Statistical inferences

Lucie Lu

May 13, 2018

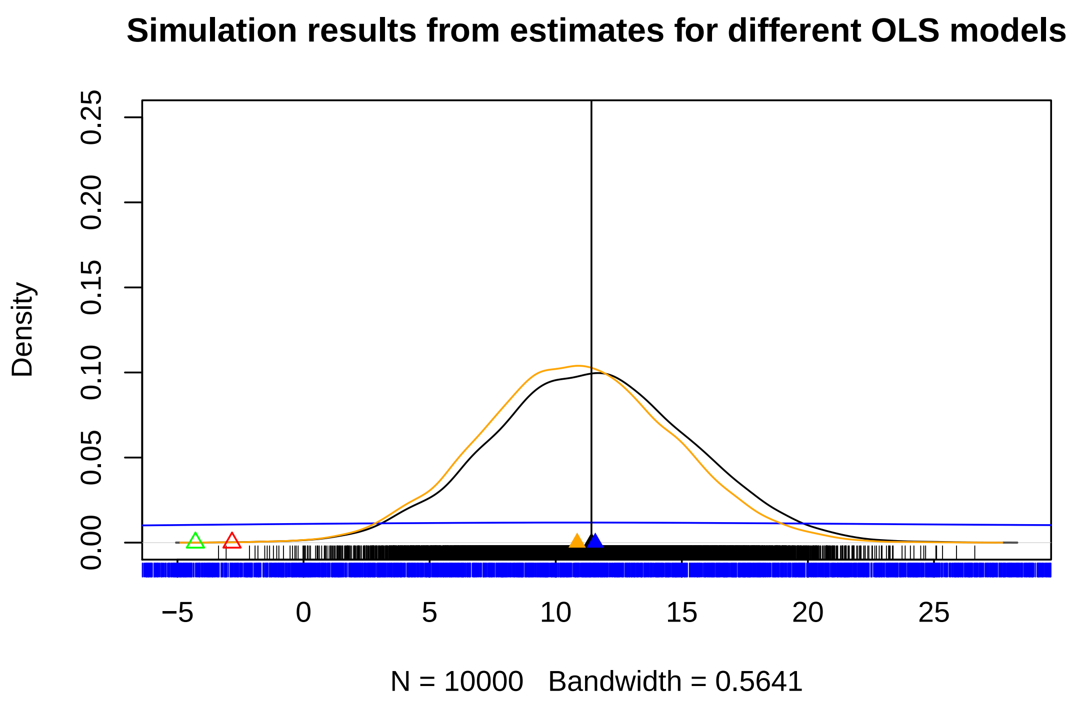
Checking unbiasedness and consistency of the estimates in the OLS model

% latex table generated in R 3.4.3 by xtable 1.8-2 package % Sun May 13 16:35:48 2018

	bestATE	unbiasedATE	bestATE2
trueATE	11.42	11.42	11.42
sampmeans	11.42	11.57	10.85
bias	0.00	0.15	0.57
sd	3.96	31.82	3.78
MSE	15.64	1012.36	14.58

Table 1: Simulation results from different estimates for OLS model

```
Warning in rug(estdists["unbiasedATE", ], col = "blue", line = 0.5): some values will be clipped
```



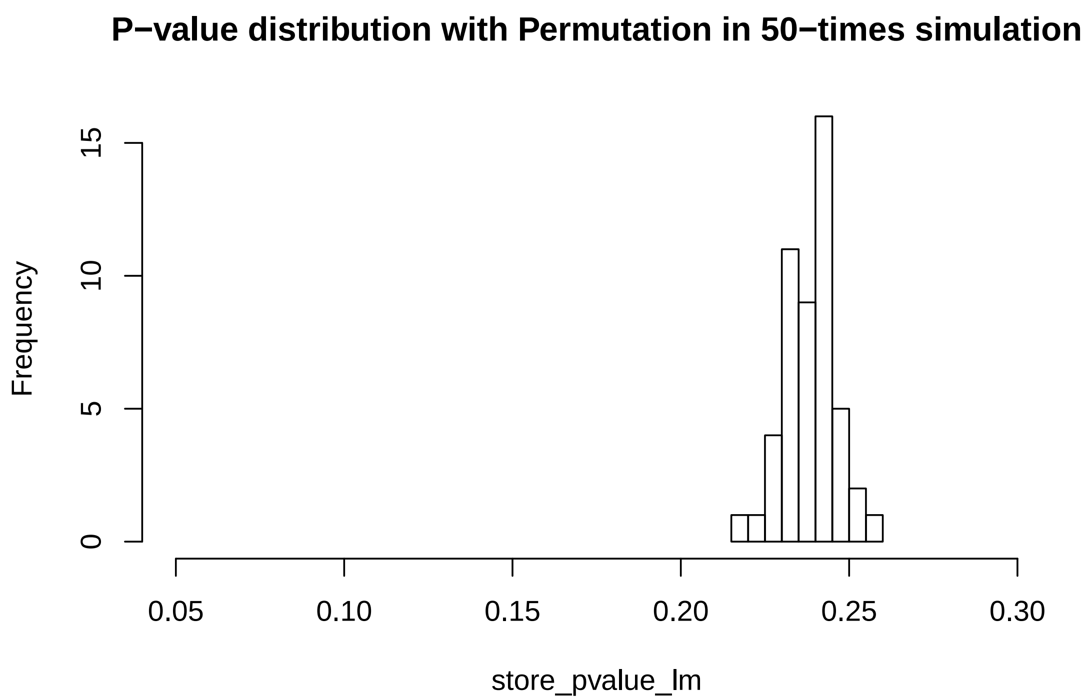
The *bestATE* is the estimaor in a *lm* function with all the relevant covariates in the model. The *unbiasedATE* is the estimaor in a *lm* function with no covariates in the model at all. The *bestATE2* is the estimaor in a residual-based function.

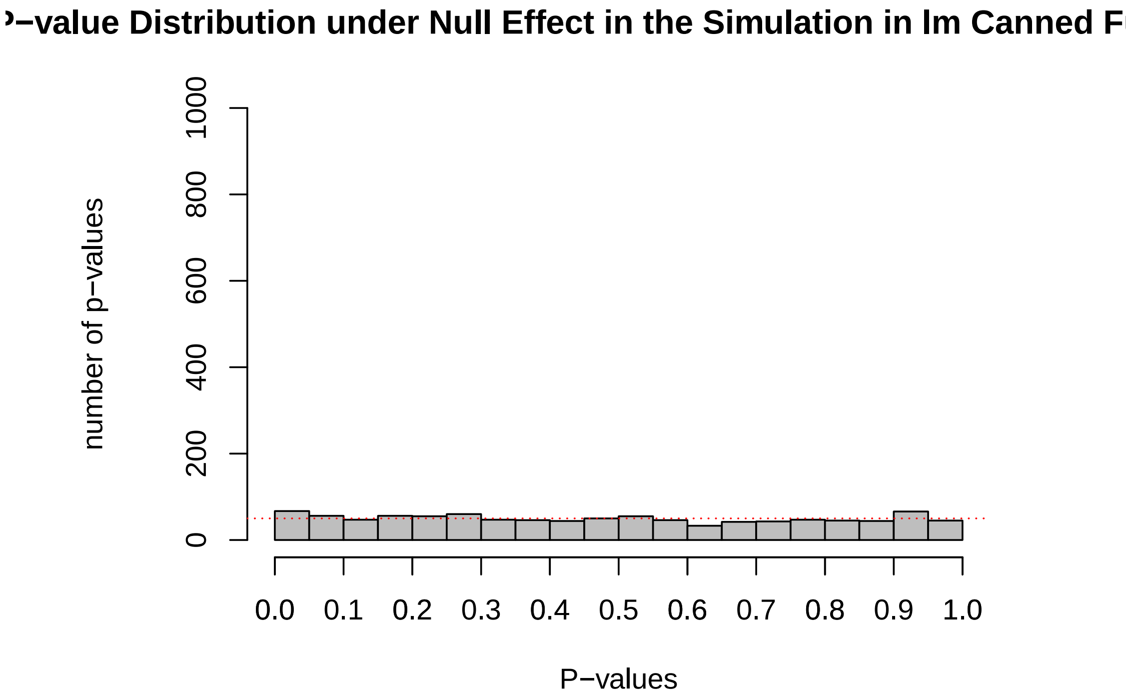
To assess biasedness, I compare sample means of the estimators with the true means I created in the simulation test. In a simple way, I can check the thrid row bias (denoted as the absoluate positive values: the differences between trueATE and the means of the estimates). All of the three estimators are close to the true mean (11.416) I created in the simulation test. This suggest all three of them are (pretty much) unbiased. In fact, the *bestATE* has the lowest bias out of the three. Its absolute distance to the true mean is the smallest one, with bias equals to 0.0025. The *bestATE2* is slightly biased (with biase 0.5661).

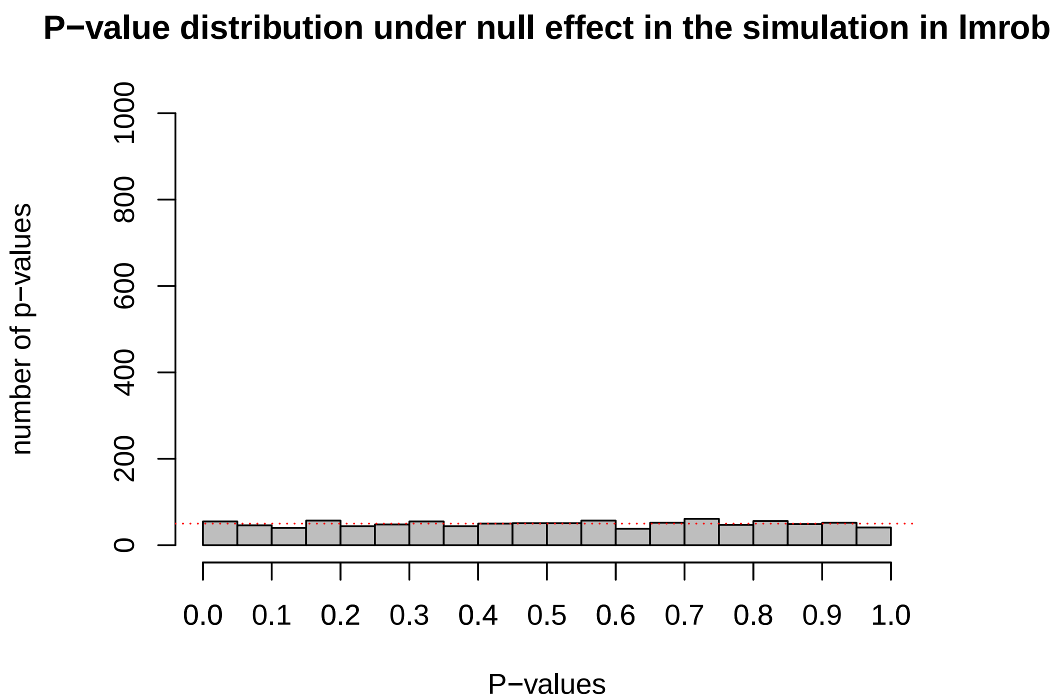
To assess consistency or efficiency, the MSEs for bestATE and bestATE2 are relatively the same, but the bestATE2 behave slightly better ($14.575 < 15.6409$). The estimators *bestATE* and *bestATE2* are efficient, compared to the unbiasedATE. The standard error for unbiasedATE is very high (31.8188), suggesting that this estimate is very inefficient and inconsistent. It does not converge to the true mean at all. From this simulation test, because the bestATE has the lowest bias and is the most efficient one, this estimate is preferred.

P-values from the permutation test

[1] 0.2268





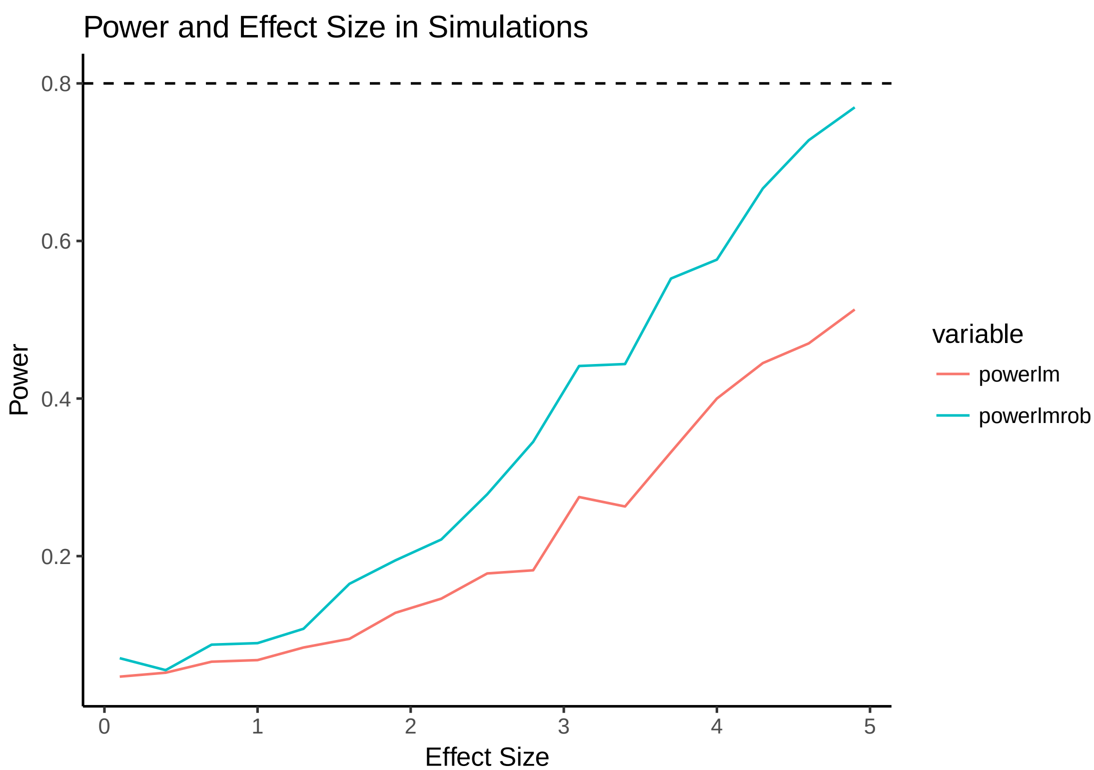


Power and effect size

% latex table generated in R 3.4.3 by xtable 1.8-2 package % Sun May 13 16:52:08 2018

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
effectsize	0.10	0.40	0.70	1.00	1.30	1.60	1.90	2.20	2.50	2.80	3.10	3.40	3.70	4.00	4.30
somepower	0.05	0.05	0.07	0.07	0.08	0.10	0.13	0.15	0.18	0.18	0.28	0.26	0.33	0.40	0.44
somepower_lmrob	0.07	0.06	0.09	0.09	0.11	0.16	0.19	0.22	0.28	0.35	0.44	0.44	0.55	0.58	0.67

Table 2: Simulation results of Power and Effect Size for lm and lmrob model



Power analysis of two models:

We can use simulations to estimate the statistical power of a model. The statistical power is the probability of observing a statistically significant result, if there is a true effect. When there is an effect, I hope that my statistical test is able to detect it. This denotes to high power in my study.

Cohen describes effect size as "the degree to which the null hypothesis is false." In this simulation test, I generate different hypothetical effect sizes (from 0.1 to 5), and I calculate the number of p-values that are are lower than 0.05 ("reject the null") when I know there is a true effect (the null is false). When the effect size increases, the powers in both functions also increase.

For a given sample size, the *lmrob* model has larger statistical power given an effect size. As effect size increases, the power of the *lmrob* model is also increasing faster than that of the *lm* model. This is probably due to a relatively small sample size in this study (81 countries). To achieve an ideal 80% statistical power, the *lmrob* model requires an effect size larger than 5. 80% statistical power essentially means when there is a true effect, there is 80 percent that I will observe a signifiant effect. For this *lm* model, I need a bigger effect size to achieve the same level of power as *lmrob* model requires.