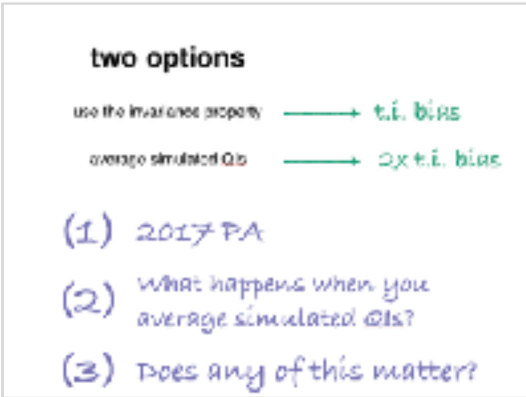
Unnecessary Bias



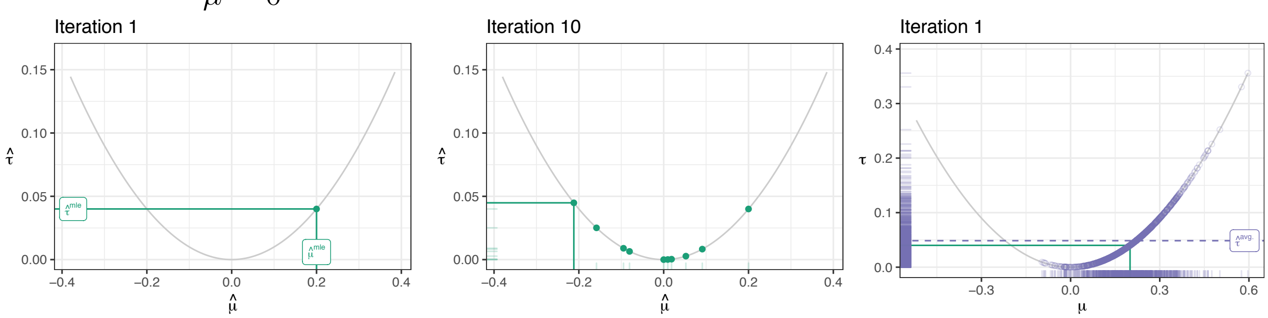
Key Point: Averaging simulated quantities of interest roughly doubles transformation induced bias. Instead, use the invariance principle to compute maximum likelihood estimates of your quantity of interest.

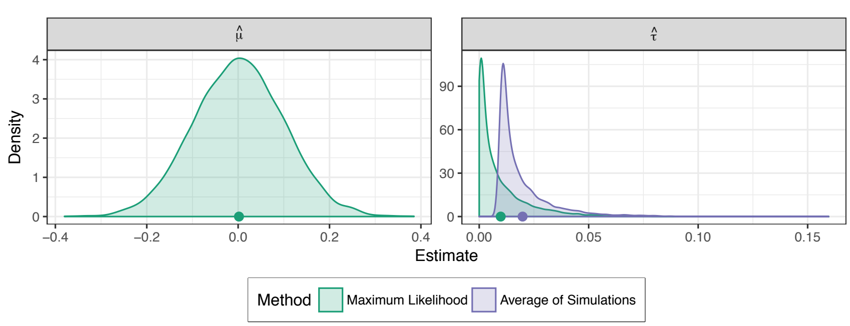
$$\text{total } \tau\text{-bias} = \underbrace{E[\tau(\hat{\beta})] - \tau[E(\hat{\beta})]}_{\text{transformation-induced}} + \overbrace{\tau[E(\hat{\beta})] - \tau(\beta)}^{\text{coefficient-induced}} .$$

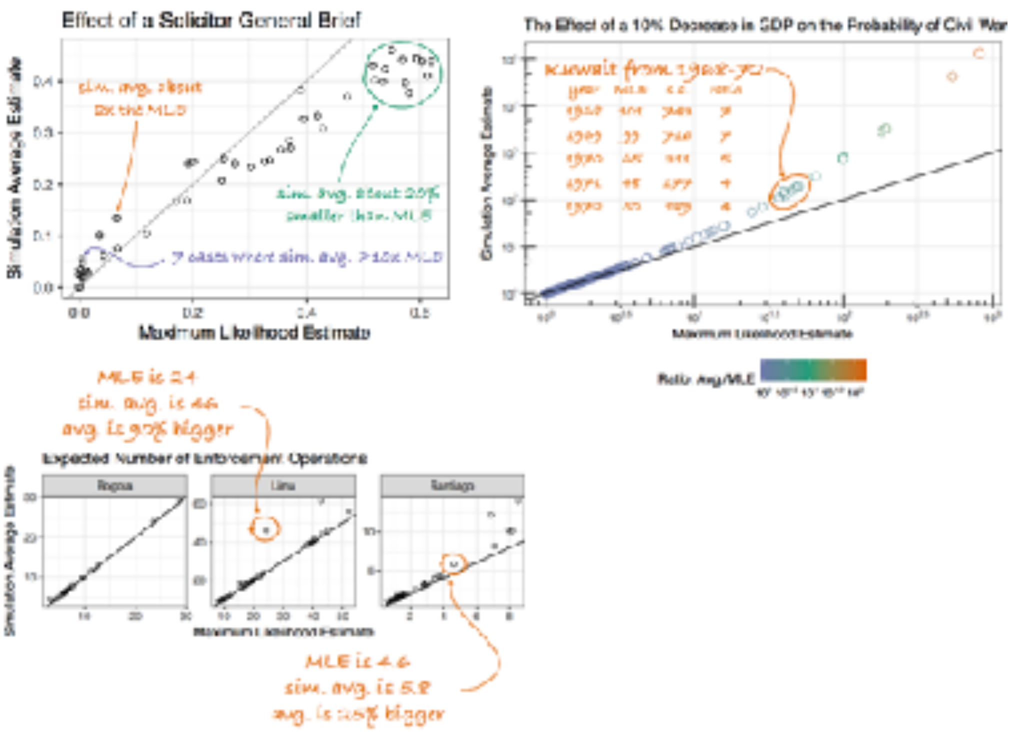
← def. of t.i. bias

$$y_i \sim \mathrm{N}(\mu, 1), \text{ for } i = 1, 2, ..., 100$$
$$\tau(\mu) = \mu^2$$
$$\mu = 0$$

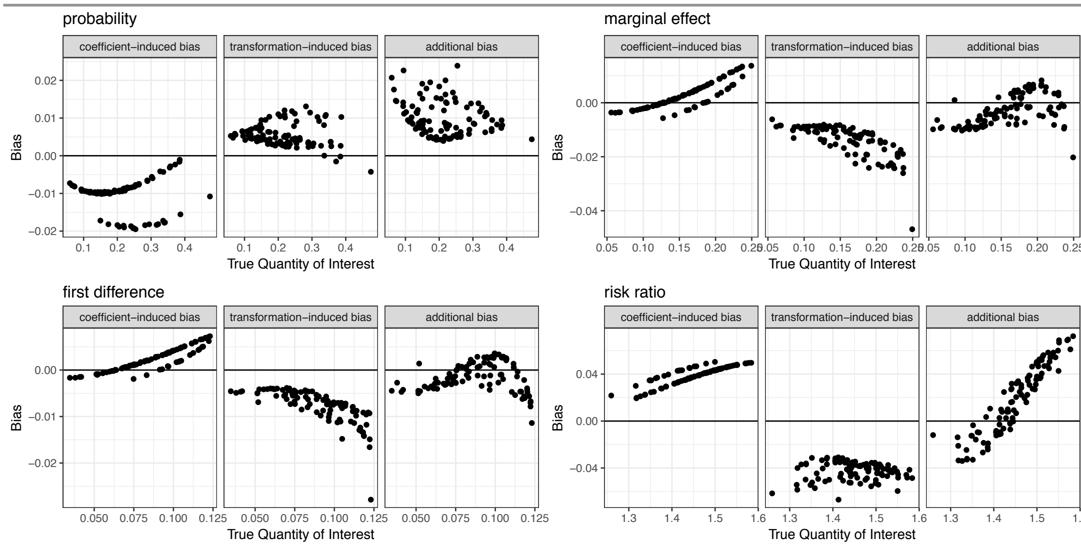
← stark illustration of the bias in the ML and sim. avg. estimates.







← diff. btw. ML and sim. avg. ests.



← comparing c.i. bias, t.i. bias. and additional bias in sim. avg. in logit model.

$$y_i \sim \text{Bernoulli}(\pi_i), \text{ for } i = 1, 2, ..., 100$$
$$\text{logit}(\pi_i) = -1.5 + x_i + 0.2w_i + 0.5z_i$$

Theorem 1 (t.i. bias, Rainey 2017) *Suppose a non-degenerate estimator $\hat{\beta}$. Then any strictly convex (concave) τ creates upward (downward) transformation-induced τ-bias.*

Proof The proof follows directly from Jensen's inequality. Suppose that the non-degenerate sampling distribution of $\hat{\beta}$ is given by $S_\beta(b)$ so that $\hat{\beta} \sim S_\beta(b)$. Then $\text{E}(\hat{\beta}) = \int_B bS_\beta(b)db$ and $\text{E}[\tau(\hat{\beta})] = \int_B \tau(b)S_\beta(b)db$. Suppose first that τ is convex. By Jensen's inequality, $\int_B \tau(b)S_\beta(b)db > \tau\left[\int_B bS_\beta(b)db\right]$, which implies that $\text{E}[\tau(\hat{\beta})] > \tau[\text{E}(\hat{\beta})]$. Because $\text{E}[\tau(\hat{\beta})] - \tau[\text{E}(\hat{\beta})] > 0$, the transformation-induced τ-bias is upward. By similar argument, one can show that for any strictly *concave* τ, $\text{E}[\tau(\hat{\beta})] - \tau[\text{E}(\hat{\beta})] > 0$ and that the transformation-induced τ-bias is downward. ■

Theorem 1 *Suppose a maximum likelihood estimator $\hat{\beta}^{mle}$. Then for any strictly convex or concave τ, the transformation-induced τ-bias for $\hat{\tau}^{avg.}$ is strictly greater in magnitude than the transformation-induced τ-bias for $\hat{\tau}^{mle}$.*

Proof According to Theorem 1 of Rainey (2017), $\text{E}\left(\hat{\tau}^{\text{mle}}\right) - \tau\left[\text{E}\left(\hat{\beta}^{\text{mle}}\right)\right] > 0$. Lemma 1 shows that for any convex τ, $\hat{\tau}^{\text{avg.}} > \hat{\tau}^{\text{mle}}$. It follows that $\underbrace{\text{E}\left(\hat{\tau}^{\text{avg.}}\right) - \tau\left[\text{E}\left(\hat{\beta}^{\text{mle}}\right)\right]}_{\text{t.i. } \tau\text{-bias in } \hat{\tau}^{\text{avg.}}} > \underbrace{\text{E}\left(\hat{\tau}^{\text{mle}}\right) - \tau\left[\text{E}\left(\hat{\beta}^{\text{mle}}\right)\right]}_{\text{t.i. } \tau\text{-bias in } \hat{\tau}^{\text{mle}}} > 0$. For the concave case, it follows similarly that $\underbrace{\text{E}\left(\hat{\tau}^{\text{avg.}}\right) - \tau\left[\text{E}\left(\hat{\beta}^{\text{mle}}\right)\right]}_{\text{t.i. } \tau\text{-bias in } \hat{\tau}^{\text{avg.}}} < \underbrace{\text{E}\left(\hat{\tau}^{\text{mle}}\right) - \tau\left[\text{E}\left(\hat{\beta}^{\text{mle}}\right)\right]}_{\text{t.i. } \tau\text{-bias in } \hat{\tau}^{\text{mle}}} < 0$

■

Lemma 1 *Suppose a maximum likelihood estimator $\hat{\beta}^{mle}$. Then any strictly convex (concave) τ guarantees that $\hat{\tau}^{avg.}$ is strictly greater [less] than $\hat{\tau}^{mle}$.*

Proof By definition,

$$\hat{\tau}^{\text{avg.}} = \text{E}\left[\tau\left(\tilde{\beta}\right)\right].$$

Using Jensen's inequality, we know that $\text{E}\left[\tau\left(\tilde{\beta}\right)\right] > \tau\left[\text{E}\left(\tilde{\beta}\right)\right]$, so that

$$\hat{\tau}^{\text{avg.}} > \tau\left[\text{E}\left(\tilde{\beta}\right)\right].$$

However, because $\tilde{\beta} \sim N\left[\hat{\beta}^{\text{mle}}, \hat{V}\left(\hat{\beta}^{\text{mle}}\right)\right]$, $\text{E}\left(\tilde{\beta}\right) = \hat{\beta}^{\text{mle}}$, so that

$$\hat{\tau}^{\text{avg.}} > \tau\left(\hat{\beta}^{\text{mle}}\right).$$

Of course, $\hat{\tau}^{\text{mle}} = \tau\left(\hat{\beta}^{\text{mle}}\right)$ by definition, so that

$$\hat{\tau}^{\text{avg.}} > \hat{\tau}^{\text{mle}}.$$

The proof for concave τ follows similarly. ■