p-uniform: Unbiased effect size estimation in the presence of publication bias

Robbie C.M. van Aert (R.C.M.vanAert@uvt.nl), M. A. L. M. van Assen, & J. M. Wicherts
Tilburg University, Department of Methodology and Statistics

1. Background

Publication bias is omnipresent in science and yields overestimated effect sizes in meta-analysis. *No* current meta-analysis technique can accurately estimate effect size in the presence of publication bias. However, p-uniform is the only method able to:

(i) Test if publication bias exists
(ii) Adequately test the null-hypothesis of no effect
(iii) Provide an unbiased effect size estimate (in theory) when there is publication bias

2. p-uniform

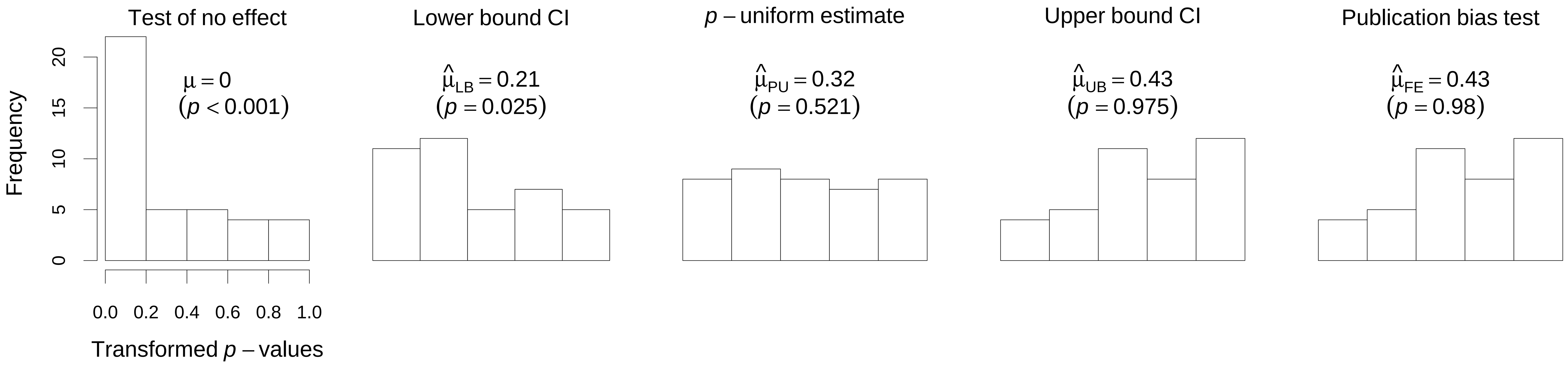
p-uniform only considers statistically significant studies, and discards all the non-significant studies. Two assumptions are underlying p-uniform:

(i) Homogeneous population effect size
(ii) The probability of selecting a statistically significant study in a meta-analysis is independent of its p-value:

$$f(p_i) = \text{Constant for } p_i \leq \alpha$$

The main idea behind p-uniform is that the distribution of p-values conditional on the true effect size is uniform.

3. Illustration: Distribution of p-values as a function of effect



4. Simulations

Results of p-uniform were compared to traditional fixed-effect meta-analysis and the trim-and-fill method for effect size estimation and testing H_0: $\mu = 0$ and to the Test of Excess Significance (TES) for examining publication bias.

Conditions:

$N = 25$ and the expected number of significant studies is 8

$K = 160$	$d = \mu = 0$	$\alpha = .05$
$K = 40$	$d = \mu = 0.16$	$1 - \beta = 0.2$
$K = 16$	$d = \mu = 0.33$	$1 - \beta = 0.5$
$K = 10$	$d = \mu = 0.5$	$1 - \beta = 0.8$

Publication bias (the probability of including a non-significant study in a meta-analysis) was varied from 0 0.025 0.05 0.25 0.5 1

5. Results: Publication bias test

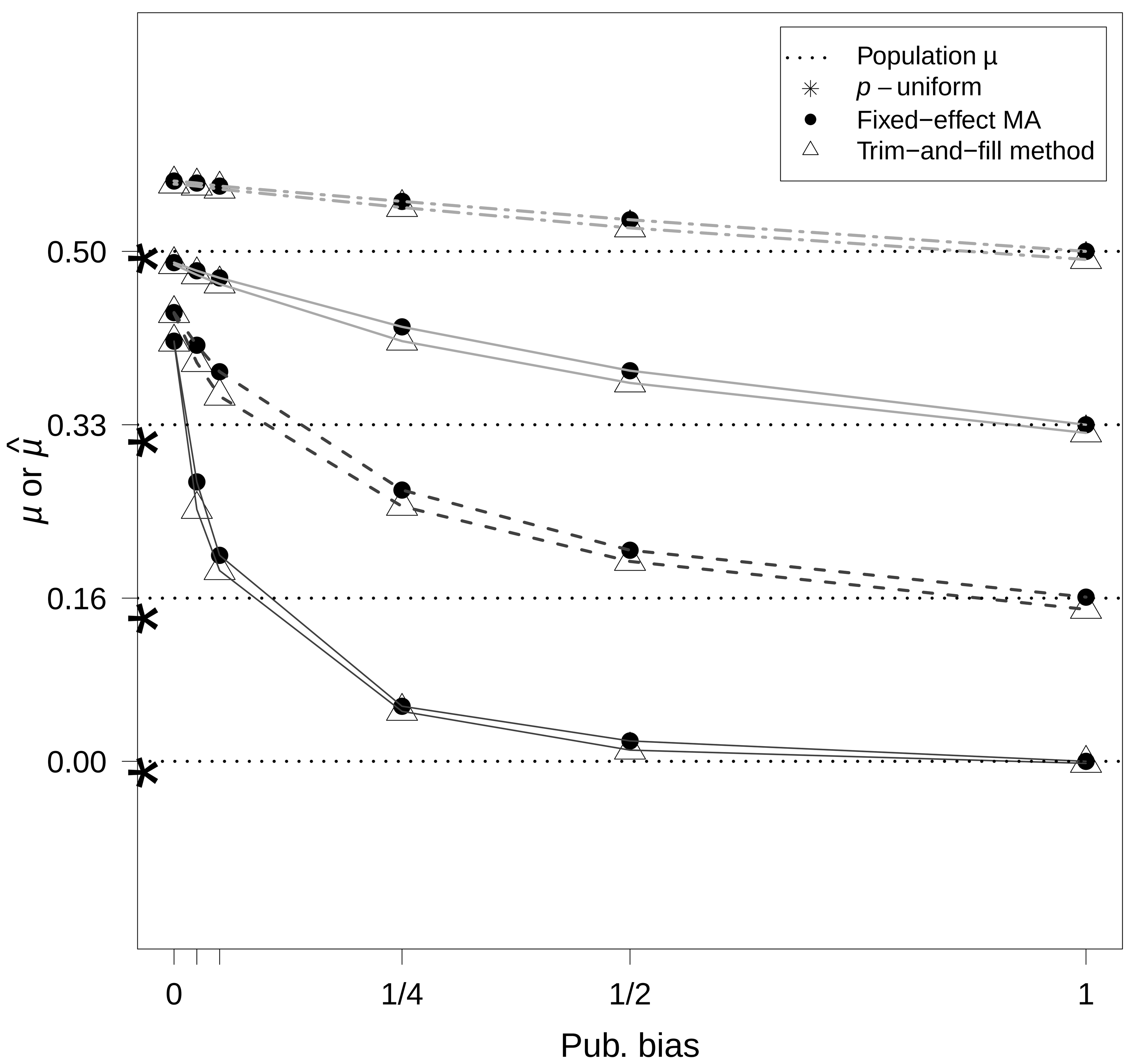
p-uniform was generally superior to the TES with respect to statistical power. Type-I error rates of p-uniform were close to 0.05 for $\mu < 0.5$.

Table: Results of simulations (10,000 replications) on Type-I error rates and statistical power

			Pub. bias					
			0	1/40	1/20	1/4	1/2	1
μ (K)	0 (160)	p-uniform	**0.902**	0.519	0.340	0.090	0.063	**0.051**
		TES	0.555	0.570	0.644	0.565	0.239	0.022
	0.16 (40)	p-uniform	**0.748**	**0.620**	**0.520**	**0.184**	**0.092**	**0.050**
		TES	0.338	0.245	0.185	0.065	0.029	0.006
	0.33 (16)	p-uniform	**0.365**	**0.342**	**0.319**	**0.182**	**0.100**	**0.043**
		TES	0.074	0.068	0.061	0.023	0.005	0.002
	0.5 (10)	p-uniform	**0.033**	**0.032**	**0.031**	**0.024**	**0.019**	**0.012**
		TES	0.001	0.001	0.001	0.001	0.002	0.003

5. Results: Effect size estimation

The fixed-effect meta-analysis and the trim-and-fill method overestimate μ in case of publication bias while μ is only slightly underestimated by p-uniform.



5. Results: Test of no effect

The Type-I error rate of p-uniform is exactly 0.05 while the Type-I error rates of the other methods is way too high. Statistical power of p-uniform is reasonable when $\mu \geq 0.33$ and power of the other methods is deceivingly high.

Table: Results of simulations (10,000 replications) on Type-I error rates and statistical power

			Pub. bias					
			0	1/40	1/20	1/4	1/2	1
μ (K)	0 (160)	Fixed-effect model	1.000	0.985	0.952	0.566	0.249	0.053
		Trim-and-fill	1.000	0.978	0.939	0.524	0.208	0.035
		p-uniform						**0.050**
	0.16 (40)	Fixed-effect model	1.000	1.000	1.000	0.998	0.999	0.999
		Trim-and-fill	1.000	1.000	0.999	0.996	0.996	0.990
		p-uniform						**0.259**
	0.33 (16)	Fixed-effect model	1.000	1.000	1.000	1.000	1.000	1.000
		Trim-and-fill	1.000	1.000	1.000	1.000	1.000	1.000
		p-uniform						**0.722**
	0.5 (10)	Fixed-effect model	1.000	1.000	1.000	1.000	1.000	1.000
		Trim-and-fill	1.000	1.000	1.000	1.000	1.000	1.000
		p-uniform						**0.980**

6. Conclusion

- p-uniform outperforms other techniques in case of publication bias

Limitation:

- p-uniform can only be used as a sensitivity analysis when effects are heterogeneous

Future research:

- Building a web application where applied researchers can use p-uniform
- Examining the effect of questionable research practices on meta-analysis

NWO Netherlands Organisation for Scientific Research
Tilburg University