APPENDIX
A Careful Consideration of CLARIFY
Simulation-Induced Bias in Point Estimates of Quantities of Interest

Carlisle Rainey[†]

December 13, 2022

A Proofs

A.1 Proof of Lemma 1

Proof By definition,

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

Using Jensen's inequality (Casella and Berger 2002, p. 190, Thm. 4.7.7), $\text{E}\left[\tau\left(\tilde{\beta}\right)\right] > \tau\left[\text{E}\left(\tilde{\beta}\right)\right]$, so that

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

However, because $\tilde{\beta} \sim \text{MVN}\left[\hat{\beta}^{\text{mle}}, \hat{V}\left(\hat{\beta}^{\text{mle}}\right)\right]$, $\text{E}\left(\tilde{\beta}\right) = \hat{\beta}^{\text{mle}}$, 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. ■

A.2 Proof of Theorem 1

Proof According to Theorem 1 of Rainey (2017, p. 405), $\text{E}\left(\hat{\tau}^{\text{mle}}\right) - \tau\left[\text{E}\left(\hat{\beta}^{\text{mle}}\right)\right] > 0$. Lemma ?? shows that for any convex τ, $\hat{\tau}^{\text{avg}} > \hat{\tau}^{\text{mle}}$. It follows that $\underbrace{\text{E}\left(\hat{\tau}^{\text{avg}}\right) - \tau\left[\text{E}\left(\hat{\beta}^{\text{mle}}\right)\right]}_{\text{s.i. and t.i. } \tau\text{-bias in } \hat{\tau}^{\text{avg}}} >$

[†]Carlisle Rainey is Associate Professor of Political Science, Florida State University, 540 Bellamy, Tallahassee, FL, 32306. (crainey@fsu.edu).

$\underbrace{\mathrm{E}\left(\hat{\tau}^{\text{mle}}\right) - \tau\left[\mathrm{E}\left(\hat{\beta}^{\text{mle}}\right)\right]}_{\text{t.i. } \tau\text{-bias in } \hat{\tau}^{\text{mle}}} > 0.$

For the concave case, it follows similarly that $\underbrace{\mathrm{E}\left(\hat{\tau}^{\text{avg}}\right) - \tau\left[\mathrm{E}\left(\hat{\beta}^{\text{mle}}\right)\right]}_{\text{s.i. and t.i. } \tau\text{-bias in } \hat{\tau}^{\text{avg}}} < \underbrace{\mathrm{E}\left(\hat{\tau}^{\text{mle}}\right) - \tau\left[\mathrm{E}\left(\hat{\beta}^{\text{mle}}\right)\right]}_{\text{t.i. } \tau\text{-bias in } \hat{\tau}^{\text{mle}}} < 0.$ ■

B Additional Analysis of the Drastic, Convex Transformation

In the main text, I develop an intuition for the simulation-induced τ-bias in $\hat{\tau}^{\text{avg}}$ using the simple (unrealistic, but heuristically useful) scenario in which $y_i \sim \mathrm{N}(0,1)$, for $i \in \{1, 2, \ldots, n = 100\}$, and the researcher wishes to estimate μ^2. Suppose that the researcher knows that the variance equals one but does not know that the mean μ equals zero. The researcher uses the unbiased ML estimator $\hat{\mu}^{\text{mle}} = \frac{\sum_{i=1}^n y_i}{n}$ of μ, but ultimately cares about the quantity of interest $\tau(\mu) = \mu^2$. The researcher can use the plug-in estimator $\hat{\tau}^{\text{mle}} = \left(\hat{\mu}^{\text{mle}}\right)^2$ of $\tau(\mu)$. Alternatively, the researcher can use the average-of-simulations estimator, estimating $\tau(\mu)$ as $\hat{\tau}^{\text{avg}} = \frac{1}{M}\sum_{i=1}^M \tau\left(\tilde{\mu}^{(i)}\right)$, where $\tilde{\mu}^{(i)} \sim \mathrm{N}\left(\hat{\mu}^{\text{mle}}, \frac{1}{\sqrt{n}}\right)$ for $i \in \{1, 2, \ldots, M\}$.

Below, I calculate the bias of each estimator.[1]

B.1 The Bias in the ML Estimator

To simplify the notation below, I use $\hat{\mu}$ in place of $\hat{\mu}^{\text{mle}}$.

First, note that $\hat{\mu} = \frac{\sum_{i=1}^n y_i}{n}$ is an *unbiased* estimator so that $\mathrm{E}(\hat{\mu}) = \mu = 0$. We then have the common identity for mean-squared error: $\mathrm{E}\left((\hat{\mu} - \mu)^2\right) = \mathrm{Var}\left(\hat{\mu}\right) - \mathrm{E}(\hat{\mu} - \mu)^2$. Substituting $\mu = 0$, we have $\mathrm{E}\left(\hat{\mu}^2\right) = \mathrm{Var}\left(\hat{\mu}\right) - \mathrm{E}(\hat{\mu})^2$. Substituting $\mathrm{E}(\hat{\mu}) = \mu = 0$, we have $\mathrm{E}\left(\hat{\mu}^2\right) = \mathrm{Var}\left(\hat{\mu}\right)$. Then $\mathrm{E}\left(\hat{\mu}^2\right) = \mathrm{Var}\left(\frac{\sum_{i=1}^n y_i}{n}\right) = \frac{1}{n^2}\mathrm{Var}\left(\sum_{i=1}^n y_i\right)$. Then, using the identify that the variance of the sum of independent random variables is the sum of their variances, we have $\mathrm{E}\left(\hat{\mu}^2\right) = \frac{1}{n^2}(n \times 1) = \frac{1}{n}$.

Since $\tau = \mu^2 = 0$, the bias in $\hat{\tau} = \left[\hat{\mu}^{\text{mle}}\right]^2$ is $\frac{1}{n} - 0 = \frac{1}{n}$. Because there is no coefficient-induced bias, this is also the transformation-induced bias.

B.2 The Bias in the Average-of-Simulations Estimator

To simplify the notation below, I use $\bar{\tau}$ in place of $\hat{\tau}^{\text{avg}}$.

[1] I thank a reviewer for pointing out these results.

First, compute $\mathrm{E}(\bar{\tau} \mid \hat{\mu}) = \mathrm{E}\left[\frac{1}{M}\sum_{i=1}^{M}\left(\tilde{\mu}^{(i)}\right)^2\right] = \frac{1}{M}\sum_{i=1}^{M}\mathrm{E}\left[\left(\tilde{\mu}^{(i)}\right)^2\right]$. Then we have $\mathrm{E}(\bar{\tau} \mid \hat{\mu}) = \frac{1}{M}\sum_{i=1}^{M}\left[\mathrm{Var}(\tilde{\mu}^{(i)}) + \mathrm{E}\left(\tilde{\mu}^{(i)}\right)^2\right]$. Substituting known values, we have $\mathrm{E}(\bar{\tau} \mid \hat{\mu}) = \frac{1}{M}\sum_{i=1}^{M}\left[\frac{1}{n} + \hat{\mu}^2\right]$. Simplifying, we have $\mathrm{E}(\bar{\tau} \mid \hat{\mu}) = \frac{1}{M}\left[\frac{M}{n} + M\hat{\mu}^2\right] = \frac{1}{n} + \hat{\mu}^2$.

Next, apply the law of iterated expectations to find $\mathrm{E}(\bar{\tau}) = \mathrm{E}(\bar{\tau} \mid \hat{\mu})$. Substituting, we have $\mathrm{E}(\bar{\tau}) = \mathrm{E}(\frac{1}{n} + \hat{\mu}^2)$. Then, simplifying, we have $\mathrm{E}(\bar{\tau}) = \frac{1}{n} + \mathrm{E}(\hat{\mu}^2) = \frac{1}{n} + \frac{1}{n} = \frac{2}{n}$.

The bias in $\hat{\tau}^{\text{avg}}$ is therefore $\frac{2}{n}$. Because simulation-induced bias is defined as $\mathrm{E}\left(\hat{\tau}^{\text{avg}}\right) - \mathrm{E}\left(\hat{\tau}^{\text{mle}}\right)$, the simulation-induced bias in this example is $\frac{2}{n} - \frac{1}{n} = \frac{1}{n}$. Thus, the simulation-induced and transformation-induced bias are exactly equal and the average-of-simulations estimator exactly doubles the bias in the ML estimator.

References

Casella, George, and Roger L. Berger. 2002. *Statistical Inference*. 2nd ed. Pacific Grove, CA: Duxbury.

Rainey, Carlisle. 2017. "Transformation-Induced Bias: Unbiased Coefficients Do Not Imply Unbiased Quantities of Interest." *Political Analysis* 25:402–409.