***Test for Excess Significance.*** Another power-based test for publication bias is the Test for Excess Significance (Ioannidis & Trikalinos, 2007). This test estimates the number of expected studies with statistical significance given some anticipated effect size (usually the naïve meta-analytic estimate), then compares that expectation against the number of observed significant results. A significant test suggests censoring of nonsignificant results or the manipulation of results into statistical significance.

This test has a number of weaknesses. It has poor statistical power (Ioannidis & Trikalinos, 2007), and the validity of its *p*-value rests on strong, perhaps unwarranted assumptions about researcher behavior (Morey, 2013). Like other tests for bias, its results may be spurious when there exists genuine between-study heterogeneity. The reader is urged to interpret the result with considerable caution.

[The results are…]

***Multiplicative-error meta-regression models.*** Random-effects meta-analysis models heterogeneity across studies through the use of an additive error term τ2. An alternative strategy for handling heterogeneity in meta-regression is instead to fit models with a multiplicative error term φ. (see, e.g., Moreno et al., 2009).

We fit additional multiplicative-error models. [The results are…]

**Influence of excluded studies.** One might be curious as to which is responsible for our different conclusions: the novel bias-adjustment techniques or the exclusion of datapoints? [Blah blah blah]

**Sensitivity analyses.** Some data points appeared to be outliers but their inclusion or exclusion did not have sufficient influence on the results to bear mention in the main report. We report these here. [Blah blah blah]