The Effect of Screening for Publication Bias on the Outcomes of Meta-Analyses.

Haben Michael[1]

[1]University of Massachusetts, Amherst, MA (hmichael@math.umass.edu)

Summary: Conducting a meta-analysis on a body of studies subject to publication bias is a type of post-selection inference that may invalidate findings. Therefore, analysts often run a hypothesis test to check for publication bias prior to conducting a meta-analysis. However, conducting meta-analyses conditional on the outcome of such preliminary tests is itself a form of post-selection inference. We investigate the effect of conducting meta-analyses conditional on a null finding at the preliminary stage. We find that in many situations there is no or little bias in the findings at the main stage.
Keywords: Meta-analysis; Publication bias; Post-selection inference.

1 Introduction

Meta-analysis is a popular technique for summarizing a body of studies. Key to the soundness of the results of a meta-analysis is that the subset of studies used in forming the summary be representative of all the studies conducted. This requirement may fail to be met when publication bias is present, that is, when the availability of a study is tied to its findings. Several hypothesis tests have been proposed with the goal of alerting an analyst to the presence of publication bias in a body of studies before they are used to carry out a meta-analysis.

Two issues present themselves by this type of procedure, in which a preliminary hypothesis test is used to screen data as suitable for a subsequent main analysis. The first is that failing to reject the null in the preliminary stage is treated as a basis for proceeding as though the null were true. Therefore the power of the preliminary hypothesis test requires investigation (Michael and Ghebremichael, b,a).

The second issue is that screening may affect inference in the main analysis. Such biases have been observed widely for unadjusted post-selection inference in general, and screening tests in particular. Whether publication bias is present or not, a body of studies that passes a screening test may differ from one that fails in a way that bears on the subsequent meta-analysis. It would be unfortunate for a test of publication bias to itself bias the outcome of a meta-analysis.

This paper will consider the effect of screening meta-analyses for publication bias. To do so we look for dependency between the test statistics of the preliminary screening test and main test. Two tests for publication bias, Egger's test and Begg's test, will be considered, while the meta-analysis test statistic will be the simple fixed effects summary estimate. As analysts typically do not proceed to the second stage on a finding of significance at the preliminary stage, we focus on data conditional on a null result at the screening stage. We further restrict our focus to true nulls, i.e., studies unaffected by publication bias. The question investigated is: In a world without publication bias, what is the effect of applying versus not applying publication bias tests on the outcome of meta-analyses?

We find that in many situations, a validly applied screening test does not itself bias the outcome of a subsequent meta-analysis. We find the possibility of strong bias only in the case of Begg's test with certain non-gaussian data. In this case, the main consequence is a loss of power in the meta-analysis. Moreover, the data in which this issue arises is arguably unlikely to be encountered in practice, though we do not examine this empirical question any further.

Previous work. There is an extensive literature on "post-selection inference," i.e., inference on data under models chosen using the same data. Taylor and Tibshirani (2015) gives an overview of recent work . In the context of meta-analysis, post-selection inference usually centers on publication bias, i.e., the issue which publication bias tests are designed to address. The collection Rothstein et al. (2005) gives a comprehensive overview of publication bias in meta-analysis. We are not aware of literature discussing the post-selection effects of screening by the result of publication bias tests. Screening is a particularly simple form of selection. An early study of the effect of a preliminary screening test is Olshen (1973), which found that applying Scheffé's method to form intervals for regression coefficients conditionally on rejection by a preliminary F-test decreases the coverage rate relative to an unconditional procedure. More recently, Schucany and Tony Ng (2006) and Rochon et al. (2012) observe that screening for normality using the Shapiro-Wilks test may lead to inflated Type 1 error rates or loss of power, depending on the data.

Organization of the remainder of the paper. In Section 2 we introduce the meta-analysis test statistic and the two publication bias tests used for the preliminary analysis. In Section 3 we consider the effect of screening with finite sample sizes. The main result is that under a gaussian assumption screening does not affect the main analysis. In Section 4 we drop the gaussian assumption and consider the asymptotic effects of screening. While Egger's test does not have any asymptotic effect under many conditions, Begg's test may. In Section 5 we illustrate the theoretical results using synthetic data, and in 6 we conclude and offer future directions.

2 Background

2.1 Meta-analysis model

The data is modeled as pairs $(Y_1, \sigma_1), \ldots, (Y_n, \sigma_n)$ representing the estimated effect sizes and sampling variances of n studies with a common mean effect size $\theta \in \mathbb{R}$. The study effects are assumed to be mutually independent conditionally on the sampling variances:

$$\begin{aligned}(Y_1, \sigma_1), \ldots, (Y_n, \sigma_n) &\text{ independent} \\ \mathrm{E}(Y_1, \ldots, Y_n \mid \sigma_1, \ldots, \sigma_n) &= \theta \mathbb{1}_n \\ \mathrm{Var}(Y_1, \ldots, Y_n \mid \sigma_1, \ldots, \sigma_n) &= \mathrm{diag}(\sigma_1^2, \ldots, \sigma_n^2).\end{aligned} \tag{1}$$

The study variances σ_j^2 are usually treated as fixed, with analyses carried out conditionally, but we will also consider random $\vec{\sigma}$ in the formulation of certain results (Michael and Ghebremichael, b; Lin and Chu, 2018). The typical number of studies, n, depends on the area of research and can be small (Davey et al., 2011). Model (1) is known as the "fixed-effects" meta-analysis model to distinguish it from models in which the common effect θ is treated as random.

The goal of inference of a meta-analysis is the mean effect size θ. The most common estimator is the weighted sample average of the study effects, with weights given by the study precisions $1/\sigma_1^2, \ldots, 1/\sigma_n^2$,

$$\hat{\theta} = \frac{\sum_{j=1}^n Y_j/\sigma_j^2}{\sum_{j=1}^n 1/\sigma_j^2}. \tag{2}$$

Conditionally on $\vec{\sigma}$, the estimator $\hat{\theta}$ is an unbiased estimator of θ with variance $\sigma_{\hat{\theta}}^2 = \mathrm{Var}(\hat{\theta} \mid \vec{\sigma}) = (\sum_j 1/\sigma_j^2)^{-1}$. The usual procedure is to refer $(\hat{\theta} - \theta)/\sigma_{\hat{\theta}}$ to a standard normal distribution in carrying out inference (Konstantopoulos and Hedges, 2019). Sufficient conditions for asymptotic normality are given as part of Theorem 3 below.

2.2 Description of Egger's and Begg's tests for publication bias

Egger's and Begg's tests both test for the presence of publication bias on the basis of the relationship between reported study effects and variances. The premise appears to be that the net effect of selective publication will be a trend between corresponding effects and variances, such as by publication favoring a very precise estimate of an unremarkable effect, or an imprecise estimate of a remarkable effect, or by some other means.

Egger's procedure tests the null of a zero constant coefficient in the simple linear regression of Y/σ against $1/\sigma$. Under model (1), $E(Y_j/\sigma_j \mid \sigma_j) = \theta/\sigma_j$ and $\mathrm{Var}(Y_j/\sigma_j) = 1$. Therefore,

$$y_j/\sigma_j = \beta_0 + \beta_1/\sigma_j + \epsilon \tag{3}$$

is a correctly specified linear model with independent homoskedastic errors ϵ and $\beta_0 = 0$. The t-statistic for β_0,

$$\begin{aligned}\hat{t} &= \frac{\hat{\beta}_0}{\sqrt{\hat{\mathrm{Var}}(\hat{\beta}_0)}} \\ &= \sqrt{\frac{n-1}{RSS}} n^{-1/2} \frac{1}{\sqrt{m_2(m_2 - m_1^2)}} \sum_j Y_i/\sigma_i(m_2 - m_1/\sigma_i),\end{aligned} \tag{4}$$

$$\text{where } m_k = \sum_{i=1}^n 1/\sigma_i^k$$

may therefore serve as a consistent test statistic. The null of no publication bias is rejected when $|\hat{t}| > t_{n-2,1-\alpha/2}$.

Begg's procedure tests the null that Y_j is uncorrelated with σ_j, $j = 1, \ldots, n$. The test statistic is Kendall's rank correlation coefficient,

$$\hat{\tau} = \binom{n}{2}^{-1} \sum_{j<k} 2\left\{(u_j - u_k)(v_j - v_k) > 0\right\} - 1, \tag{5}$$

applied to the sequence of pairs (u_j, v_j) given by

$$(u_j, v_j) = \left(\frac{Y_j - \hat{\theta}}{\sqrt{\sigma_j^2 - \sigma_{\hat{\theta}}^2}}, \sigma_j \right), j = 1, \ldots, n.$$

The test statistic counts the number of corresponding pairs of studentized effect sizes $u_j = (Y_j - \hat{\theta})/\sqrt{\sigma_j^2 - \sigma_{\hat{\theta}}^2}$ and variances $v_j = \sigma_j$ that concord in the sense that either $u_j < u_k$ and $v_j < v_k$ or $u_j > u_k$ and $v_j > v_k$. The null of no correlation is to be interpreted as no publication bias, and is rejected at level α when $\sqrt{9n/4}|\hat{\tau}| > \Phi^{-1}(1 - \alpha/2)$.

2.3 Location invariance of publications bias tests

There seems little reason to think that the location of the grand mean of the data under analysis, θ in model (1), is relevant to an assessment of the presence or absence of publication bias. Therefore, a desirable property of hypothesis tests for publication bias is invariance to location shifts of θ. Both Egger's and Begg's tests satisfy the property. Shifting $\vec{Y}$ by $\theta' \in \mathbb{R}$, $\vec{Y} \mapsto \vec{Y} + \theta'\mathbb{1}$, the sum in Egger's statistic (4) is

$$\begin{aligned}\sum_i (Y_i + \theta')/\sigma_i(m_2 - m_1/\sigma_i) &= \sum_i Y_i/\sigma_i(m_2 - m_1/\sigma_i) + \theta' \sum_i 1/\sigma_i(m_2 - m_1/\sigma_i) \\ &= \sum_i Y_i/\sigma_i(m_2 - m_1/\sigma_i) + \theta'(nm_1m_2 - nm_2m_1) \\ &= \sum_i Y_i/\sigma_i(m_2 - m_1/\sigma_i).\end{aligned}$$

The RSS is also unchanged . Likewise, the Begg statistic (5) depends on Y_i only through the differences $Y_i - \hat{\theta}$, which cancel out any shift. It is therefore unsurprising that the test statistics for such hypothesis tests should not share much dependency with the meta-analysis test statistic, which targets θ. (But see Macaskill et al. (2001) who claim based on simulations an effect of location on power.) The next sections show that the publication bias and meta-analysis test statistics are indeed often independent or nearly so, with the result that screening does not bias the outcome of meta-analyses.

3 Finite-sample, gaussian effect sizes

The study effects in (1) are often modeled as gaussian by appealing to the CLT, e.g., in the original paper Begg and Mazumdar (1994). In this situation, the test statistics for

Egger's and Begg's tests are conditionally independent of the meta-analysis test statistic given the primary study variances. It follows that the publication bias test statistics are marginally orthogonal to the meta-analysis test statistic. There is therefore no harm of bias in screening under conditional analyses with even a small number n of studies, provided the gaussian assumption holds, besides the other meta-analysis and publication bias test assumptions.

A simple connection between the publication bias test statistics and the meta-analysis test statistics illustrates this result. The meta-analysis test statistic (2), $\hat{\theta} = (\sum Y_j/\sigma_j^2)/(\sum \sigma_j^2)$, may be viewed as the coefficient of the regression of $\frac{\vec{Y}}{\sigma}$ on $\frac{\vec{1}}{\sigma}$. Therefore, $\frac{\vec{Y}}{\sigma} - \hat{\theta}\frac{\vec{1}}{\sigma}$ is orthogonal to $\hat{\theta}\frac{\vec{1}}{\sigma}$, and then, conditionally on $\vec{\sigma}$, orthogonal to $\hat{\theta}$. Since the Begg statistic is a function of $\frac{\vec{Y}}{\sigma} - \hat{\theta}\frac{\vec{1}}{\sigma} = (\frac{Y_1-\hat{\theta}}{\sigma_1}, \ldots, \frac{Y_n-\hat{\theta}}{\sigma_n})$, it too is orthogonal to $\hat{\theta}$ given $\vec{\sigma}$, when $\vec{Y}$ is gaussian.

As for Egger's test, rewriting the Egger regression $\frac{\vec{Y}}{\sigma} = \hat{\beta}_0\mathbb{1} + \hat{\beta}_1\frac{\vec{1}}{\sigma}$ as

$$\frac{\vec{Y}}{\sigma} - \hat{\theta}\frac{\vec{1}}{\sigma} = \hat{\beta}_0\mathbb{1} + (\hat{\beta}_1 - \hat{\theta})\frac{\vec{1}}{\sigma},$$

it follows that the coefficient $\hat{\beta}_0$ used in Egger's test may be obtained from the regression of $\frac{\vec{Y}}{\sigma} - \hat{\theta}\frac{\vec{1}}{\sigma}$ on $\left(\mathbb{1}, \frac{\vec{1}}{\sigma}\right)$. Therefore, $\hat{\beta}_0$ is a function of $\frac{\vec{Y}}{\sigma} - \hat{\theta}\frac{\vec{1}}{\sigma}$ which is orthogonal to $\hat{\theta}\frac{\vec{1}}{\sigma}$ and so to $\hat{\theta}$, conditionally on $\vec{\sigma}$. See Fig. 1. Theorem 1 formalizes this argument, using the gaussian assumption to convert the orthogonality to independence.

Theorem 1. *Assume*

$$(Y_1, \ldots, Y_n) \mid (\sigma_1, \ldots, \sigma_n) \sim \mathcal{N}(\theta\mathbb{1}, \operatorname{diag}(\sigma_1^2, \ldots, \sigma_n^2)).$$

Then 1. $\hat{t} \perp\!\!\!\perp \hat{\theta} \mid \sigma_1, \ldots, \sigma_n$ *and 2.* $\hat{\tau} \perp\!\!\!\perp \hat{\theta} \mid \sigma_1, \ldots, \sigma_n$.

Proof. 1. First, for gaussian $\vec{Y}$, the residual sum of squares in the Egger regression, RSS, is independent of $\hat{\beta}_0$. Second, RSS is also independent of $\hat{\theta}$ in light of the remarks preceding the theorem, since $\frac{\vec{1}}{\sigma}$ is in the column space of the Egger regression. Third, $\mathrm{E}(\hat{t} \mid \vec{\sigma}) = \mathrm{E}\left(1/\sqrt{\operatorname{Var}(\hat{\beta}_0)} \mid \vec{\sigma}\right)\mathrm{E}(\hat{\beta}_0 \mid \vec{\sigma}) = 0$ by OLS theory, since the linear model (3) is correct. Therefore,

$$\begin{aligned}
\operatorname{Cov}(\hat{t}, \hat{\theta} \mid \vec{\sigma}) &= \mathrm{E}(\hat{t}\hat{\theta} \mid \vec{\sigma}) \\
&= \mathrm{E}\left(\left.\frac{\hat{\beta}_0\hat{\theta}}{\sqrt{\operatorname{Var}(\hat{\beta}_0)}}\right| \vec{\sigma}\right) \\
&= \mathrm{E}\left(\left.\frac{1}{\sqrt{\operatorname{Var}(\hat{\beta}_0)}}\right| \vec{\sigma}\right)\mathrm{E}\left(\hat{\beta}_0\hat{\theta}\middle| \vec{\sigma}\right) \\
&= \mathrm{E}\left(\left.\frac{1}{\sqrt{\operatorname{Var}(\hat{\beta}_0)}}\right| \vec{\sigma}\right)\frac{1}{n(m_2 - m_1^2)}\frac{1}{m_2}\mathrm{E}\left(\left(\sum_j Y_j/\sigma_j(m_2 - m_1/\sigma_j)\right)\sum_j Y_j/\sigma_j^2\middle| \vec{\sigma}\right),
\end{aligned}$$

while the second expectation is

$$\mathrm{E}\left(\left(\sum_j Y_j/\sigma_j(m_2 - m_1/\sigma_j)\right)\sum_j Y_j/\sigma_j^2 \middle| \vec{\sigma}\right) = \mathrm{Cov}\left(\sum_j Y_j/\sigma_j(m_2 - m_1/\sigma_j), \sum_j Y_j/\sigma_j^2 \middle| \vec{\sigma}\right)$$

$$= \sum_j \mathrm{Cov}(Y_j/\sigma_j(m_2 - m_1/\sigma_j), Y_j/\sigma_j^2 \mid \vec{\sigma})$$

$$= \sum_j 1/\sigma_j(m_2 - m_1/\sigma_j) = 0.$$

2. As noted in the remarks preceding the theorem statement, $\mathrm{Cov}(\frac{\vec{Y}}{\sigma} - \hat{\theta}\frac{\vec{1}}{\sigma}, \frac{\vec{1}}{\sigma}) = 0$. For any j, then, $\mathrm{Cov}(Y_j - \hat{\theta}, \hat{\theta} \mid \vec{\sigma}) = \sigma_j^2 \,\mathrm{Cov}(Y_j/\sigma_j - \hat{\theta}/\sigma_j, \hat{\theta}/\sigma_j \mid \vec{\sigma}) = 0$. Explicitly,

$$\begin{aligned}\mathrm{Cov}(Y_j - \hat{\theta}, \hat{\theta} \mid \vec{\sigma}) &= \mathrm{Cov}(Y_j, \hat{\theta} \mid \vec{\sigma}) - \mathrm{Var}(\hat{\theta} \mid \vec{\sigma}) \\ &= \mathrm{Cov}\left(Y_j, \frac{Y_j/\sigma_j^2}{nm_2} \middle| \vec{\sigma}\right) - \mathrm{Var}\left(\frac{Y_j/\sigma_j^2}{nm_2} \middle| \vec{\sigma}\right) \\ &= \frac{1}{nm_2} - \frac{1}{nm_2} = 0.\end{aligned}$$

Since the data is gaussian, $(y_1 - \hat{\theta}, \ldots, y_n - \hat{\theta})$ is independent of $\hat{\theta}$ given $\vec{\sigma}$. Since $\hat{\tau}$ is a function of $(y_1 - \hat{\theta}, \ldots, y_n - \hat{\theta})$ given $\vec{\sigma}$, it too is conditionally independent of $\hat{\theta}$.

□

Corollary 2. *Assuming*

$$(Y_1, \ldots, Y_n) \mid (\sigma_1, \ldots, \sigma_n) \sim \mathcal{N}(\theta\mathbb{1}, \mathrm{diag}(\sigma_1^2, \ldots, \sigma_n^2)).$$

$\mathrm{Cov}(\hat{t}, \hat{\theta}) = \mathrm{Cov}(\hat{\tau}, \hat{\theta}) = 0.$

Proof. $\mathrm{Cov}(\hat{t}, \hat{\theta}) = \mathrm{E}\,\mathrm{Cov}(\hat{t}, \hat{\theta} \mid \vec{\sigma}) + \mathrm{Cov}\left(\mathrm{E}\left(\hat{t} \mid \vec{\sigma}\right), \mathrm{E}\left(\hat{\theta} \mid \vec{\sigma}\right)\right)$. The first term is 0 by Theorem 1 and the second since $\mathrm{E}\left(\hat{\theta} \middle| \vec{\sigma}\right) = \theta$ is constant. □

In showing that the Egger statistic is independent, normality is only used to obtain independence of the residual errors from the coefficient estimate. Writing

$$\mathrm{Cov}(\hat{t}, \sqrt{n}\hat{\theta}) = \mathrm{Cov}\left(\hat{t}\sqrt{\frac{RSS}{n-2}}, \sqrt{n}\hat{\theta}\right) + \mathrm{Cov}\left(\hat{t}\left(1 - \sqrt{\frac{RSS}{n-2}}\right), \sqrt{n}\hat{\theta}\right),$$

the first term is shown to be 0 in the first part of the proof, and the second is ordinarily of order $n^{-1/2}$. Therefore, under the general model (1), any correlation between the Egger test statistic and meta-analysis test statistic vanishes as the number of studies grows. In contrast, without the gaussian assumption Begg's test statistic and the meta-analysis test statistic may be correlated even in the limit, as Theorem 5 below shows.

4 Asymptotic, non-parametric study effects

We extend the analysis beyond gaussian study effects by considering the association between the test statistics as the number of studies n grows.

Theorem 3. *Assuming:*

1. $(Y_1, \ldots, Y_n) \mid (\sigma_1, \ldots, \sigma_n)$ *are independent with means* 0 *and variances* $\sigma_1^2, \ldots, \sigma_n^2$
2. *Existence and finiteness of* $\mu_1 = \lim_{n\to\infty} m_1 = \lim_{n\to\infty} n^{-1} \sum_{j=1}^n 1/\sigma^j$

For some $\delta > 0$,

3. $\sup_j \mathrm{E}\left(\frac{Y_j-\theta}{\sigma_j}\right)^{2+\delta} < \infty$
4. $m_2 - m_1^2 > \delta$
5. $\sum_j (\sigma_j j)^{-1-\delta} < \infty$

Then

$$\left.\left(\hat{t}, \frac{\hat{\theta}-\theta}{\sigma_{\hat{\theta}}}\right)\right| \sigma_1, \ldots, \sigma_n \rightsquigarrow \mathcal{N}(0, I).$$

Proof. 1. Use the Cramer-Wold device to show

$$\left(\sqrt{\frac{RSS}{n-1}}\hat{t}, \frac{\hat{\theta}-\theta}{\sigma_{\hat{\theta}}}\right) \rightsquigarrow N(0, I) \tag{6}$$

then 2. show $RSS/(n-1) \to_p 1$.

1. The random variables in the statement, re-written in terms of the study precisions $S_1, \ldots, S_n$, are

$$\begin{aligned}
\hat{t} &= \sqrt{\frac{n-1}{RSS}} n^{-1/2} \frac{1}{\sqrt{m_2(m_2-m_1^2)}} \sum_j Y_j S_j (m_2 - m_1 S_j) \\
&= \sqrt{\frac{n-1}{RSS}} n^{-1/2} \frac{1}{\sqrt{m_2(m_2-m_1^2)}} \sum_j (Y_j-\theta) S_j (m_2 - m_1 S_j), \\
\frac{\hat{\theta}-\theta}{\sigma_{\hat{\theta}}} &= \frac{1}{\sqrt{m_2}} \sum_j (Y_j - \theta) S_j^2.
\end{aligned}$$

Given $a, b \in \mathbb{R}$, the linear combination

$$a\sqrt{\frac{RSS}{n-1}}\hat{t} + b\frac{\hat{\theta}-\theta}{\sigma_{\hat{\theta}}} = n^{-1/2} \sum_{j=1}^n \frac{S_j^2}{\sqrt{m_2}} \left(\frac{a(m_2/S_j - m_1)}{\sqrt{m_2(m_2-m_1^2)}} + b\right)(Y_j - \theta)$$

has conditional mean and variance

$$\mathrm{E}\left(a\sqrt{\frac{RSS}{n-1}}\hat{t}+b\frac{\hat{\theta}-\theta}{\sigma_{\hat{\theta}}}\mid\vec{\sigma}\right)=0$$
$$\mathrm{Var}\left(a\sqrt{\frac{RSS}{n-1}}\hat{t}+b\frac{\hat{\theta}-\theta}{\sigma_{\hat{\theta}}}\mid\vec{\sigma}\right)=a^2+b^2.$$

It suffices to show that the linear combination is asymptotically normal, as (6) then follows by the Cramer-Wold device. By the Lyapunov CLT, asymptotic normality follows if for some $\delta>0$, given $\vec{S}$,

$$\begin{aligned}&\sum_{j=1}^{n}\mathrm{E}\left|n^{-1/2}\frac{S_j^2}{\sqrt{m_2}}\left(\frac{a(m_2/S_j-m_1)}{\sqrt{m_2(m_2-m_1^2)}}+b\right)(Y_j-\theta)\right|^{2+\delta}\\&=\sum_{j=1}^{n}\left(n^{-1/2}\frac{S_j}{\sqrt{m_2}}\left(\frac{a(m_2/S_j-m_1)}{\sqrt{m_2(m_2-m_1^2)}}+b\right)\right)^{2+\delta}\mathrm{E}\left|\frac{Y_j-\theta}{\sigma_j}\right|^{2+\delta}\end{aligned}$$

converges to 0. First, the terms $\mathrm{E}\left|\frac{Y_j-\theta}{\sigma_j}\right|^{2+\delta}$ are assumed bounded. Second, $b\frac{S_j}{\sqrt{m_2}}\leq b$. Third,

$$\frac{m_2-m_2S_j}{m_2\sqrt{m_2-m_1^2}}=\frac{1-\frac{m_1}{m_2}S_j}{\sqrt{m_2-m_1^2}}\leq\frac{1}{\sqrt{m_2-m_1^2}}.$$

Therefore,

$$\begin{aligned}\sum_{j=1}^{n}\left(n^{-1/2}\frac{S_j}{\sqrt{m_2}}\left(\frac{a(m_2/S_j-m_1)}{\sqrt{m_2(m_2-m_1^2)}}+b\right)\right)^{2+\delta}&\leq\sum_{j=1}^{n}\left(n^{-1/2}\left(b+\frac{1}{\sqrt{m_2-m_1^2}}\right)\right)^{2+\delta}\\&\leq\sum_{j=1}^{n}(n^{-1/2}(b+1/\delta))^{2+\delta}\\&=(b+1/\delta)^{2+\delta}n^{-\delta/2}\to 0.\end{aligned}$$

2. In terms of the precisions,

$$\frac{RSS}{n}=\overline{Y^2S^2}-(\overline{YS})^2-\frac{(\overline{YS^2}-m_1\overline{YS})^2}{m_2-m_1^2}.$$

Let $\mu_k=\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}S_j^k$ for $k=1,2$. Show $\frac{RSS}{n}\to_p 1$ by showing a. $\overline{Y^2S^2}\to_p 1+\theta^2\mu_2$, b. $\overline{YS}\to_p\theta\mu_1$, and c. $\overline{YS^2}\to_p\theta\mu_2$. In all 3 cases we use a variant of the LLN (e.g., Chapter 2, Ex. 4.5, of Rao (1973)): Given independent integrable RVs $X_1,X_2,\ldots$, and $\delta>0$ such that $\sum_j\mathrm{E}\left|(X_j-\mathrm{E}\,X_j)/j\right|^{1+\delta<\infty}$, conclude $\left|\frac{1}{n}\sum_{j=1}^{n}X_j-\frac{1}{n}\sum_{j=1}^{n}\mathrm{E}\,X_j\right|\to_{a.s.}0$.

a.

$$\mathrm{E}\left((Y_jS_j)^2 - (1+\theta^2 S_j^2)\right)^{1+\delta} = \mathrm{E}\left(((Y_j-\theta)S_j)^2 + 2\theta(Y_j-\theta)S_j^2 - 1\right)^{1+\delta}$$
$$\leq 2^{1+\delta}\left(\mathrm{E}\left(\frac{Y_j-\theta}{\sigma_j}\right)^{2(1+\delta)} + (2\theta S_j)^{1+\delta}\,\mathrm{E}\left(\frac{Y_j-\theta}{\sigma_j}\right)^{1+\delta} + 1\right).$$

As $\sup_j \mathrm{E}\left(\frac{Y_j-\theta}{\sigma_j}\right)^{2+\delta}$ is assumed finite for some $\delta > 0$, the same holds for $\sup_j \mathrm{E}\left(\frac{Y_j-\theta}{\sigma_j}\right)^{1+\delta}$, and it is assumed that $\sum_j (S_j/j)^{1+\delta} < \infty$.

b. With $\delta = 2$, $\sum_j \mathrm{E}\,|(y_jS_j - \theta\mu_1)/j|^{1+\delta} = \sum_j \mathrm{Var}(Y_jS_j) = \sum_j 1/j^2 < \infty$, so $\overline{ys} \to_{a.s.} \theta\mu_1$.

c.

$$\sum_j \mathrm{E}\left|(Y_jS_j^2 - \theta S_j^2)/j\right|^{1+\delta} = \sum_j (S_j/j)^{1+\delta}\,\mathrm{E}\left|\frac{Y_j-\theta}{\sigma_j}\right|^{1+\delta} < \infty$$

as before.

□

Many IID data models meet the conditions of the theorem.

Corollary 4. *Assuming: 1.* $(Y_1,\ldots,Y_n) \mid (\sigma_1,\ldots,\sigma_n)$ *are independent with means* 0 *and variances* $\sigma_1^2,\ldots,\sigma_n^2$*. 2.* $((Y_1-\theta)/\sigma_1,\ldots,(Y_n-\theta)/\sigma_n) \mid (\sigma_1,\ldots,\sigma_n)$ *follow a common distribution such that for some* $\delta > 0$, $\mathrm{E}((Y-\theta)/\sigma \mid \sigma)^{2+\delta} < \infty$*. 3. The precisions* $S_1, S_2, \ldots$ *are IID, non-constant, with a finite second moment. Then almost surely,*

$$\left.\left(\hat{t}, \frac{\hat{\theta}-\theta}{\sigma_{\hat{\theta}}}\right)\right| \sigma_1,\sigma_2,\ldots,\sigma_n \rightsquigarrow N(0,I).$$

Proof. $m_2 - m_1^2 \to \mathrm{Var}(S)$ almost surely, with $0 < \mathrm{Var}(S) < \infty$, so almost surely there is a random $\delta > 0$ such that $m_2 - m_1^2 > \delta$. By a truncation argument (e.g., Ex. 2.3.25 of Dembo (2016)) the assumed moment condition on S likewise implies $\sum_j (S_j/j)^{1+\delta}$ converges almost surely for any $1 < \delta \leq 2$. □

The corresponding analysis for Begg's statistic is relatively complicated due to the statistics $\sigma_{\hat{\theta}}$ and $\hat{\theta}$ common to the terms of the double sum (5). We impose several simplifying assumptions. First, as noted in Section 2.3, $\hat{\tau}$ is invariant to shifting the location of the study effects, $\vec{Y} \mapsto \vec{Y} + \theta$, so there is no loss in assuming $\theta = 0$:

$$\mathrm{E}(Y_1,\ldots,Y_n \mid \sigma_1,\ldots,\sigma_n) = 0. \tag{7}$$

Second, the $O(1/n)$ terms $\sigma_{\hat{\theta}}$ are asymptotically negligible in many common situations (Michael and Ghebremichael, b),

$$\hat{\tau} = \binom{n}{2}^{-1} \sum_{j<k} 2\left\{\left(\frac{Y_j-\hat{\theta}}{\sigma_j} - \frac{Y_k-\hat{\theta}}{\sigma_k}\right)(\sigma_j - \sigma_k) > 0\right\} - 1 + o_P(n^{-1/2}). \tag{8}$$

Next, let $Z_j = Y_j/\sigma_j, j = 1, \dots, n$, denote the standardized effect sizes. From (1), $\mathrm{Var}(Z) = 1$ and by (7), $\mathrm{E}(Z) = 0$. We further assume that $Z_1, \dots, Z_n$, follow a common distribution, i.e., the study effects $Y_1, \dots, Y_n$, belong to a scale family. Finally, we assume the study precisions $(S_1, \dots, S_n)$ are IID. In summary,

$$\begin{aligned} Z_1, \dots, Z_n &\overset{IID}{\sim} F_Z \\ S_1, \dots, S_n &\overset{IID}{\sim} F_S \\ Z_j \mid S_j &\sim Z_j, \\ Y_j &= Z_j/S_j, j = 1, \dots, n. \end{aligned} \tag{9}$$

Rewritten in terms of the standardized effect sizes and their precisions, (8) is

$$\hat{\tau} = \binom{n}{2}^{-1} \sum_{j<k} 2 \left\{ \frac{Z_j - Z_k}{S_j - S_k} < \hat{\theta} \right\} - 1 + o_P(n^{-1/2}). \tag{10}$$

The double sum is a U-statistic with estimated parameter $\hat{\theta}$ (Nolan and Pollard, 1988). The estimate is ordinarily of order $1/\sqrt{n}$ and affects the asymptotics, and Begg's test, ignoring this effect, can be biased (Michael and Ghebremichael, b). The following result includes the correct asymptotic distribution under the IID model (9).

Theorem 5. *Assuming:*

1. *the IID model* (9)
2. $0 < \mathrm{E}\, S^2 < \infty$
3. $\int_{-\infty}^{\infty} f_Z(z)^2 dz < \infty$

Then

$$\begin{pmatrix} \sqrt{n}\hat{\tau} \\ (\hat{\theta} - \theta)/\sigma_{\hat{\theta}} \end{pmatrix}$$

is asymptotically normal with

$$\begin{aligned} \mathrm{Var}(\sqrt{n}\hat{\tau}) &\to \frac{4}{9} + \frac{4\,(\mathrm{E}\,|S - S'|)^2}{\mathrm{E}\, S^2} \mathrm{E}(f_Z(Z))\,(\mathrm{E}(f_Z(Z)) - 2\,\mathrm{E}(ZF_Z(Z)))\,, \\ \mathrm{Cov}(\sqrt{n}\hat{\tau}, (\hat{\theta} - \theta)/\sigma_{\hat{\theta}}) &\to \frac{2\,\mathrm{E}\,|S - S'|}{\sqrt{\mathrm{E}\, S^2}} (\mathrm{E}\, f_Z(Z) - \mathrm{E}(ZF_Z(Z)))\,, \\ \mathrm{Var}((\hat{\theta} - \theta)/\sigma_{\hat{\theta}}) &\to 1. \end{aligned}$$

The theorem follows from a lemma given in Michael and Ghebremichael (a) that, drawing on the theory of U-processes (Nolan and Pollard, 1988), rewrites (10) as an asymptotically equivalent IID sum to which the CLT may be applied. Let

$$\Pi\hat{\tau}(\theta) : \theta \mapsto 2\left(\frac{1}{n} \sum_{j=1}^{n} 2\,\mathrm{P}\left(\frac{Z_j - Z}{S_j - S} < \theta \mid Z_j, S_j \right) - 1 \right) - \left(2\,\mathrm{P}\left(\frac{Z - Z'}{S - S'} < \theta \right) - 1 \right).$$

Lemma 6. *Under the assumptions of Theorem 5,*

$$\sqrt{n}\hat{\tau} = \sqrt{n}\left(\hat{\theta} 2\,\mathrm{E}(f_Z(Z))\,\mathrm{E}\,|S - S'| + \Pi\hat{\tau}(0)\right) + o_P(1)$$

The asymptotic covariance between $\hat{\tau}$ and $\hat{\theta}$ given in Theorem 5 depends on both the distribution of the study precisions, through the parameter

$$\frac{\mathrm{E}\,|S - S'|}{\sqrt{\mathrm{E}\,S^2}}, \tag{11}$$

and that of the study effects, through the parameter

$$\zeta(F_Z) = \mathrm{E}\, f_Z(Z) - \mathrm{E}(Z F_Z(Z)).$$

The ratio (11) approaches 0 as the precision distribution approximates a nonzero constant. A loose upper bound of $\sqrt{2}$ follows from $(\mathrm{E}\,|S - S'|)^2 \leq \mathrm{E}((S - S')^2) = 2\,\mathrm{Var}(S) \leq 2\,\mathrm{E}(S^2)$. Michael and Ghebremichael (b) gives a tight upper bound of $\sqrt{2/3}$.

The parameter $\zeta(F_Z)$ takes a larger range of values and also determines the sign of the correlation, in turn determining whether the power of the subsequently conducted meta-analysis will be too low or too high. For some distributions of Z, ζ is 0. For example, for standard normal Z, $\mathrm{E}\, f_Z(Z) = \mathrm{E}(Z F_Z(Z)) = 1/(2\sqrt{\pi})$, and for centered and scaled uniform Z, $\mathrm{E}\, f_Z(Z) = \mathrm{E}(Z F_Z(Z)) = 1/(2\sqrt{3})$. When the standardized effects follow these distributions, Begg's test does not bias the meta-analysis, under the conditions of Theorem 5.

In general, however, ζ may be arbitrarily large. In terms of the centered but not necessarily scaled study effects $Y = \sigma Z$,

$$\int f_Z(z)^2 = \sigma \int f_Y(y)^2.$$

The expression on the right may blow up due to either factor σ or $f_Y(y)^2$. An example of the former is Student's t. As the degrees of freedom p approach 2 from above, $\sigma \to \infty$, while $\int f_Y^2(y) \propto \int (1 + y^2/p)^{1+p}$ is bounded away from 0. An example of the latter is any unbounded density that diverges faster than $1/\sqrt{x}$, such as the "peaked" distribution,

$$f_Y(y) = |y|^p \text{ on } |y| < \left(\frac{p+1}{2}\right)^{\frac{1}{p+1}}, \quad p > -1. \tag{12}$$

For this density, $\int f_Y^2(y) \propto \int |y|^{2p} \to \infty$ as $p \downarrow -1/2$, while for $p = -1/2$, $\sigma^2 = \int y^2 f_Y(y) = 1/5(1/4)^4$. . Another example is a centered, symmetric beta distribution with common shape parameter p,

$$f_Y(y) = \frac{(1/4 - y^2)^{p-1}}{\mathrm{B}(p, p)} \text{ on } |y| < 1/2, p > 0. \tag{13}$$

Here $\int f_Y^2(y) \propto (1/4 - y^2)^{2(p-1)} \to \infty$ as $p \downarrow 1/2$. While the density (12) is increasingly peaked $\int f_Y^2(y) \to \infty$, with mass moving to the origin, the centered beta (13) is increasingly U-shaped, with mass moving to $\pm 1/2$.

Though $\zeta \to \infty$ due either to $\sigma \to \infty$ or $\mathrm{E}\, f_Y = \int f_Y^2(y) \to \infty$, the two possibilities say different things about the operational characteristics of the publication bias tests. When σ is large, the basic meta-analysis model (1) is nearly violated and there are other difficulties with attempting a meta-analysis. For example, the fixed effects estimator $\hat{\theta}$ may have a poor rate of convergence under the CLT. A large value of $\mathrm{E}\, f_Z = \int f_Z^2$ poses problems specifically for Begg's test. Student's t and the densities (12), (13), are examined using synthetic data in Section 5.

When ζ is large, screening out data for which Begg's test does not give a significant result then affects in kind the positively correlated meta-analysis test statistic, reducing its power. Very negative ζ, which would lead to poor FPR control on the meta-analysis, does not appear to be as significant an issue.

Theorem 7. *Let A denote the set of monotonic differentiable real-valued functions such that $\lim_{z\to-\infty} F(z) = 0, \lim_{z\to\infty} F(z) = 1, \int zF' = 0, \int z^2F' = 1$. Let lower-case f denote the derivative F' for $F \in A$.*

1. *The functional $F \mapsto \int (F'(z))^2 dz = \int f(z)^2$ is convex on the set of differentiable functions with square-integrable derivatives, and is minimized on A by $f(z) \propto 0 \wedge (1 - z^2)$, with value $3/25$.*

2. *The functional $F \mapsto \int zF(z)F'(z)dz = \int zF(z)f(z)dz$ is concave on A and is maximized when f is proportional to a centered uniform distribution, with value $1/(2\sqrt{3})$.*

3. *The functional $F \mapsto \int (F'(z))^2 dz - \int zF(z)F'(z)dz = \zeta(F)$ is convex.*

Proof. 1. For $\lambda \in [0, 1]$,

$$\lambda \int f^2 + (1-\lambda) \int g^2 - \int (\lambda f + (1-\lambda)g)^2 = \lambda(1-\lambda) \int (f-g)^2 \geq 0.$$

For the minimization, see, e.g., Chapter 14, Ex. 8, of Van der Vaart (2000).

2. For $\lambda \in [0, 1]$, $F, G \in A$,

$$\begin{aligned}
&\int z(\lambda F + (1-\lambda)G)(\lambda f + (1-\lambda)g) - \lambda \int zFf - (1-\lambda) \int zGg \\
&= -\lambda(1-\lambda) \int z(f-g)(F-G) \\
&= -\lambda(1-\lambda) \left(\frac{z}{2}(F-G)^2 \Big|_{-\infty}^{\infty} - \frac{1}{2} \int (F-G)^2 \right) \\
&= \frac{1}{2}\lambda(1-\lambda) \int (F-G)^2 \geq 0.
\end{aligned}$$

The variational calculus gives conditions for stationarity of $\int zFF'$ subject to $\int F' = 1, \int zF' = 0, \int z^2F' = 1$. The Lagrangian is $zFF' - \lambda_1 zF' - \lambda_2 z^2 F'$ with Euler-Lagrange equation $F = \lambda_1 + 2\lambda_2 3$, implying a density f proportional to a constant.

3. Convexity follows from the previous parts since the sum of convex functions is convex. □

Theorem 7 implies a loose lower bound

$$\zeta \geq -1/(2\sqrt{3}).$$

This bound is not achieved by the uniform distribution given in the Theorem. As mentioned earlier, $\zeta = 0$ for the centered and scaled uniform.

5 Simulation

We use synthetic data to investigate the effect on a meta-analysis of conditioning data on a non-significant publication bias test outcome. First, a synthetic body of studies $(Y_1, \sigma_1, \ldots, Y_n, \sigma_n)$ is generated under model (1). Next, Egger's test for publication bias,

$$|\hat{t}| > t_{n-2,1-\alpha_0/2}$$

and Begg's test for publication bias

$$\sqrt{9n/4}|\hat{\tau}| > \Phi^{-1}(1 - \alpha_0/2)$$

are carried out at significance level α_0. Finally, the test for a non-zero grand mean in the meta-analysis,

$$\hat{\theta}/\sigma_{\hat{\theta}} > \Phi^{-1}(1 - \alpha_1/2)$$

is carried out at significance level α_1. The process is iterated $10,000$ times, giving a set of $10,000$ triples of rejection indicators. With these we approximate the error rates of the meta-analysis test conditional on not rejecting during the publication bias test.

5.1 Parameters of simulation

In generating the body of studies $(Y_1, \sigma_1, \ldots, Y_n, \sigma_n)$ we considered three families for the distribution of the responses $\vec{Y}$:

1. Student's t distributions with degrees of freedom ranging between $(2, 6)$. When the degrees of freedom are large, the data approaches the gaussian model, in which Theorem 1 asserts the publication bias tests exert no influence on the meta-analysis distribution. When the degrees of freedom approach 2, the variance blows up and the data approach the boundary of the basic meta-analysis model (1).

2. The power law-type distribution (12), with the exponent p in the range $(-1, 0]$. When the exponent is $p = 0$, the distribution is a centered and scaled uniform, for which $\zeta = 0$, with the peakedness about the origin increasing as $p \to -1$. When $p \to -1/2$, $\zeta \to \infty$, posing difficulties for Begg's test according to Theorem 5. For $1 < p \leq .5$, the theorem is inapplicable as $\mathrm{E}\, f_Z$ is not finite, and only perhaps suggestive of the behavior of Begg's test.

3. Symmetric and centered beta distributions (13), with common shape parameter p in the range $[.1, 1]$. When $p = 1$, the distribution is again a centered and scaled uniform. When $p \to 0$, the density becomes increasingly U-shaped. For $p < 1/2$ Theorem 5 is again only suggestive as $\mathrm{E}\, f_Z$ is not finite.

In all cases, the distribution of the standard deviations was uniform on $[1, 4]$, following Lin and Chu (2018). The meta-analysis sample size was $n = 25$ or $n = 75$, following Begg and Mazumdar (1994), who based the choice on literature reviews of the medical and social sciences literature, respectively. The grand mean θ was 0 or .2, with 0 representing the null case in the meta-analysis test. The significance level of the screening publication bias test α_0 was 0.05 or 0.15 following recommendations in Egger et al. (1997), Begg (1994), and elsewhere , while in all cases the level of the meta-analysis test was $\alpha_1 = .05$, a standard recommendation (Konstantopoulos and Hedges, 2019).

5.2 Results of simulation

Main results are given in Tables 1 and 2.

1. Student's t distributions. The power of the meta-analysis is similar whether a publication bias test is used or not. That power is around the nominal rate when the degrees are in the range 4–6 or larger, dropping as the degrees of freedom approach 2. The drop is expected as model (1) does not hold for degrees of freedom ≤ 2. It is unrelated to screening and attributable to the slow CLT convergence of of averages such as $\hat{\theta}$ under these distributions.

2. Power law-type distributions. The power of Begg's test drops as the peakedness increases, while Egger's test is consistent. The behavior is expected from Theorem 5, which shows the dependence of $\mathrm{Cov}(\hat{\tau}, \hat{\theta})$ on $\mathrm{E}\, f_Z$, and Theorem 3, which asserts that $\hat{t}$ and $\hat{\theta}$ are asymptotically independent.

3. Beta distributions. As with the power law-type distribution, the power of Begg's test drops relative to Egger's test as the distribution becomes less uniform and increasingly U-shaped. This result, too, is expected from Theorem 5, as the change in shape corresponds to increasing $\mathrm{E}\, f_Z$.

In the problem area, i.e., Begg's test for very peaked or very U-shaped distributions, Tables 1 and 2 reveal two other contributors to lower power. The first is that the power is poorer as the sample size increases. For example, for the most peaked power law-type distribution in Table 1, the power of Begg's test drops from a reasonable .045 when $n = 25$ to .007 when $n = 75$. In contrast, for Student's t, the power improves with sample size, as the CLT approximation improves. Theorem 5 is qualitative in that it provides no rates of convergence to the asymptotic covariance between $\hat{\tau}$ and $\hat{\theta}$. Evidently the asymptotic covariance overstates the severity, particularly for the relatively small values of n typical of meta-analyses. A more refined asymptotic analysis might offer a more accurate picture of the finite sample behavior.

A second, less dramatic contributor to lower power is a larger nominal level for the screening publication bias test α_0. . This effect is at odds with the frequent recommendation to opt for higher nominal levels in testing for publication bias as a way to boost power (Begg, 1994; Egger et al., 1997; Macaskill et al., 2001). However, the magnitude of a test statistic being inversely related to the p-value, the magnitude of the publication bias test statistic conditional on being $\geq \alpha_0$ tends to be smaller as α_0 is larger. Consequently the positively correlated meta-analysis test statistic also tends to be smaller in magnitude, reducing the power of the meta-analysis test. See Fig. 2.

Both of these "counterintuitive results" emerge in the simulation analysis conducted by Schucany and Tony Ng (2006) investigating the effect of screening for normality with the Shapiro-Wilks test, though it is not clear whether the underlying mechanisms are the same.

6 Conclusion

The recommendation to test for publication bias before conducting a meta-analysis is tentatively supported by many of the results presented here. The exceptions involve screening on Begg's test with nonnormal data, in which case the main effect is a loss of power of the meta-analysis, arguably of less concern than a loss of FPR control. These conclusions of course presuppose validity of the test assumptions, such as the idealized fixed-effects model.

When screening by Begg's test does reduce the power of the subsequent meta-analysis, it appears that larger sample sizes and using a larger nominal level in the publication bias test both aggravate the issue. That the power decreases with sample size is an artifact of our asymptotic theory, which predicts a more severe dependency than exhibited in finite samples. It is possible that with other data-generating mechanisms the reverse trend occurs.

A lower nominal level is generally cautioned against as the two-stage testing procedure involves accepting a narrow null hypothesis and power is desirable. One possible remedy to choosing the appropriate level is to use the asymptotic conditional distribution of the meta-analysis statistic implied by Theorem 5.

We give three avenues for further work. The foregoing analysis, relying on the fixed-effects model, ignores the effect of heterogeneity in screening. Second, the analysis only looks at the effect of screening on true nulls. The effect of screening when publication bias is present but the generally under-powered publication bias test fails to identify it. This type of analysis requires a model for publication bias. Finally, another possible source of post-selection inference in conducting meta-analysis is the test for heterogeneity. The effect of choosing which type of meta-analysis to conduct based on the outcome of the preliminary heterogeneity test may influence the main analysis.

References

Begg, C. (1994). Publication bias. In H. Cooper and L. Hedges (Eds.), *The Handbook of Research Synthesis*, pp. 399–409. New York: Russel Sage Foundation.

Begg, C. and M. Mazumdar (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, 1088–1101.

Davey, J., R. M. Turner, M. J. Clarke, and J. Higgins (2011). Characteristics of meta-analyses and their component studies in the cochrane database of systematic reviews: a cross-sectional, descriptive analysis. *BMC Medical Research Methodology 11*(1), 1–11.

Dembo, A. (2016, September). Probability Theory: STAT310/MATH230.

Egger, M., G. D. Smith, M. Schneider, and C. Minder (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ 315*(7109), 629–634.

Konstantopoulos, S. and L. V. Hedges (2019). Statistically analyzing effect sizes: Fixed- and random-effects models. In H. Cooper, L. V. Hedges, and J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis*, pp. 245–280. New York, NY: Russell Sage Foundation.

Lin, L. and H. Chu (2018). Quantifying publication bias in meta-analysis. *Biometrics 74*(3), 785–794.

Macaskill, P., S. D. Walter, and L. Irwig (2001). A comparison of methods to detect publication bias in meta-analysis. *Statistics in Medicine 20*(4), 641–654.

Michael, H. and M. Ghebremichael. Comparison of the powers of begg's and egger's tests for publication bias. forthcoming.

Michael, H. and M. Ghebremichael. A correction to begg's test for publication bias. forthcoming.

Nolan, D. and D. Pollard (1988). Functional limit theorems for u-processes. *The Annals of Probability 16*(3), 1291–1298.

Olshen, R. A. (1973). The conditional level of the F-Test. *Journal of the American Statistical Association 68*(343), 692–698.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications.* John Wiley & Sons.

Rochon, J., M. Gondan, and M. Kieser (2012). To test or not to test: Preliminary assessment of normality when comparing two independent samples. *BMC medical research methodology 12*(1), 1–11.

Rothstein, H. R., A. J. Sutton, and M. Borenstein (2005). *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments.* West Sussex, England: John Wiley & Sons.

Schucany, W. R. and H. Tony Ng (2006). Preliminary goodness-of-fit tests for normality do not validate the one-sample student t. *Communications in Statistics-Theory and Methods 35*(12), 2275–2286.

Taylor, J. and R. J. Tibshirani (2015). Statistical learning and selective inference. *Proceedings of the National Academy of Sciences 112*(25), 7629–7634.

Van der Vaart, A. W. (2000). *Asymptotic Statistics.* Cambridge University Press.

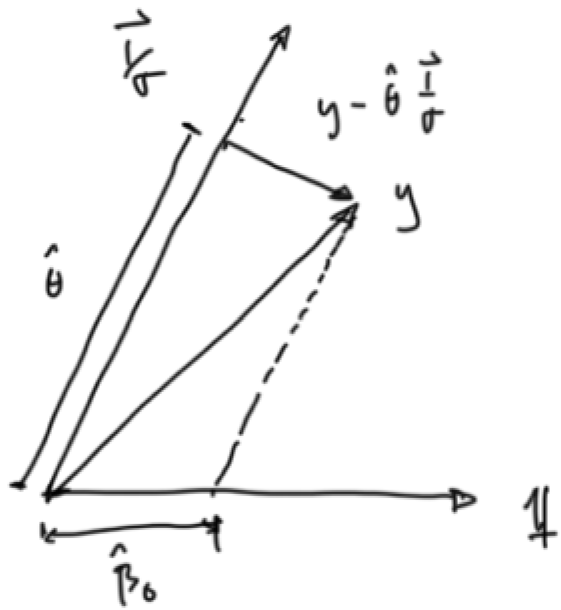
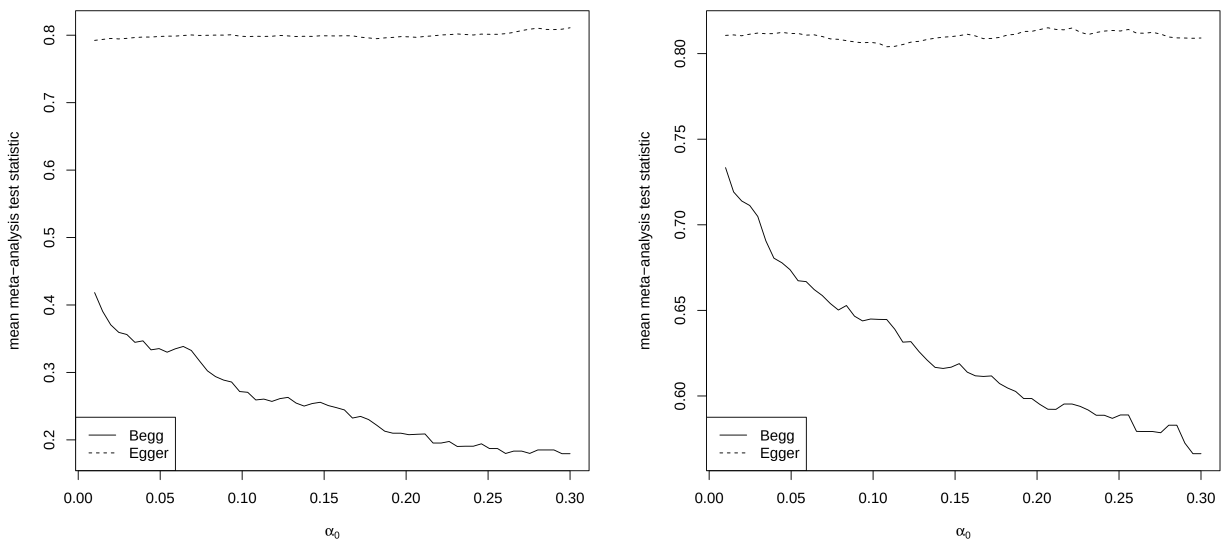


Figure 1: Schematic description of the orthogonality of $\hat{\beta}_0$ and $\hat{\theta}$ in Egger regression.



(a) Power-type f_Z.

(b) Beta f_Z.

Figure 2: The effect of the nominal level α_0 of the publication bias screen on the monte carlo mean of the meta-analysis test statistic.

condition	α_0	f_Z n | zeta	t low	med	high	power low	med	high	beta low	med	high
begg	0.05	25	0.050	0.046	0.009	0.050	0.050	0.045	0.049	0.049	0.031
		75	0.050	0.047	0.008	0.051	0.049	0.007	0.050	0.049	0.020
	0.15	25	0.051	0.046	0.007	0.049	0.050	0.040	0.049	0.048	0.026
		75	0.050	0.047	0.006	0.050	0.048	0.005	0.051	0.048	0.015
egger	0.05	25	0.050	0.048	0.012	0.050	0.050	0.052	0.049	0.049	0.049
		75	0.051	0.048	0.012	0.050	0.049	0.051	0.050	0.050	0.049
	0.15	25	0.051	0.048	0.012	0.049	0.050	0.053	0.049	0.049	0.048
		75	0.051	0.048	0.012	0.051	0.049	0.052	0.051	0.050	0.049
unconditional	0.05	25	0.050	0.047	0.012	0.050	0.050	0.052	0.049	0.050	0.049
		75	0.050	0.048	0.012	0.051	0.049	0.051	0.050	0.049	0.049
	0.15	25	0.050	0.047	0.012	0.050	0.050	0.052	0.049	0.050	0.049
		75	0.050	0.048	0.012	0.051	0.049	0.051	0.050	0.049	0.049

Table 1: $\theta = 0$.

		f_Z	t			power			beta		
condition	α_0	n | zeta	low	med	high	low	med	high	low	med	high
begg	0.05	25	0.158	0.148	0.032	0.167	0.165	0.113	0.166	0.165	0.132
		75	0.393	0.389	0.279	0.396	0.394	0.049	0.395	0.393	0.272
	0.15	25	0.158	0.147	0.029	0.167	0.165	0.095	0.166	0.163	0.117
		75	0.393	0.386	0.266	0.396	0.392	0.025	0.396	0.390	0.235
egger	0.05	25	0.159	0.150	0.039	0.167	0.167	0.160	0.166	0.167	0.170
		75	0.395	0.392	0.306	0.395	0.397	0.398	0.396	0.397	0.398
	0.15	25	0.159	0.151	0.040	0.168	0.166	0.162	0.166	0.167	0.168
		75	0.395	0.394	0.307	0.396	0.398	0.398	0.397	0.397	0.398
unconditional	0.05	25	0.159	0.149	0.037	0.168	0.166	0.158	0.167	0.167	0.169
		75	0.394	0.393	0.303	0.395	0.397	0.398	0.395	0.397	0.398
	0.15	25	0.159	0.149	0.037	0.168	0.166	0.158	0.167	0.167	0.169
		75	0.394	0.393	0.303	0.395	0.397	0.398	0.395	0.397	0.398

Table 2: $\theta = 0.2$.