Plausible Priors Precede Persuasive Posteriors

OPRE Innovative Methods Meeting:
Bayesian Methods for Social Policy Research and Evaluation

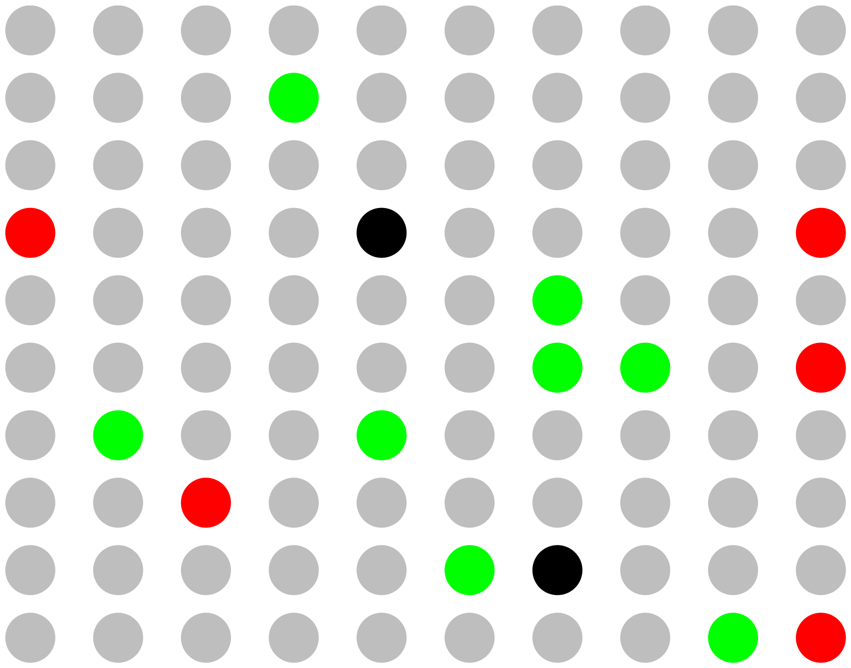
October 19, 2017

Mariel Finucane • Shira Mitchell • John Deke

We Need Plausible Priors – Example

- Federal grants fund 100 locally developed programs
- Truth (unknown to researcher or policy maker):
  - 10 programs have meaningful effects
  - 90 programs have no effect
- An RCT is conducted for 1 randomly selected program
- Statistical testing with alpha = 0.05
- Study power: the study is big enough that it has an 80% chance of detecting a meaningful effect

We Need Plausible Priors – Example: Pick a Study Out of a Barrel



100 studies COULD be conducted, but only 1 will be.

That 1 study will (in expectation) fall into 1 of 4 categories:

- **8 Green = significant and truly effective** $(power)$
- **2 Black = insignificant but effective** $(1 - power)$
- **85 Grey = insignificant and truly ineffective** $(1 - \alpha)$
- **5 Red = significant but ineffective** (α)

$$P(\text{sig impact is not real}) = \frac{5 \text{ red}}{5 \text{ red} + 8 \text{ green}} = 38\%$$

Incorporating Priors is Now Computationally Feasible

Stan…

… is new.
… is fast.
… is beginner friendly.
… makes it easy to extend models.
… has an active developer/user community.

Bayes Rule

Bayes Rule

<table>
<tr><td><u>Likelihood</u></td><td>+</td><td><u>Prior</u></td><td>=</td><td><u>Posterior</u></td></tr>
<tr><td>Take the estimate based on data from your study…</td><td></td><td>… and put it in the context of external information…</td><td></td><td>… to get a better answer to your policy question.</td></tr>
</table>

MATHEMATICA Policy Research

Bayes Rule

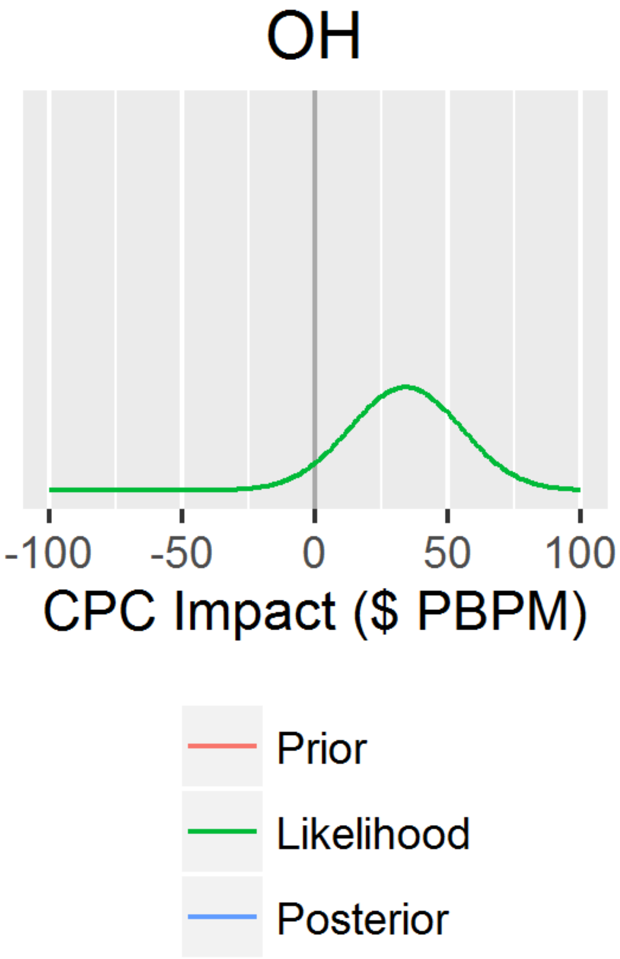
Likelihood + **Prior** = **Posterior**

Take the estimate based on data from your study…

… and put it in the context of external information…

… to get a better answer to your policy question.

Estimate = \$34
95% CI (-\$7, \$75)
p = 0.10



MATHEMATICA Policy Research

Bayes Rule

Likelihood + **Prior** = **Posterior**

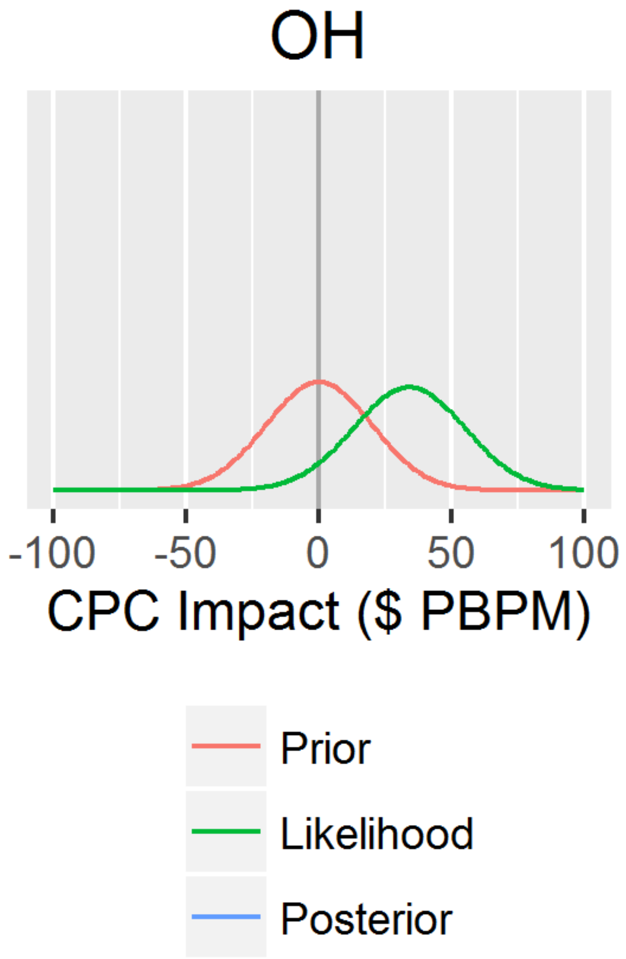
Take the estimate based on data from your study…

… and put it in the context of external information…

… to get a better answer to your policy question.

Estimate = $34
95% CI (-$7, $75)
p = 0.10

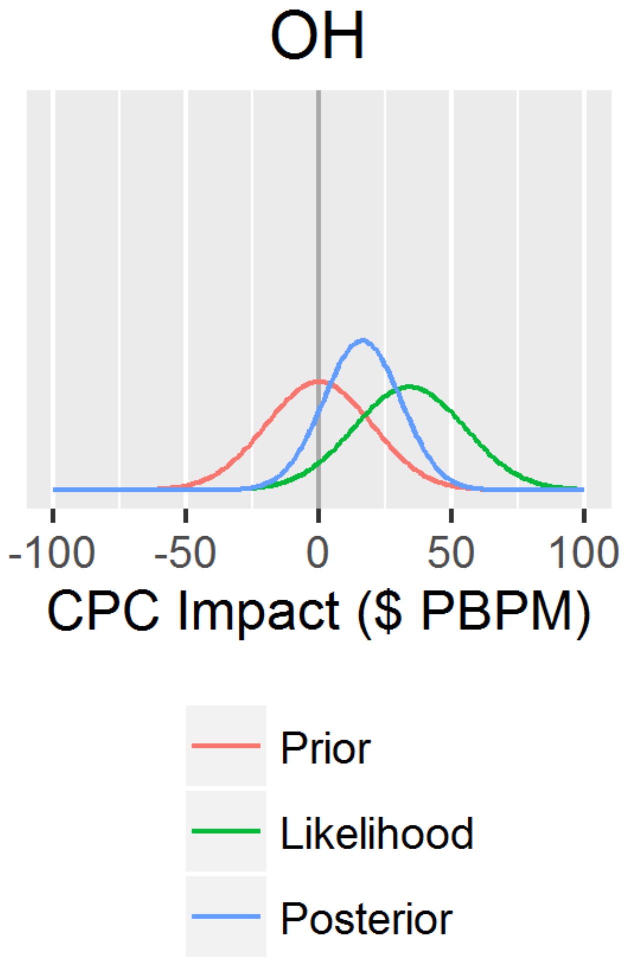
95% sure that
-$40 < impact < $40



MATHEMATICA Policy Research

Bayes Rule

Likelihood	+	Prior	=	Posterior
Take the estimate based on data from your study…		… and put it in the context of external information…		… to get a better answer to your policy question.
Estimate = $34 95% CI (-$7, $75) p = 0.10		95% sure that -$40 < impact < $40		Estimate = $16 95% CI (-$11, $43) 87% chance that Impact > $0



Bayes Rule

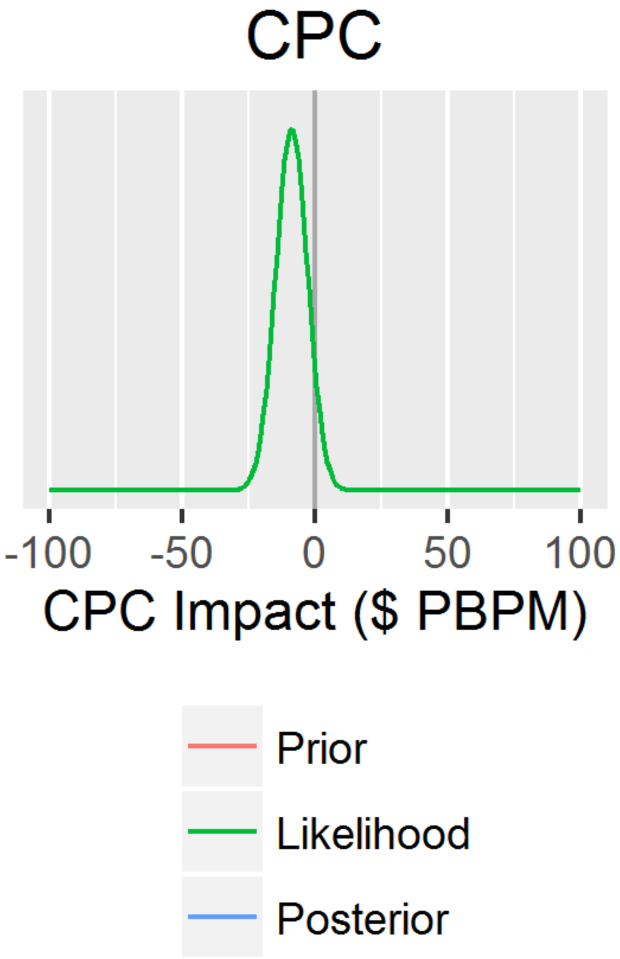
Likelihood + **Prior** = **Posterior**

Take the estimate based on data from your study…

… and put it in the context of external information…

… to get a better answer to your policy question.

Estimate = -$9
95% CI (-$21, $3)
p = 0.15



Bayes Rule

Likelihood + **Prior** = **Posterior**

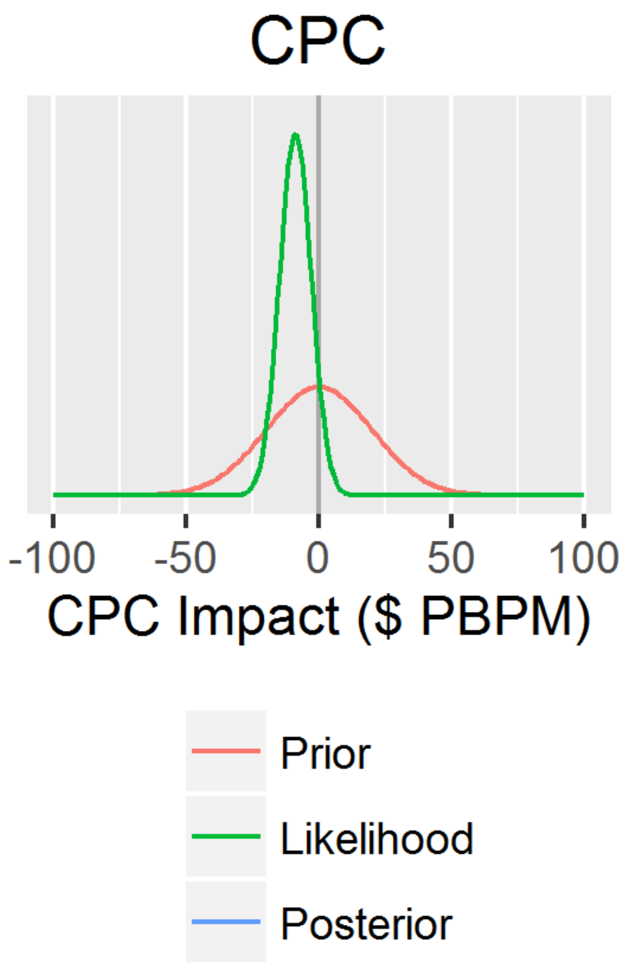
Take the estimate based on data from your study…

… and put it in the context of external information…

… to get a better answer to your policy question.

Estimate = -$9
95% CI (-$21, $3)
p = 0.15

95% sure that
-$40 < impact < $40



Bayes Rule

Likelihood + **Prior** = **Posterior**

Take the estimate based on data from your study…

… and put it in the context of external information…

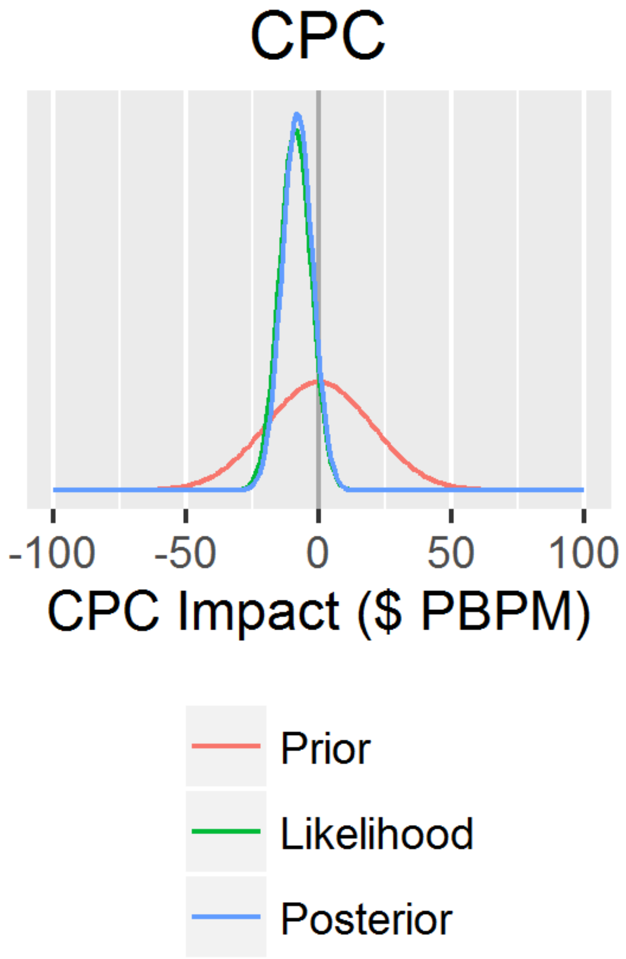
… to get a better answer to your policy question.

Estimate = -$9
95% CI (-$21, $3)
p = 0.15

95% sure that
-$40 < impact < $40

Estimate = -$8
95% CI (-$20, $4)

91% chance that impact < 0



A "Flat Prior" Fails the Sniff Test

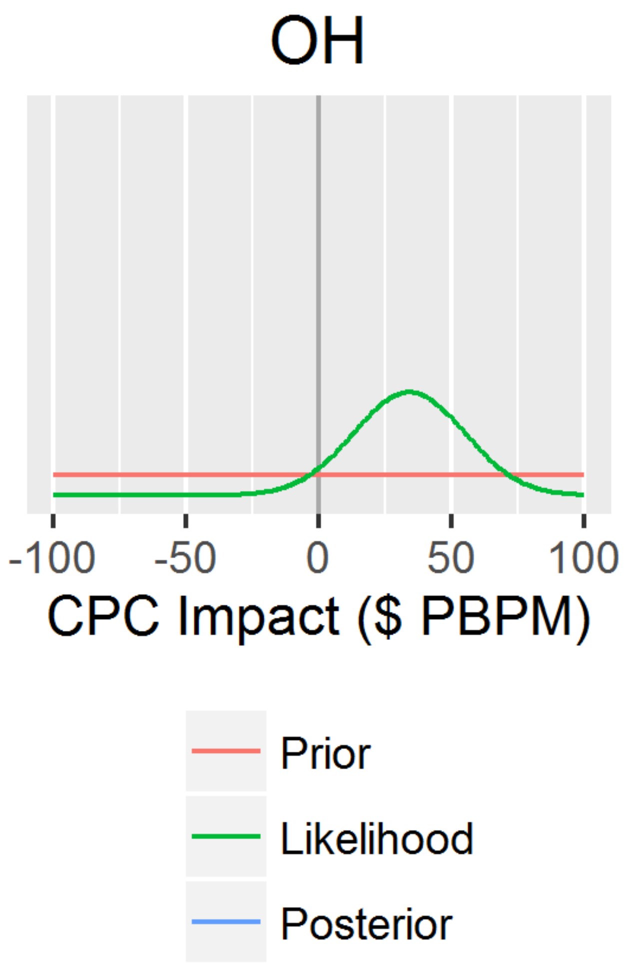
Likelihood + **Prior** = **Posterior**

Take the estimate based on data from your study…

… and put it in the context of external information…

… to get a better answer to your policy question.

Estimate = $34
95% CI (-$7, $75)
p = 0.10



A “Flat Prior” Fails the Sniff Test

Likelihood + **Prior** = **Posterior**

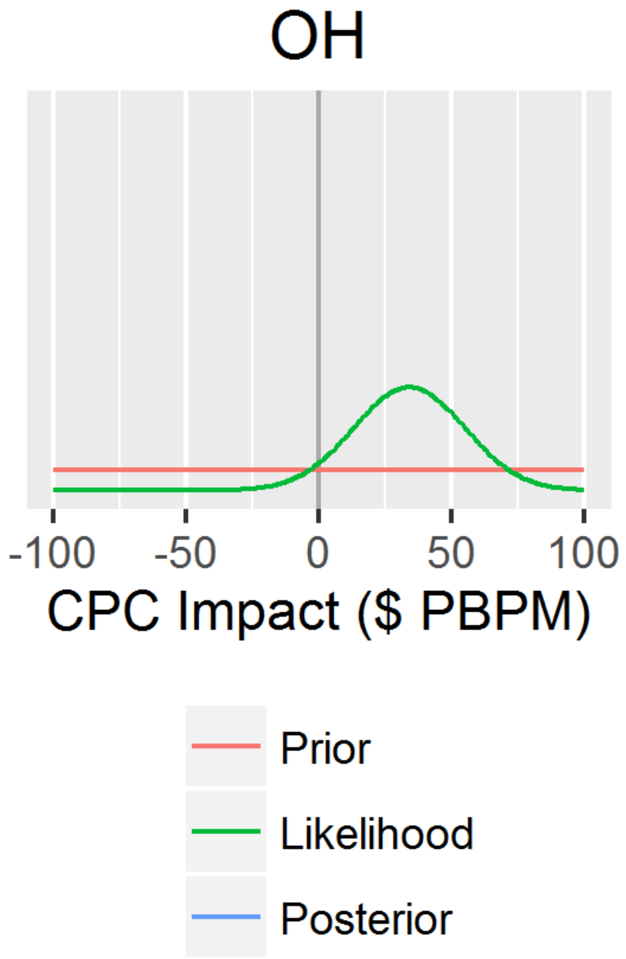
Take the estimate based on data from your study…

… and put it in the context of external information…

… to get a better answer to your policy question.

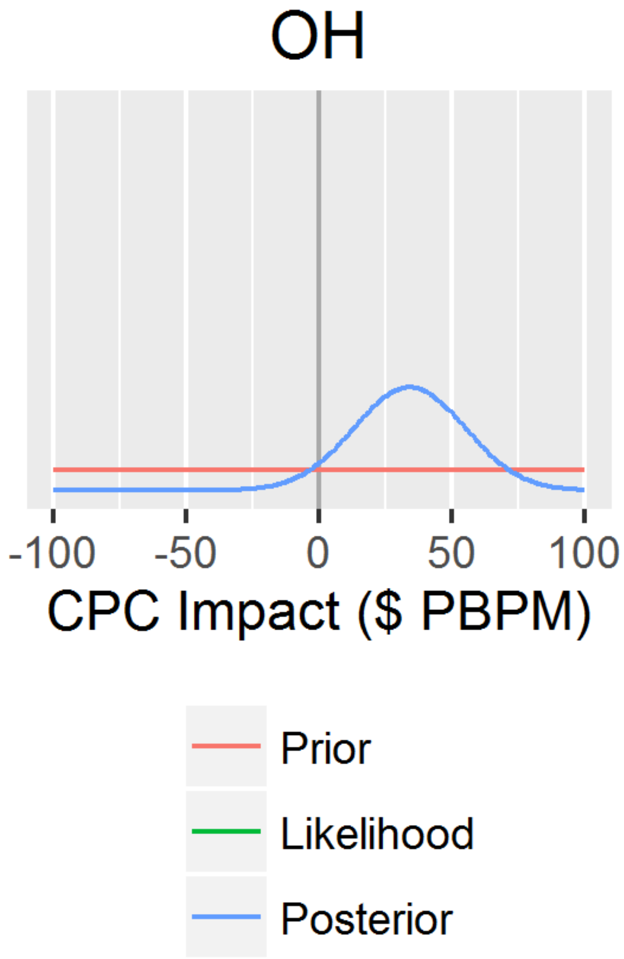
Estimate = $34
95% CI (-$7, $75)
p = 0.10

Probability that
-$50 < impact < $0
= probability that
-$1050 < impact < -$1000



A “Flat Prior” Fails the Sniff Test

Likelihood	+	Prior	=	Posterior
Take the estimate based on data from your study…		… and put it in the context of external information…		… to get a better answer to your policy question.
Estimate = \$34 95% CI (-\$7, \$75) p = 0.10 → 1-sided p=0.05		Probability that -\$50 < impact < \$0 = probability that -\$1050 < impact < -\$1000		Estimate = \$34 95% CI (-\$7, \$75) 5% chance that impact < \$0



MATHEMATICA Policy Research

The True Prior

The True Prior

Get the prior right →

- What group of questions is your research question a random draw from?
- The true prior is the distribution of true impacts in that group.

The True Prior

Get the prior right →

- What group of questions is your research question a random draw from?
- The true prior is the distribution of true impacts in that group.

IAH CPC

MCCD

etc…

The True Prior

Get the prior right →

- **What group of questions is your research question a random draw from?**
- **The true prior is the distribution of true impacts in that group.**

IAH
CPC
MCCD
etc…

→ Get the right posterior

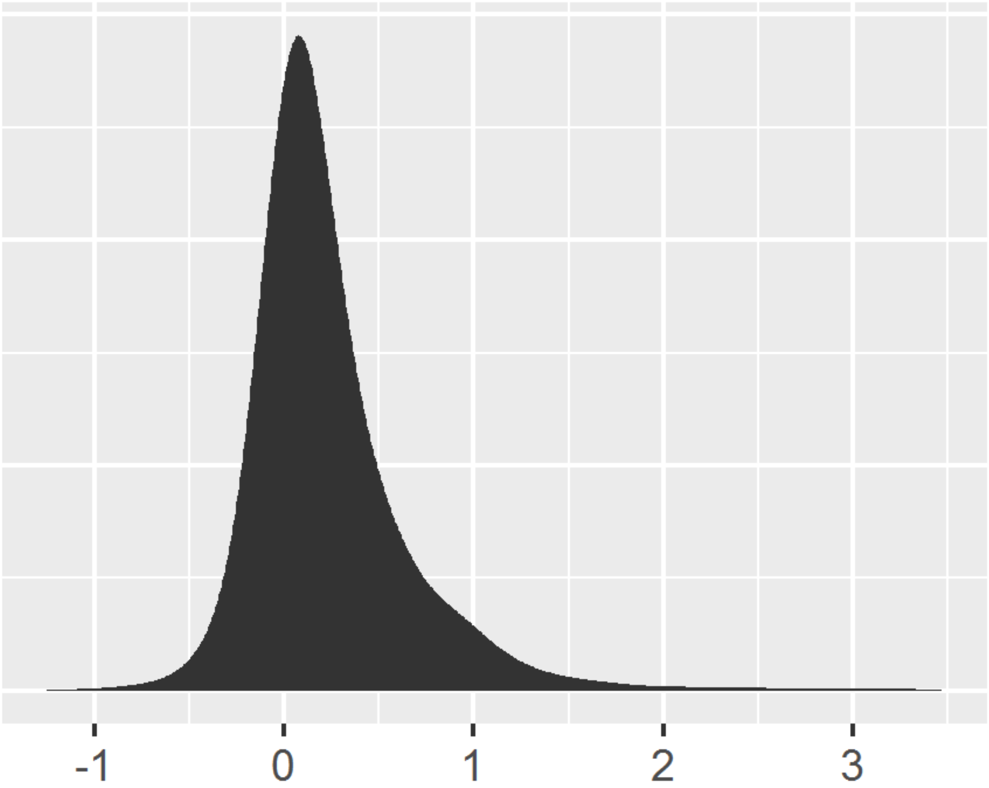
- **"There's an 80% chance that** CPC saved at least $4 PBPM."
- **"There's an 80% chance that** IAH saved at least $50 PBPM."
- **"There's an 80% chance that** MCCD reduced hospitalizations by at least 10%."
- **….Lots more P(X) = 0.80 statements….**

→ **Relative frequency of true statements = 0.80**

A Prior for the Impact of “Bright Beginnings”

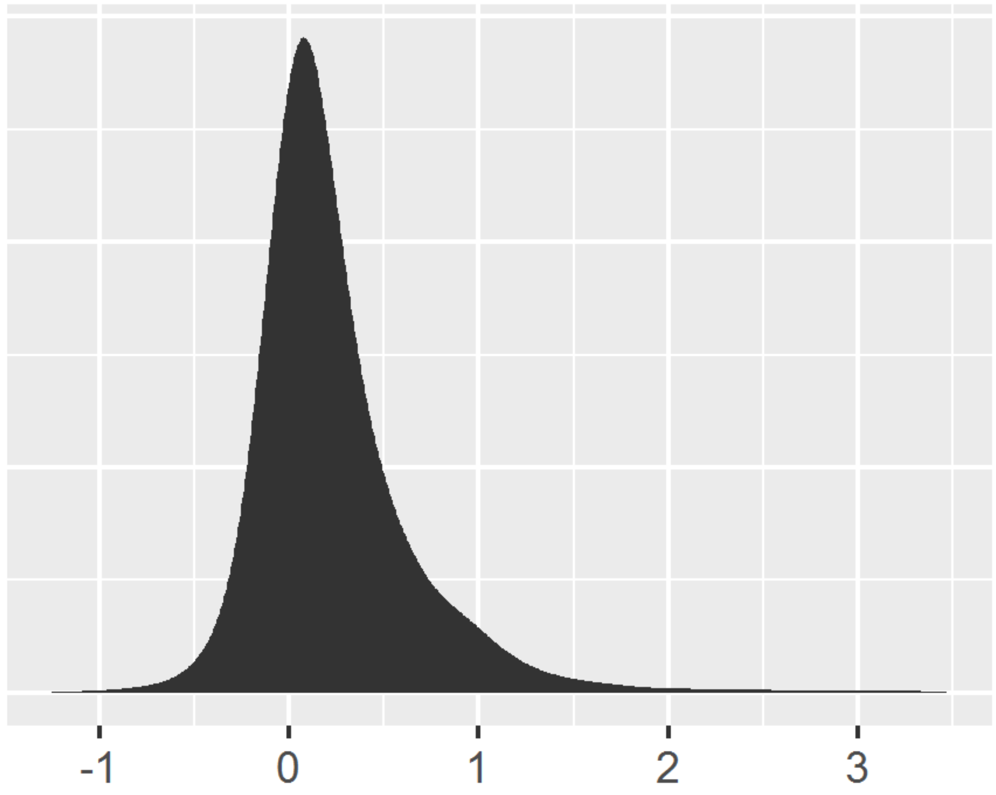
A Prior for the Impact of "Bright Beginnings"

In My Dreams: True Impacts of Bright Beginnings Replications

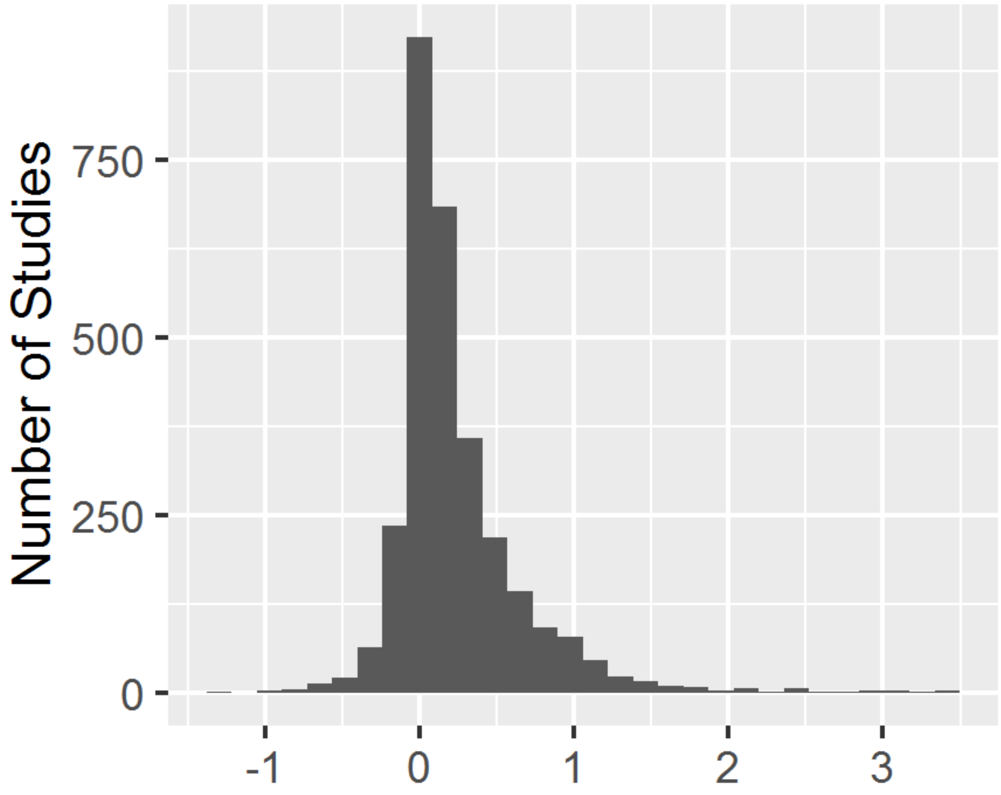


A Prior for the Impact of "Bright Beginnings"

In My Dreams: True Impacts of Bright Beginnings Replications



In Reality: WWC Impact Estimates



MATHEMATICA Policy Research

Plausible Priors Precede Persuasive Posteriors

- We need plausible priors
- A flat prior fails the sniff test
- Don’t let the perfect be the enemy of the good
- Own your prior
- Do sensitivity analyses
- Get more data

“We view much of the recent history of Bayesian inference as a set of converging messages from many directions… pointing toward the benefits of including real, subject-matter-specific, prior information in order to get more stable and accurate inferences.”

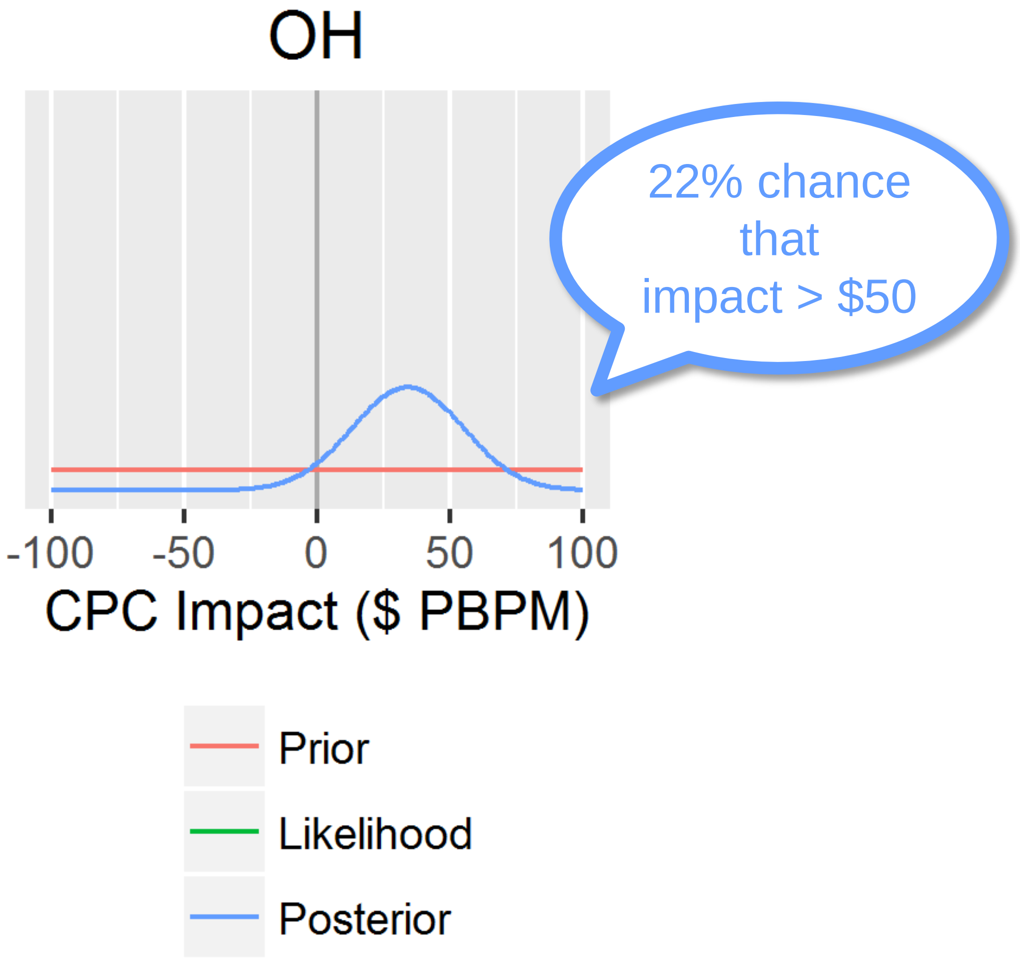
Gelman et al., 2017,
https://arxiv.org/pdf/1708.07487.pdf

- **Gelman, A., & Carlin, J. (2017). Some natural solutions to the p-value communication problem—and why they won't work.**
- **Gelman, A., & Weakliem, D. (2009). Of beauty, sex and power: Too little attention has been paid to the statistical challenges in estimating small effects. *American Scientist*, *97*(4), 310-316.**
- **Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's statement on p-values: context, process, and purpose.**
  - **Gelman, A. (2016). The problems with p-values are not just with p-values. *The American Statistician, supplemental material to the ASA statement on p-values and statistical significance*, *online discussion*.**
  - **Carlin, J.B. (2016). Is Reform Possible Without a Paradigm Shift? *The American Statistician, supplemental material to the ASA statement on p-values and statistical significance*, *online discussion*.**
- **Greenland, S., & Poole, C. (2013). Living with p values: resurrecting a Bayesian perspective on frequentist statistics. *Epidemiology*, *24*(1), 62-68.**
- **Gelman, A., Simpson, D., & Betancourt, M. (2017). The prior can generally only be understood in the context of the likelihood. *arXiv preprint arXiv:1708.07487*.**
- **www.andrewgelman.com**
  - **"Hidden dangers of noninformative priors" Nov 21, 2013**
  - **"Interpreting posterior probabilities in the context of weakly informative priors" June 28, 2015**
  - **"The general problem I have with noninformatively-derived Bayesian probabilities is that they tend to be too strong" May 1, 2015**
  - **"What are some situations in which the classical approach gives worse results than a Bayesian approach?" Nov 13, 2013**
  - **"What is the "true prior distribution"? A hard-nosed answer" April 23, 2016**

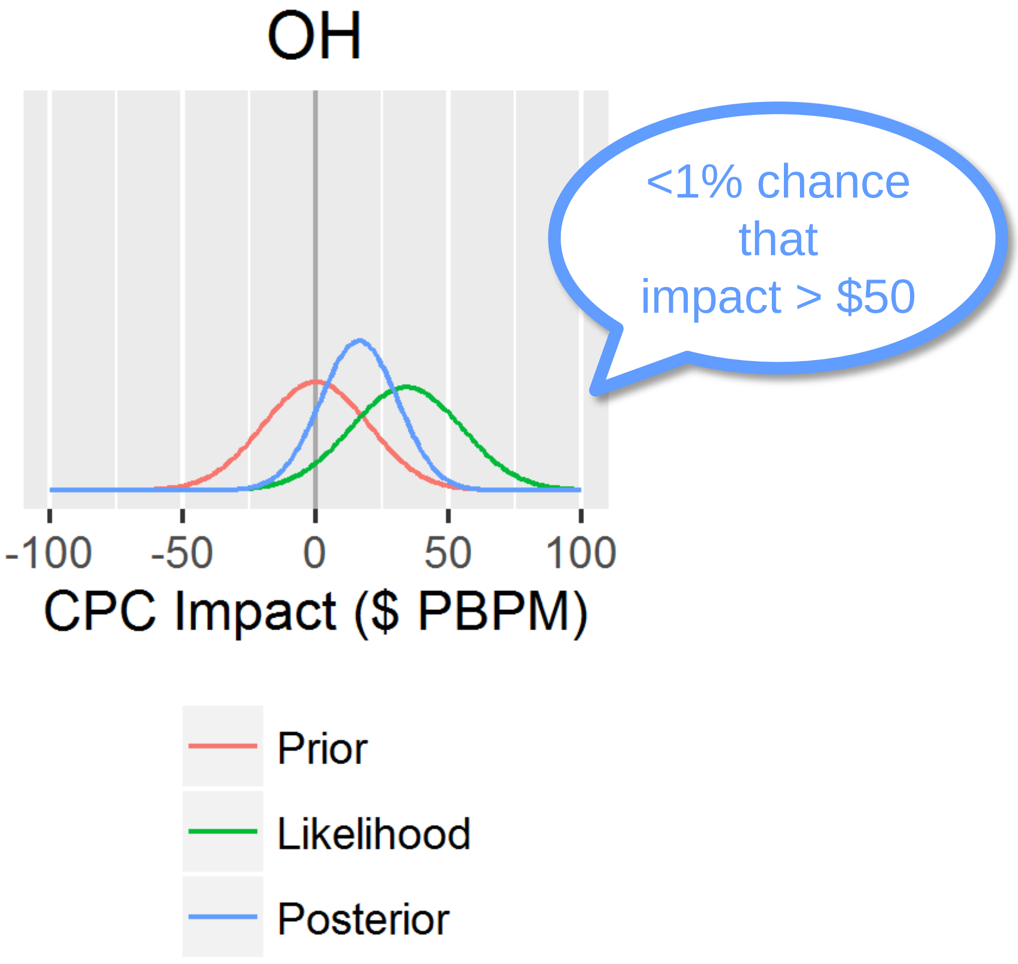
Appendix 1.
Plausible Priors Precede Persuasive Posteriors

MATHEMATICA Policy Research

Plausible Priors Precede Persuasive Posteriors



vs.



Appendix 2.
The Easy Case – A Prior Across Regions/Subgroups/Quarters

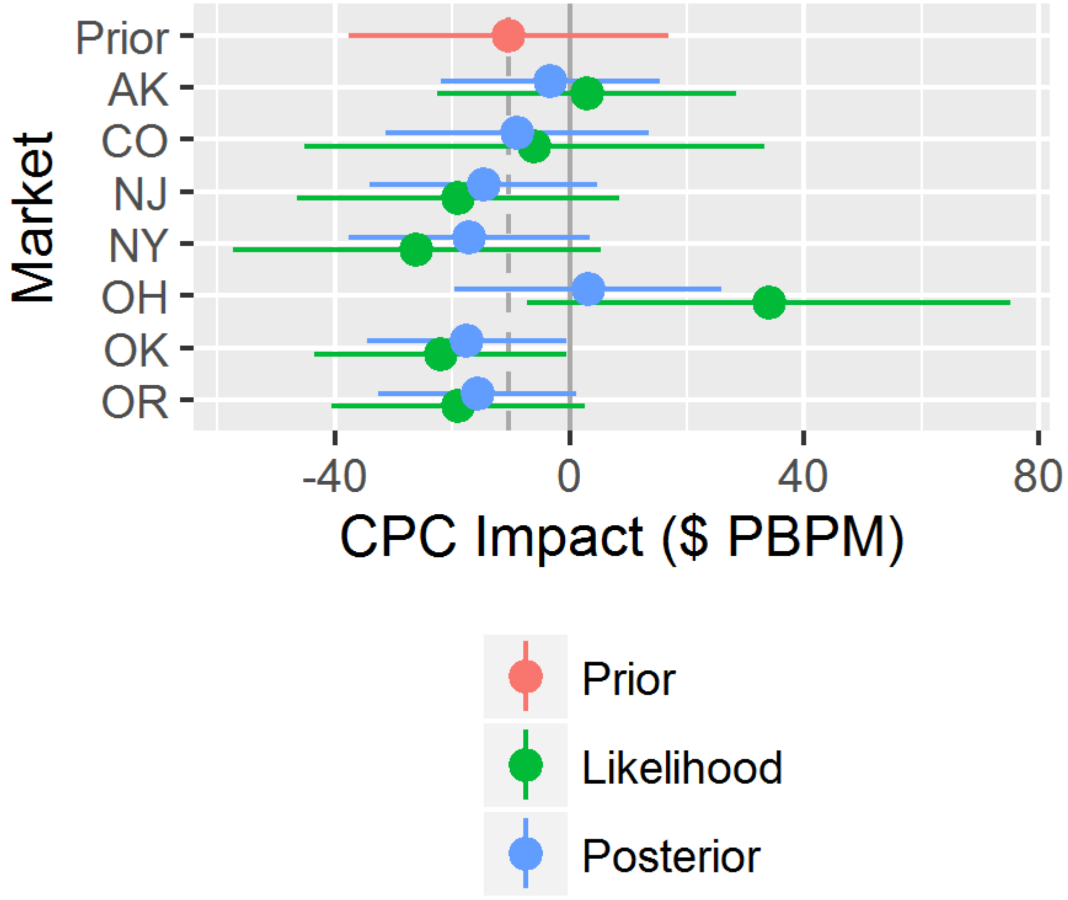
The Easy Case: A Prior Across Regions/Subgroups/Quarters

Get the prior right →

- What group of questions is your research question a random draw from?
- The true prior is the distribution of true impacts in that group.

OR NY
CO OK
NJ
OH AK

→ Get the right posterior

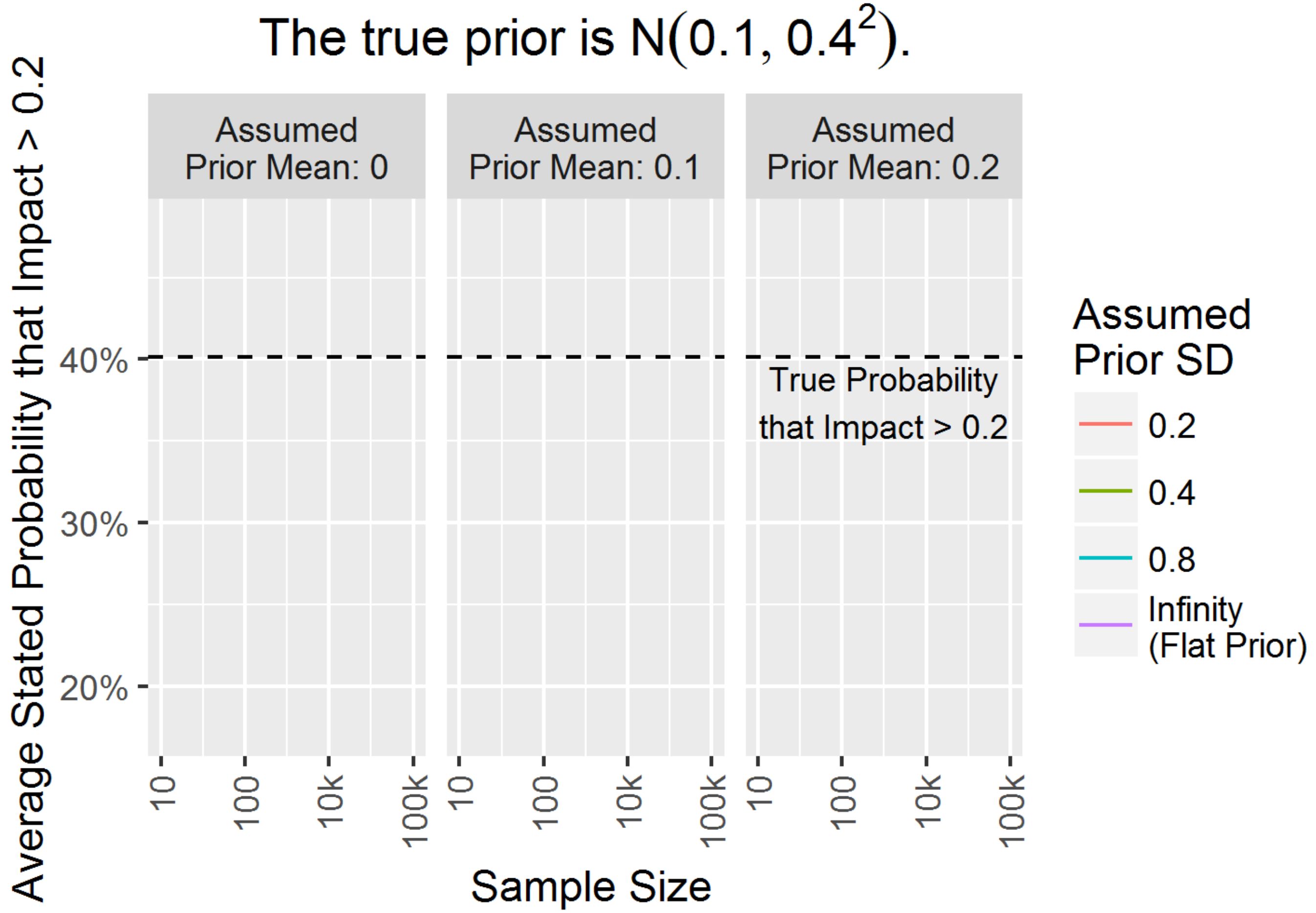


MATHEMATICA Policy Research

Appendix 3.
What Happens When We Use the Wrong Prior?

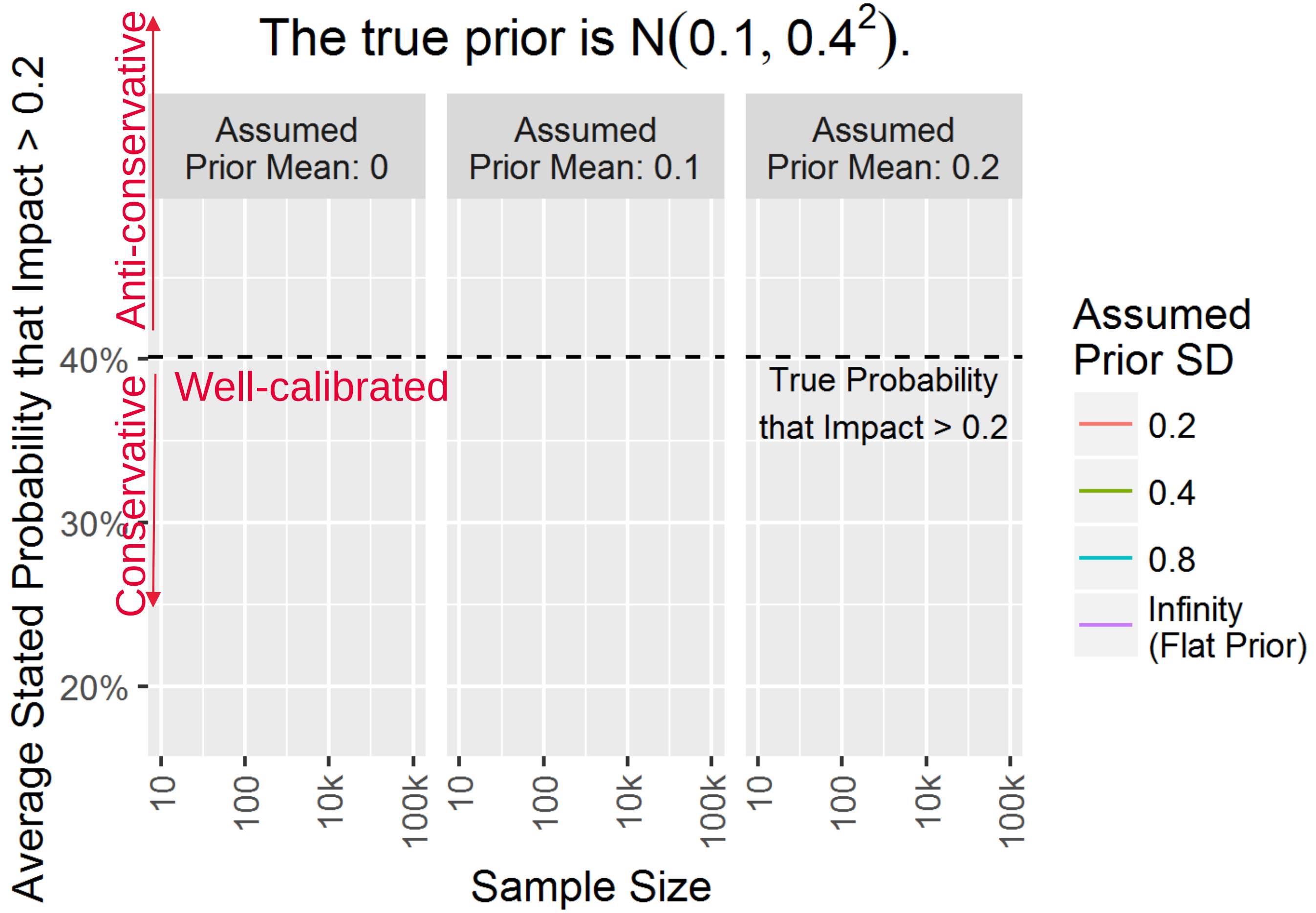
What Happens When We Use the Wrong Prior?

- **Draw impact from a true prior**
  - $\theta_s \sim N(0.1, 0.4^2)$
- **Simulate data given the drawn impact**
  - $y_s \sim P(y|\theta_s)$
- **Perform Bayesian inference assuming the wrong prior**
  - $P^*(\theta|y_s) \propto P(y_s|\theta)P^*(\theta)$
- **How bad is the resulting inference?**
  - **Compare** $P(\theta > L|y_s)$ **to** $P^*(\theta > L|y_s)$



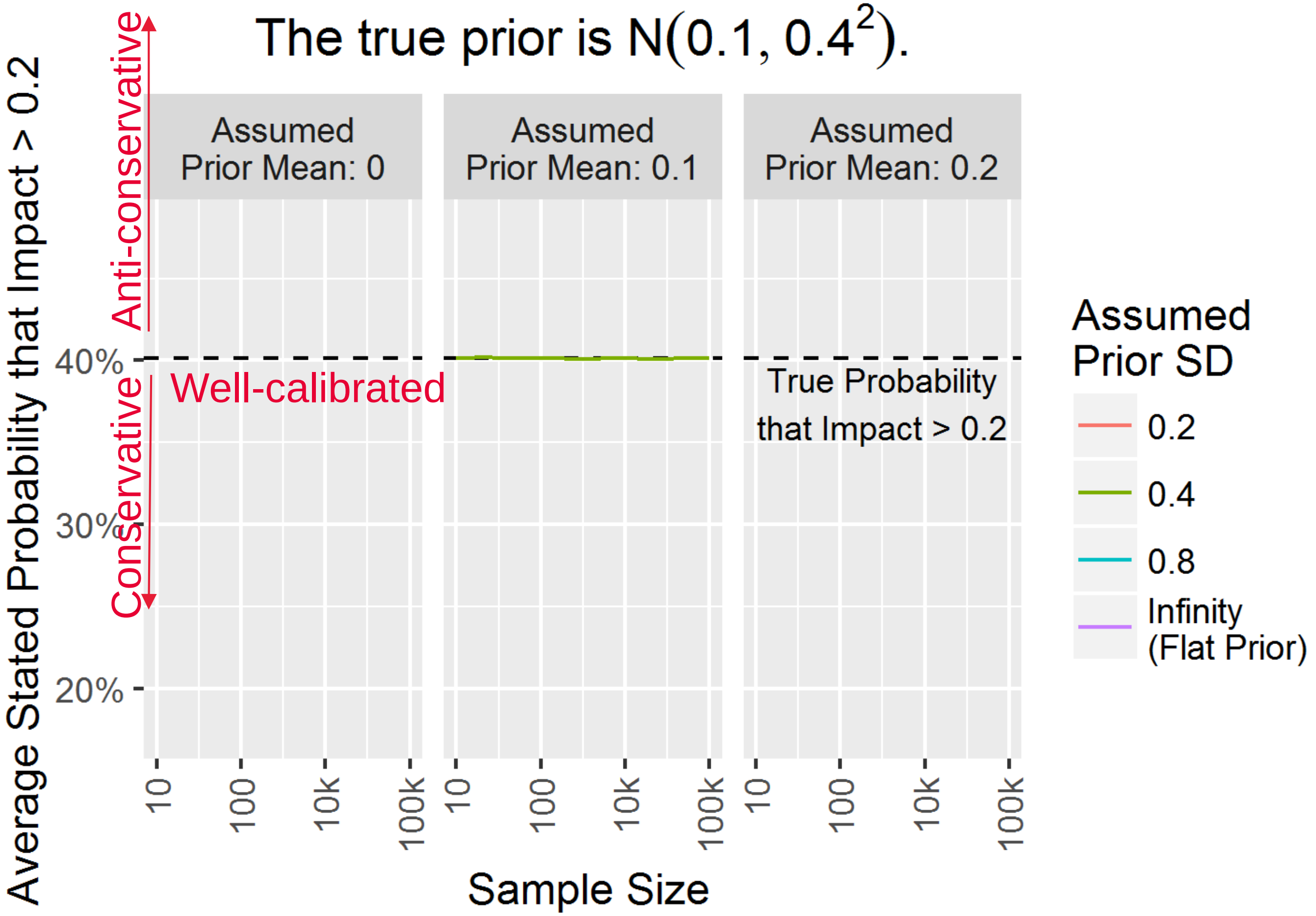
What Happens When We Use the Wrong Prior?

- **Draw impact from a true prior**
  - $\theta_s \sim N(0.1, 0.4^2)$
- **Simulate data given the drawn impact**
  - $y_s \sim P(y|\theta_s)$
- **Perform Bayesian inference assuming the wrong prior**
  - $P^*(\theta|y_s) \propto P(y_s|\theta)P^*(\theta)$
- **How bad is the resulting inference?**
  - **Compare** $P(\theta > L|y_s)$ **to** $P^*(\theta > L|y_s)$



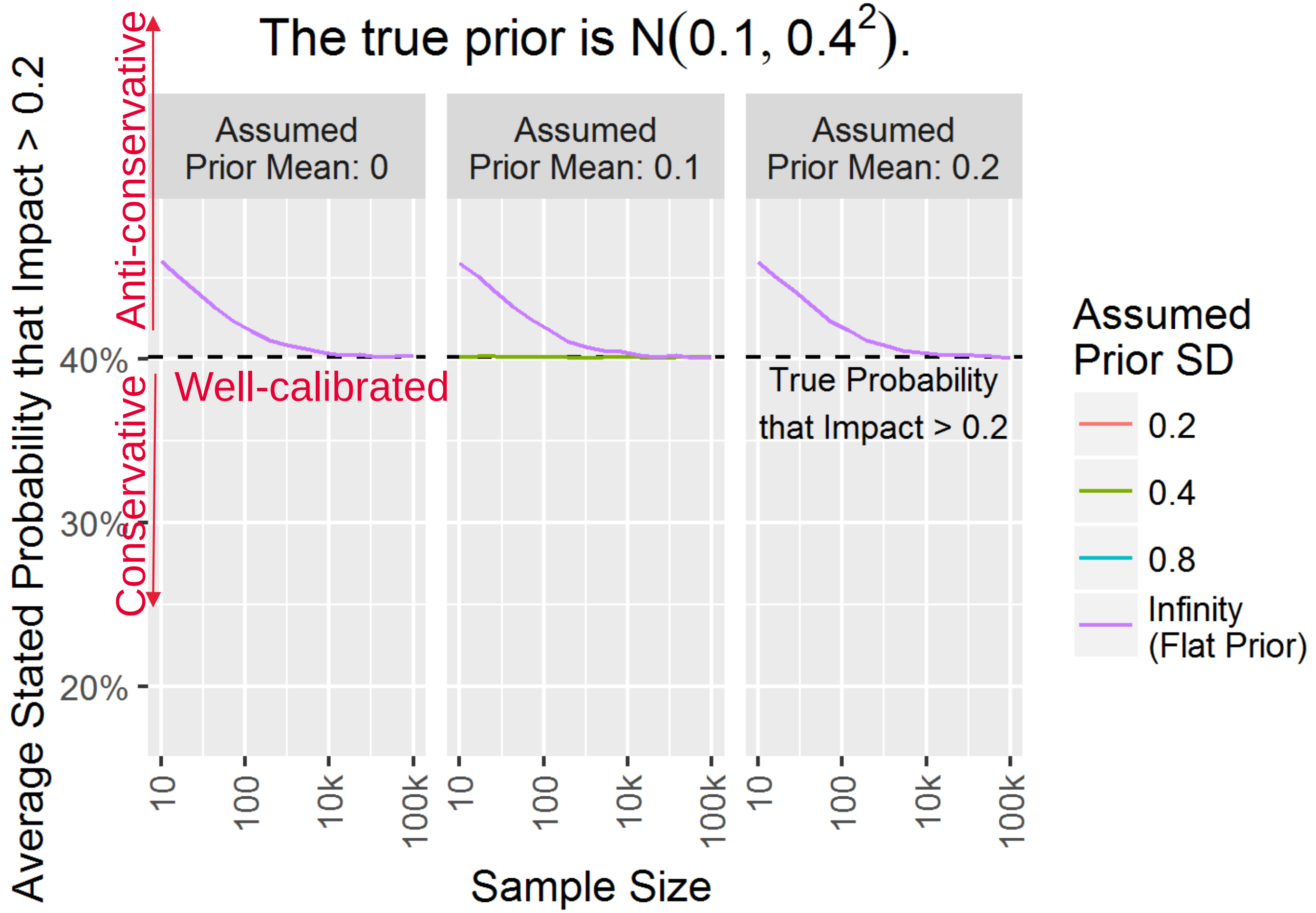
What Happens When We Use the Wrong Prior?

- **Draw impact from a true prior**
  - $\theta_s \sim N(0.1, 0.4^2)$
- **Simulate data given the drawn impact**
  - $y_s \sim P(y|\theta_s)$
- **Perform Bayesian inference assuming the wrong prior**
  - $P^*(\theta|y_s) \propto P(y_s|\theta)P^*(\theta)$
- **How bad is the resulting inference?**
  - **Compare** $P(\theta > L|y_s)$ **to** $P^*(\theta > L|y_s)$



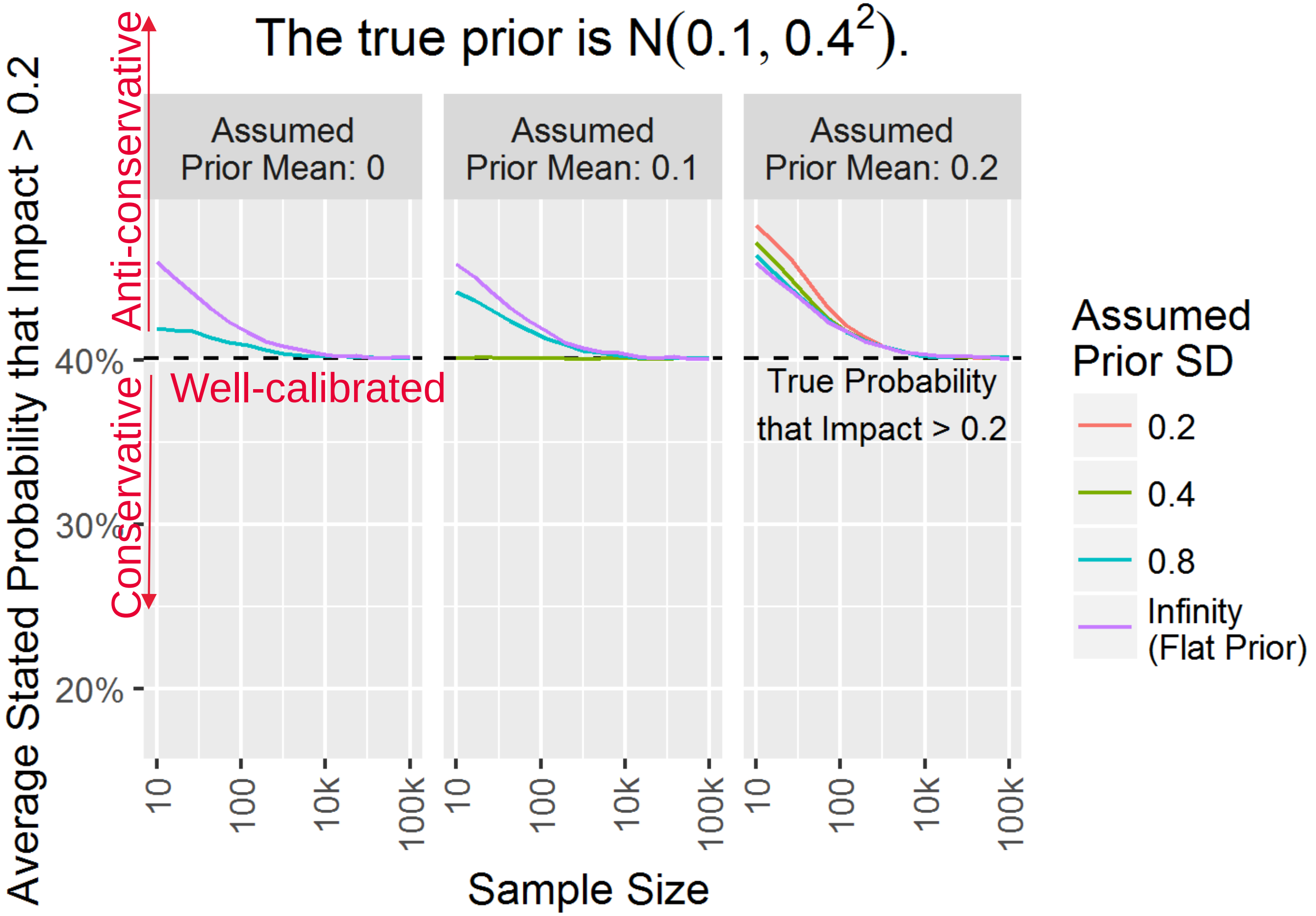
What Happens When We Use the Wrong Prior?

- **Draw impact from a true prior**
  - $\theta_s \sim N(0.1, 0.4^2)$
- **Simulate data given the drawn impact**
  - $y_s \sim P(y|\theta_s)$
- **Perform Bayesian inference assuming the wrong prior**
  - $P^*(\theta|y_s) \propto P(y_s|\theta)P^*(\theta)$
- **How bad is the resulting inference?**
  - **Compare** $P(\theta > L|y_s)$ **to** $P^*(\theta > L|y_s)$



What Happens When We Use the Wrong Prior?

- **Draw impact from a true prior**
  - $\theta_s \sim N(0.1, 0.4^2)$
- **Simulate data given the drawn impact**
  - $y_s \sim P(y|\theta_s)$
- **Perform Bayesian inference assuming the wrong prior**
  - $P^*(\theta|y_s) \propto P(y_s|\theta)P^*(\theta)$
- **How bad is the resulting inference?**
  - **Compare** $P(\theta > L|y_s)$ **to** $P^*(\theta > L|y_s)$



What Happens When We Use the Wrong Prior?

- **Draw impact from a true prior**
  - $\theta_s \sim N(0.1, 0.4^2)$
- **Simulate data given the drawn impact**
  - $y_s \sim P(y|\theta_s)$
- **Perform Bayesian inference assuming the wrong prior**
  - $P^*(\theta|y_s) \propto P(y_s|\theta)P^*(\theta)$
- **How bad is the resulting inference?**
  - **Compare** $P(\theta > L|y_s)$ **to** $P^*(\theta > L|y_s)$