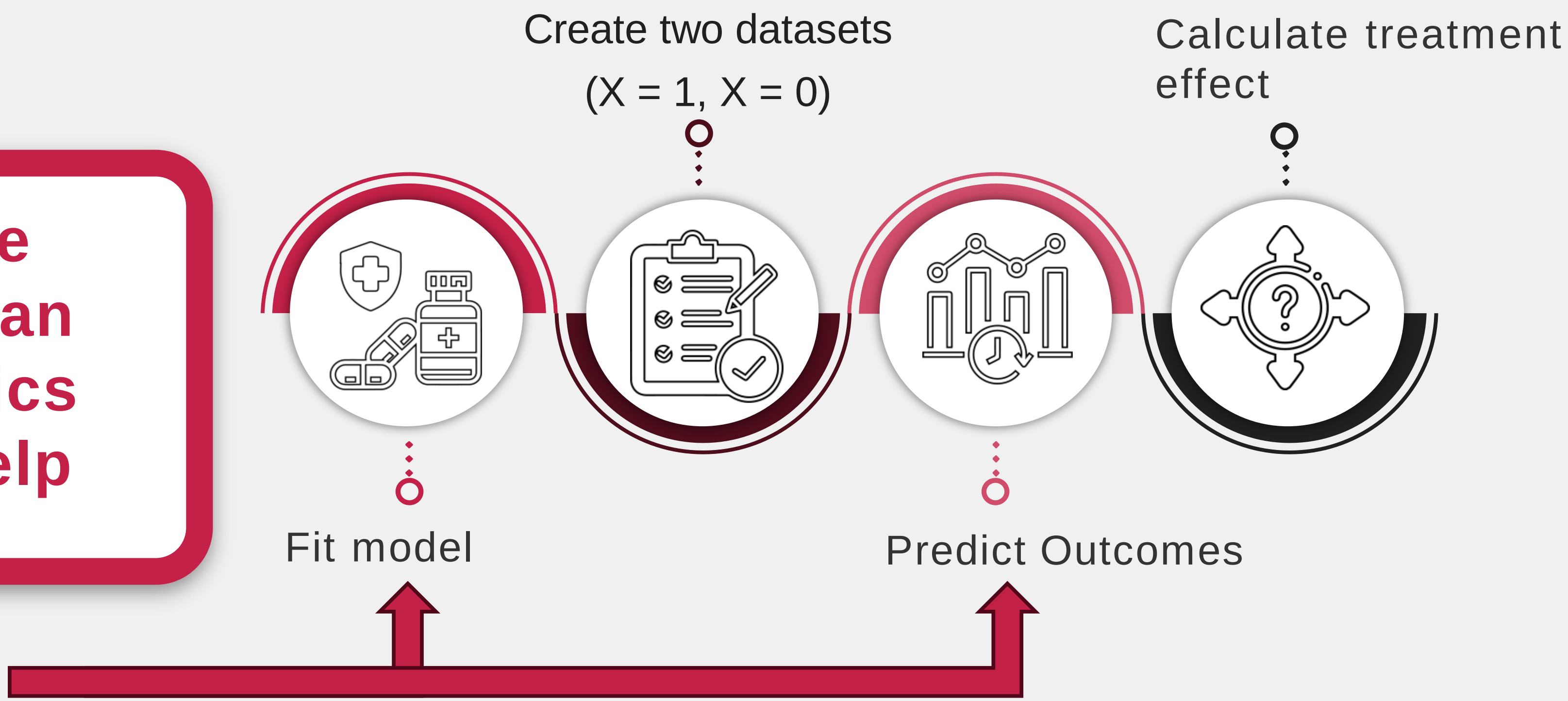
- Estimating the treatment effect with the parametric g-formula involves bootstrapping
- We can use the posterior distribution instead!
- Rather than using a single prior, we can use different priors for sensitivity analyses.



- We do know ***some*** prior information
- Use this prior information for the Z->Y relationship

Bayesian Parametric G-Formula

Where Bayesian Statistics Can Help

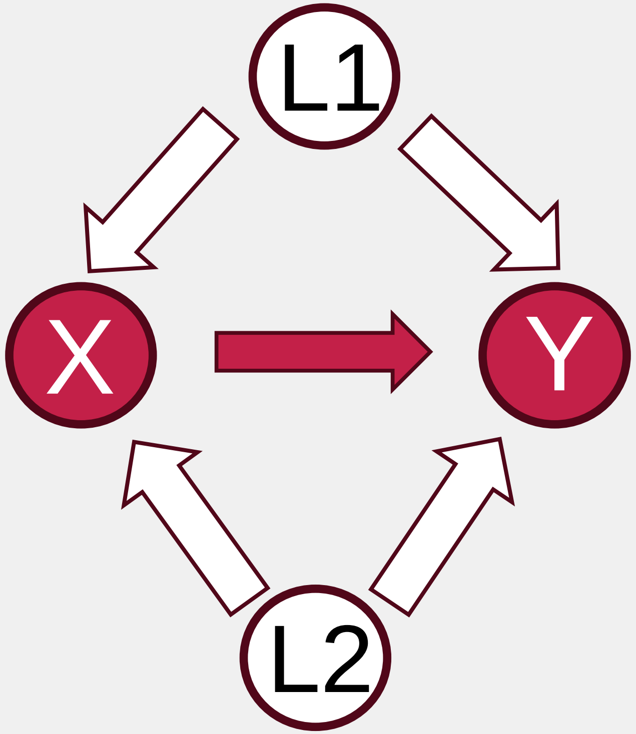


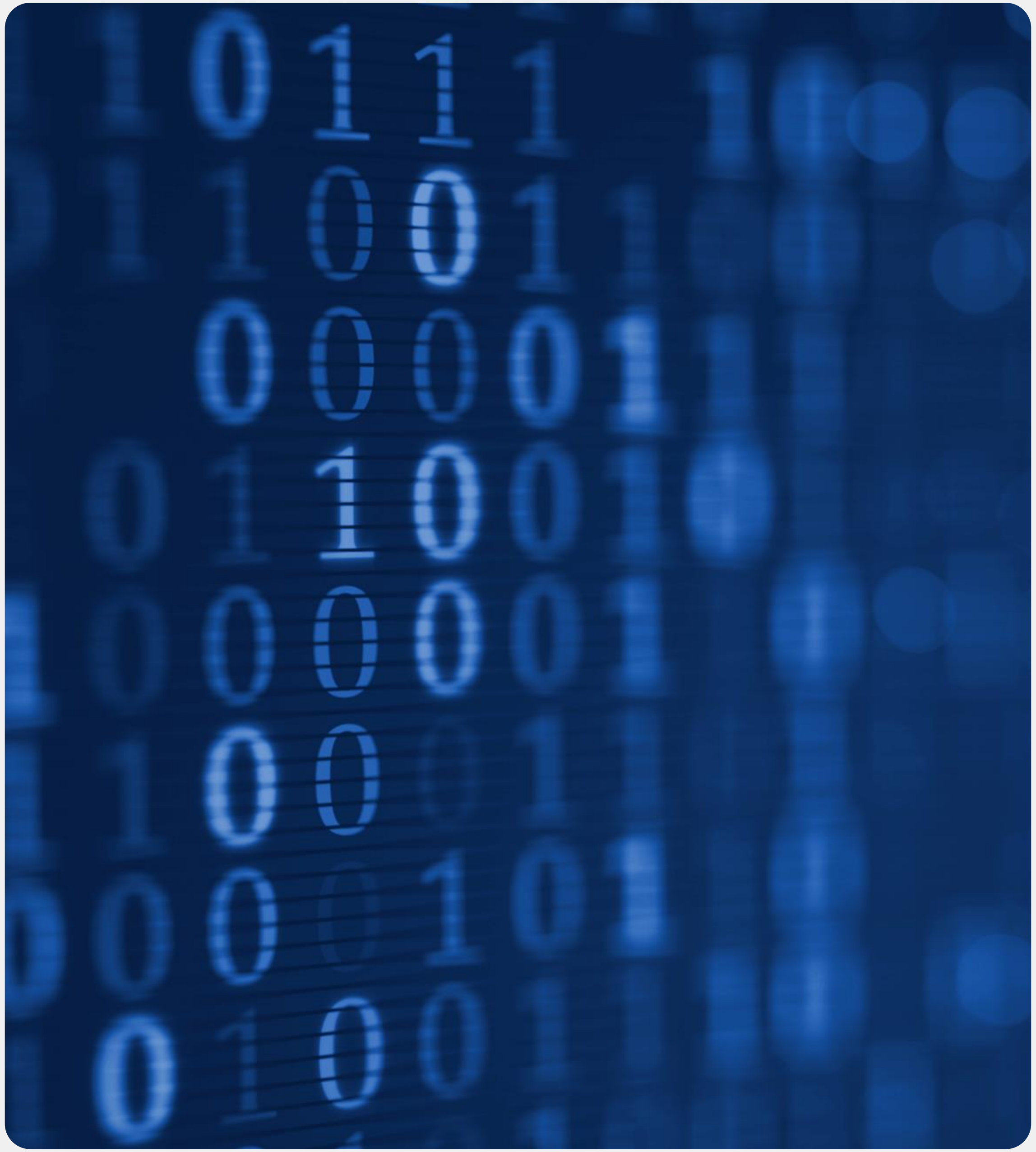
phastar

Example

Simulated Data

- Parameters
  - Sample size of 300
  - Risk difference of 0.15
  - Y = 0.15*X + 0.5 *L1 + 1.5*L2





Frequentist Results

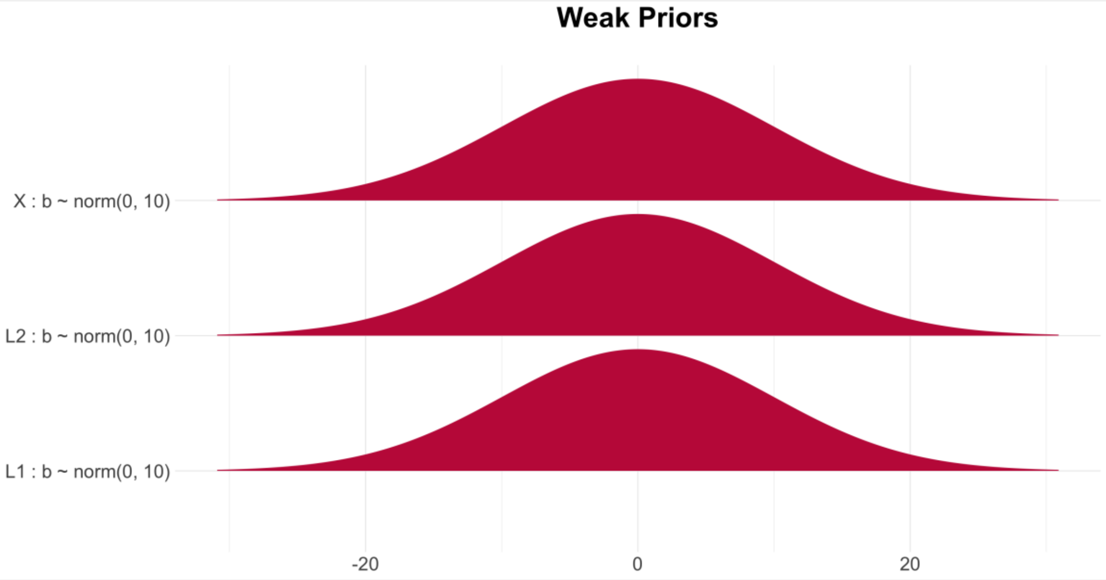
Method	Risk Difference*	95% Confidence Interval
IPTW*	**0.160**	**-0.149 to 0.468**
Parametric G-Formula		
Y | X = 1	3.552	3.381 to 3.723
Y | X = 0	3.268	3.048 to 3.488
Effect Estimate	**0.283**	**0.026 to 0.541**

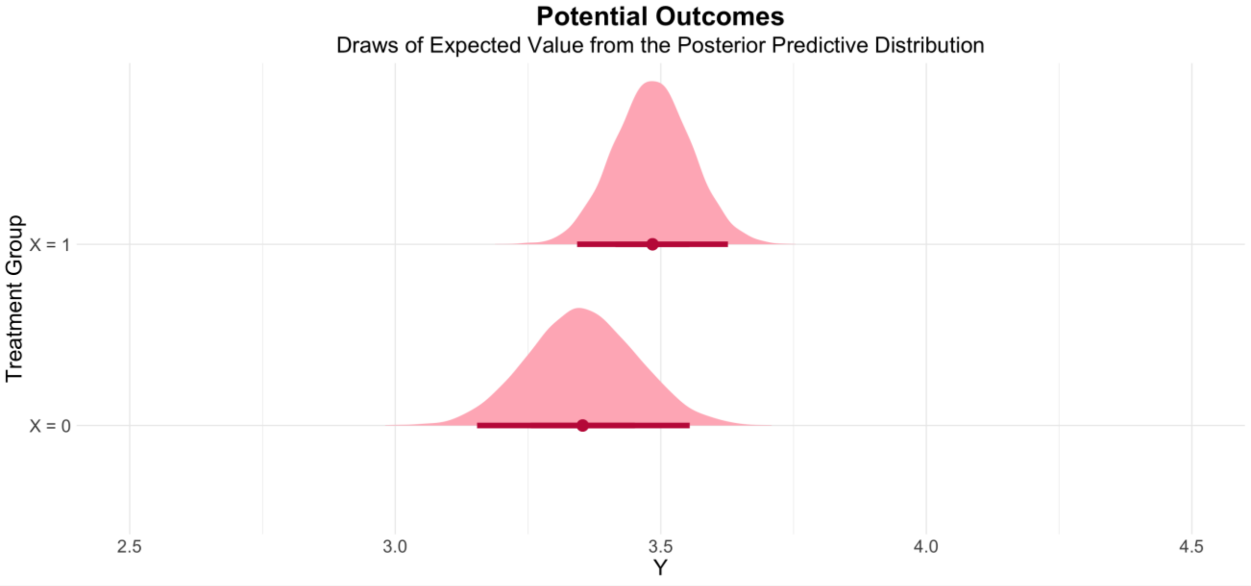
*Targeting the average treatment effect

- Both results give different answers. So which one is right?
- In this particular example, IPTW is closer to true effect but they model different things

phastar

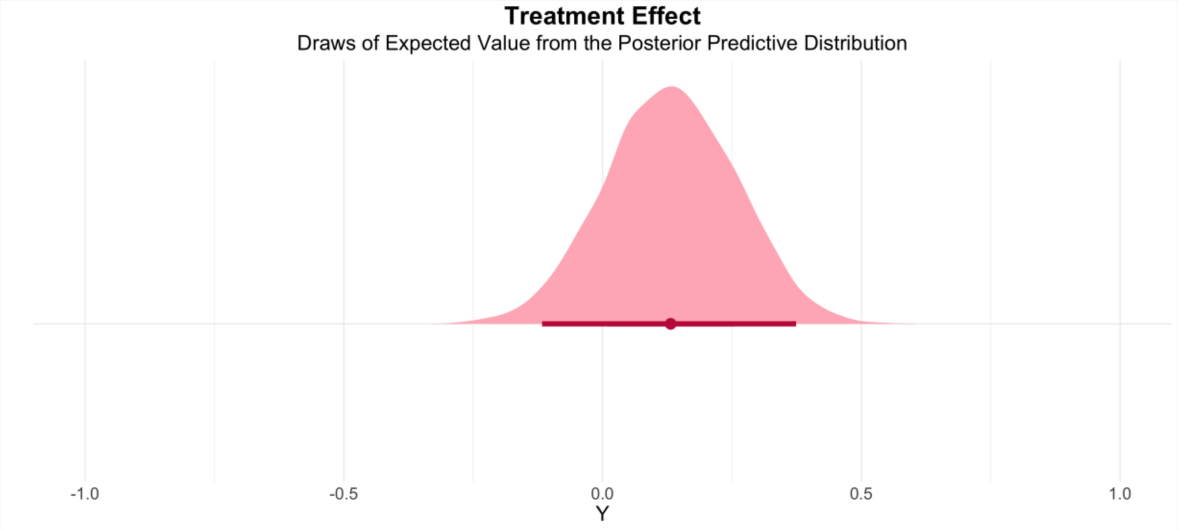
Bayesian Parametric G-Formula





- Treatment effect estimate:
  - 0.132 (-0.117 to 0.374)

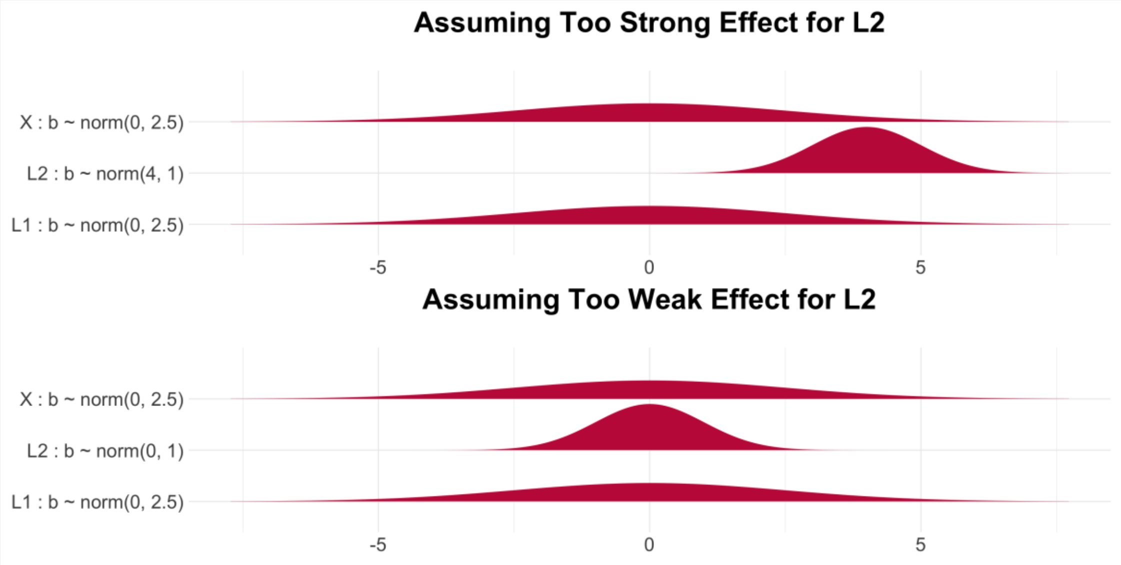
- Could we improve on this?



phastar

Scenarios - Priors

- We know ***some*** information about the confounder. After all, that's why it's included!
- What if the prior assumed too strong of a relationship?
- What if the prior assumed too weak of a relationship?



Scenarios - Results

Method	Risk Difference*	95% Credible Interval
Weak Priors	0.132	-0.117 to 0.374
Strong Confounder	0.124	-0.124 to 0.369
Weak Confounder	0.138	-0.111 to 0.383
True Effect	0.15	-

*Average treatment effect

- We need to be careful about how strong we assume the relationship is
  - Too strong – biased
  - Too weak – biased

- Like goldilocks, want the prior that is just right

phastar

Next Steps

Key Takeaways

- This method can incorporate prior information about confounder-outcome relationship
- We can assess robustness of results to different strengths of confounding
- A bonus is being able to visualize the posterior of the estimated treatment effect for each individual

Where From Here?

Next Steps

- Does this change with unmeasured confounding?
- Is there a tipping point for a strong confounder?
- Could we robustify our prior?
- Does this change with more confounders?
- Formal simulation study
- Model relationship between two confounders
- Different distributions other than the normal for the prior

Thank you!