meta bias — Tests for small-study effects in meta-analysis

Description

`meta bias` performs tests for the presence of small-study effects in a meta-analysis, also known as tests for funnel-plot asymmetry and publication-bias tests. Three regression-based tests and a nonparametric rank correlation test are available. For regression-based tests, you can include moderators to account for potential between-study heterogeneity.

Quick start

Test for small-study effects by using the Egger regression-based test

```
meta bias, egger
```

Same as above, but include a moderator `x1` to account for between-study heterogeneity induced by `x1`

```
meta bias x1, egger
```

Same as above, but assume a random-effects model with the empirical Bayes method for estimating τ^2 in the regression-based test

```
meta bias x1, egger random(ebayes)
```

With log risk-ratios, test for small-study effects by using the Harbord regression-based test with moderators `x1` and `x2` to account for between-study heterogeneity

```
meta bias x1 i.x2, harbord
```

With log odds-ratios, test for small-study effects by using the Peters regression-based test and assuming a common-effect model

```
meta bias, peters common
```

Menu

Statistics > Meta-analysis

Syntax

Regression-based tests for small-study effects

Test using meta-analysis model as declared with meta set or meta esize

`meta bias` [*moderators*] [*if*] [*in*] , *regtest* [*modelopts*]

Random-effects meta-analysis model

`meta bias` [*moderators*] [*if*] [*in*] , *regtest* `random`[`(`*remethod*`)`]
[`se(`*seadj*`)` *options*]

Common-effect meta-analysis model

`meta bias` [*if*] [*in*] , *regtest* `common` [*options*]

Fixed-effects meta-analysis model

`meta bias` [*moderators*] [*if*] [*in*] , *regtest* `fixed` [`multiplicative` *options*]

Traditional test

`meta bias` [*if*] [*in*] , *regtest* `traditional` [*options*]

Nonparametric rank correlation test for small-study effects

`meta bias` [*if*] [*in*] , `begg` [[`no`]`metashow detail`]

regtest	Description
`egger`	Egger's test
`harbord`	Harbord's test
`peters`	Peters's test

modelopts is any option relevant for the declared model.

remethod	Description
`reml`	restricted maximum likelihood; the default
`mle`	maximum likelihood
`ebayes`	empirical Bayes
`dlaird`	DerSimonian–Laird
`sjonkman`	Sidik–Jonkman
`hedges`	Hedges
`hschmidt`	Hunter–Schmidt

options	Description
Main	
tdistribution	report t test instead of z test
[no]metashow	display or suppress meta settings in the output
detail	display intermediate estimation results
Maximization	
maximize_options	control the maximization process of the between-study variance

moderators may contain factor variables; see [U] **11.4.3 Factor variables**.

collect is allowed; see [U] **11.1.10 Prefix commands**.

Options

Main

One of egger, harbord, peters, or begg (or their synonyms) must be specified. In addition to the traditional versions of the regression-based tests, their random-effects versions and extensions to allow for moderators are also available.

egger (synonym esphillips) specifies that the regression-based test of Egger, Davey Smith, and Phillips (1997) be performed. This test is known as the Egger test in the literature. This is the test of the slope in a weighted regression of the effect size, _meta_es, on its standard error, _meta_se, optionally adjusted for *moderators*. This test tends to have an inflated type I error rate for two-sample binary data.

harbord (synonym hesterne) specifies that the regression-based test of Harbord, Egger, and Sterne (2006) be performed. This test is known as the Harbord test. This is the test of the slope in a weighted regression of Z_j/V_j on $1/\sqrt{V_j}$, optionally adjusting for *moderators*, where Z_j is the score of the likelihood function and V_j is the score variance. This test is used for two-sample binary data with effect sizes log odds-ratio and log risk-ratio. It was designed to reduce the correlation between the effect-size estimates and their corresponding standard errors, which is inherent to the Egger test with two-sample binary data.

peters (synonym petersetal) specifies that the regression-based test of Peters et al. (2006) be performed. This test is known as the Peters test in the literature. This is the test of the slope in a weighted regression of the effect size, _meta_es, on the inverse sample size, $1/n_j$, optionally adjusted for *moderators*. The Peters test is used with two-sample binary data for log odds-ratios. Because it regresses effect sizes on inverse sample sizes, they are independent by construction.

begg (synonym bmazumdar) specifies that the nonparametric rank correlation test of Begg and Mazumdar (1994) be performed. This is not a regression-based test, so only options metashow, nometashow, and detail are allowed with it. This test is known as the Begg test in the literature. This test is no longer recommended in the literature and provided for completeness.

Options random(), common, and fixed, when specified with meta bias for regression-based tests, temporarily override the global model declared by meta set or meta esize during the computation. Options random(), common, and fixed may not be combined. If these options are omitted, the declared meta-analysis model is assumed; see *Declaring a meta-analysis model* in [META] **meta data**. Also see *Meta-analysis models* in [META] **Intro**.

random and random(*remethod*) specify that a random-effects model be assumed for regression-based test; see *Random-effects model* in [META] **Intro**.

remethod specifies the type of estimator for the between-study variance τ^2. *remethod* is one of reml, mle, ebayes, dlaird, sjonkman, hedges, or hschmidt. random is a synonym for random(reml). See *Options* in [META] **meta esize** for more information.

common specifies that a common-effect model be assumed for regression-based test; see *Common-effect ("fixed-effect") model* in [META] **Intro**. It uses the inverse-variance estimation method; see *Meta-analysis estimation methods* in [META] **Intro**. Also see the discussion in [META] **meta data** about common-effect versus fixed-effects models. common is not allowed in the presence of moderators.

fixed specifies that a fixed-effects model be assumed for regression-based test; see *Fixed-effects model* in [META] **Intro**. It uses the inverse-variance estimation method; see *Meta-analysis estimation methods* in [META] **Intro**. Also see the discussion in [META] **meta data** about fixed-effects versus common-effect models.

se(*seadj*) specifies that the adjustment *seadj* be applied to the standard errors of the coefficients. Additionally, the tests of significance of the coefficients are based on a Student's t distribution instead of the normal distribution. se() is allowed only with random-effects models.

seadj is khartung[, truncated]. Adjustment khartung specifies that the Knapp–Hartung adjustment (Hartung and Knapp 2001a, 2001b; Knapp and Hartung 2003), also known as the Sidik–Jonkman adjustment (Sidik and Jonkman 2002), be applied to the standard errors of the coefficients. hknapp and sjonkman are synonyms for khartung. truncated specifies that the truncated Knapp–Hartung adjustment (Knapp and Hartung 2003), also known as the modified Knapp–Hartung adjustment, be used.

traditional specifies that the traditional version of the selected regression-based test be performed. This option is equivalent to specifying options fixed, multiplicative, and tdistribution. It may not be specified with *moderators*.

multiplicative performs a fixed-effects regression-based test that accounts for residual heterogeneity by including a multiplicative variance parameter ϕ. ϕ is referred to as an "(over)dispersion parameter". See *Introduction* in [META] **meta regress** for details.

tdistribution reports a t test instead of a z test. This option may not be combined with option se().

metashow and nometashow display or suppress the meta setting information. By default, this information is displayed at the top of the output. You can also specify nometashow with meta update to suppress the meta setting output for the entire meta-analysis session.

detail specifies that intermediate estimation results be displayed. For regression-based tests, the results from the regression estimation will be displayed. For the nonparametric test, the results from ktau ([R] **spearman**) will be displayed.

Maximization

maximize_options: iterate(#), tolerance(#), nrtolerance(#), nonrtolerance (see [R] **Maximize**), from(#), and showtrace. These options control the iterative estimation of the between-study variance parameter, τ^2, with random-effects methods reml, mle, and ebayes. These options are seldom used.

from(#) specifies the initial value for τ^2 during estimation. By default, the initial value for τ^2 is the noniterative Hedges estimator.

showtrace displays the iteration log that contains the estimated parameter τ^2, its relative difference with the value from the previous iteration, and the scaled gradient.

Remarks and examples

Remarks are presented under the following headings:

Introduction

As we discussed in *Introduction* of [META] **meta funnelplot**, there is a tendency for smaller studies to report different, often larger, effect sizes than the larger studies. There are various reasons that explain this tendency, but the two more common ones are between-study heterogeneity and publication bias. We covered the between-study heterogeneity in [META] **meta summarize** and [META] **meta regress**. Here we focus on publication bias.

Publication bias often arises when the decision of whether to publish a study depends on the statistical significance of the results of the study. Typically, nonsignificant results from small studies have a tendency of not getting published. See *Publication bias* of [META] **Intro** for details.

The funnel plot ([META] **meta funnelplot**) is commonly used to investigate publication bias or, more generally, small-study effects in meta-analysis. The presence of asymmetry in the funnel plot may indicate the presence of publication bias. Graphical evaluation of funnel plots is useful for data exploration but may be subjective when detecting the asymmetry. Thus, a more formal evaluation of funnel-plot asymmetry is desired. Statistical tests were developed for detecting the asymmetry in a funnel plot; they are often called tests for funnel-plot asymmetry. They are also sometimes referred to as tests of publication bias, but this terminology may be misleading because the presence of a funnel-plot asymmetry is not always due to publication bias (for example, Sterne et al. [2011]). Thus, we prefer a more generic term—tests for small-study effects—suggested by Sterne, Gavaghan, and Egger (2000).

There are two types of tests for small-study effects: regression-based tests and a nonparametric rank-based test. The main idea behind these tests is to determine whether there is a statistically significant association between the effect sizes and their measures of precision such as effect-size standard errors.

The Egger regression-based test (Egger et al. 1997) performs a weighted linear regression of the effect sizes, $\widehat{\theta}_j$'s, on their standard errors, $\widehat{\sigma}_j$'s, weighted by the precision, $1/\widehat{\sigma}_j$'s. The test for the zero slope in that regression provides a formal test for small-study effects. In some cases, such as in the presence of a large true effect or with two-sample binary data, the Egger test tends to have an inflated type I error (for example, Harbord, Harris, and Sterne [2016]). Two alternative tests, the Harbord test and the Peters test, were proposed to alleviate the type I error problem in those cases.

The Harbord regression-based test (Harbord, Egger, and Sterne 2006) corresponds to the zero-slope test in a weighted regression of Z_j/V_j's on $1/\sqrt{V_j}$'s, where Z_j is the score of the likelihood function and V_j is the score variance. The Peters regression-based test (Peters et al. 2006) corresponds to the zero-slope test in a weighted regression of the effect sizes, $\widehat{\theta}_j$'s, on the respective inverse sample sizes, $1/n_j$'s. With two-sample binary data, these tests tend to perform better than the Egger test in terms of the type I error while maintaining similar power.

The rank correlation Begg test (Begg and Mazumdar 1994) tests whether Kendall's rank correlation between the effect sizes and their variances equals zero. The regression-based tests tend to perform better in terms of type I error than the rank correlation test. This test is provided mainly for completeness.

See Harbord, Harris, and Sterne (2016) and Steichen (2016) for more details about these tests.

As we discussed in [META] **meta funnelplot**, the presence of between-study heterogeneity may affect the symmetry of a funnel plot. Thus, any statistical method based on the funnel plot will also be affected (Sutton 2009). To account for the between-study heterogeneity, the regression-based tests can be extended to incorporate moderators that may help explain the heterogeneity (Sterne and Egger 2005).

The traditional version of the regression-based tests used a multiplicative fixed-effects meta-regression to account for residual heterogeneity (see *Introduction* of [META] **meta regress**). In addition to adjusting for moderators, a random-effects meta-regression is considered a better alternative to account for residual heterogeneity.

Ioannidis and Trikalinos (2007) provide the following recommendations for when it is appropriate to use small-study tests: a) the number of studies should be greater than 10; b) there should be at least one study with a statistically significant result; c) there should be no significant heterogeneity ($I^2 < 50\%$); and d) the ratio of the maximum to minimum variances across studies should be larger than 4; that is, $\max\left(\{\widehat{\sigma}_j^2\}_{j=1}^K\right) / \min\left(\{\widehat{\sigma}_j^2\}_{i=1}^K\right) > 4$. If a) is violated, the tests may have low power. If c) is violated, the asymmetry of the funnel plot may be induced by between-study heterogeneity rather than publication bias. If d) is violated, the funnel plot will look more like a horizontal line than an inverted funnel, and the funnel-asymmetry tests will have an inflated type I error. Also see Sterne et al. (2011) for details.

The results of the tests of small-study effects should be interpreted with caution. In the presence of small-study effects, apart from publication bias, other reasons should also be explored to explain the presence of small-study effects. If small-study effects are not detected by a test, their existence should not be ruled out because the tests tend to have low power.

Also see [META] **meta trimfill** for assessing the impact of publication bias on the results.

Using meta bias

`meta bias` performs tests for small-study effects. These tests are also known as the tests for funnel-plot asymmetry and tests for publication bias. You can choose from three regression-based tests: the Egger test (option `egger`), the Harbord test for two-sample binary data with effect sizes log odds-ratio and log risk-ratio (option `harbord`), and the Peters test for log odds-ratios (option `peters`). You can also perform the Begg nonparametric rank correlation test (option `begg`), but this test is no longer recommended in the meta-analysis literature.

Next, we describe the features that are relevant only to the regression-based tests. These tests are based on meta-regression of effect sizes and their measures of precision.

The default meta-analysis model (and method) are as declared by `meta set` or `meta esize`; see *Declaring a meta-analysis model* in [META] **meta data**. You can change the defaults by specifying one of options `random()`, `common()`, or `fixed()`.

Because the regression-based tests use meta-regression, many of the options of `meta regress` (see [META] **meta regress**) apply to `meta bias` as well. For example, you can specify that a multiplicative meta-regression be used by the test with option `multiplicative`. And you can specify to use the t test instead of a z test for inference with option `tdistribution`.

The regression-based tests support the `traditional` option, which specifies that the tests be performed as originally published. This option is a shortcut for `fixed`, `multiplicative`, and `tdistribution`.

To account for between-study heterogeneity when checking for publication bias, you can specify moderators with the regression-based tests.

Examples of using meta bias

Recall the pupil IQ data (Raudenbush and Bryk 1985; Raudenbush 1984) described in *Effects of teacher expectancy on pupil IQ (pupiliq.dta)* of [META] **meta**. Here we will use its declared version and will focus on the demonstration of various options of `meta bias` and explanation of its output.

```
. use https://www.stata-press.com/data/r19/pupiliqset
(Effects of teacher expectancy on pupil IQ; set with -meta set-)
. meta query, short
-> meta set stdmdiff se , studylabel(studylbl) eslabel(Std. mean diff.)
  Effect-size label: Std. mean diff.
   Effect-size type: Generic
        Effect size: stdmdiff
          Std. err.: se
              Model: Random effects
             Method: REML
```

From the meta summary, our data were declared by using `meta set` with variables `stdmdiff` and `se` specifying the effect sizes and their standard errors, respectively. The declared meta-analysis model is the default random-effects model with the REML estimation method.

Examples are presented under the following headings:

Example 1: Small-study effects due to a confounding moderator
Example 2: Traditional tests and detailed output
Example 3: Harbord's test for small-study effects

▷ Example 1: Small-study effects due to a confounding moderator

Our main focus is on investigating the potential presence of small-study effects by using a regression-based test. Because we are working with continuous data, we will use the Egger test.

```
. meta bias, egger
  Effect-size label: Std. mean diff.
        Effect size: stdmdiff
          Std. err.: se
Regression-based Egger test for small-study effects
Random-effects model
Method: REML
H0: beta1 = 0; no small-study effects
            beta1 =      1.83
      SE of beta1 =     0.724
                z =      2.53
       Prob > |z| =    0.0115
```

From the output header, the regression-based test uses the declared random-effects model with REML estimation to account for residual heterogeneity. The estimated slope, $\widehat{\beta}_1$, is 1.83 with a standard error of 0.724, giving a test statistic of $z = 2.53$ and a p-value of 0.0115. This means that there is some evidence of small-study effects.

In example 9 of [META] **meta summarize**, we used subgroup-analysis on binary variable `week1`, which records whether teachers had prior contact with students for more than 1 week or for 1 week or less, to account for between-study heterogeneity. It explained most of the heterogeneity present among the effect sizes, with generally higher effect sizes in the low contact group.

Moderators that can explain a substantial amount of the heterogeneity should be included in the regression-based test as a covariate. By properly accounting for heterogeneity through the inclusion of `week1`, we can test for small-study effects due to reasons other than heterogeneity. We include factor variable `week1` as a moderator as follows:

```
. meta bias i.week1, egger
  Effect-size label: Std. mean diff.
        Effect size: stdmdiff
          Std. err.: se
Regression-based Egger test for small-study effects
Random-effects model
Method: REML
Moderators: week1
H0: beta1 = 0; no small-study effects
            beta1 =      0.30
      SE of beta1 =     0.729
                z =      0.41
       Prob > |z| =    0.6839
```

Now that we have accounted for heterogeneity through moderator `week1`, the Egger test statistic is 0.41 with a p-value of 0.6839. Therefore, we have strong evidence to say that the presence of small-study effects was the result of heterogeneity induced by teacher-student prior contact time.

◁

▷ Example 2: Traditional tests and detailed output

For illustration, we perform the traditional version of the Egger regression-based test by specifying the `traditional` option. We also use the `detail` option to report the meta-regression results used to construct the Egger test.

```
. meta bias, egger traditional detail
  Effect-size label: Std. mean diff.
        Effect size: stdmdiff
          Std. err.: se
Fixed-effects meta-regression                         Number of obs   =         19
Error: Multiplicative                                 Dispersion phi =       1.69
Method: Inverse-variance                              Model F(1,17)   =       4.17
                                                      Prob > F        =     0.0571

    _meta_es   Coefficient  Std. err.      t    P>|t|     [95% conf. interval]

    _meta_se     1.627717   .7975212    2.04   0.057    -.0549052     3.31034
       _cons    -.1797108    .126835   -1.42   0.175    -.4473093    .0878876

Test of residual homogeneity: Q_res = chi2(17) = 28.77   Prob > Q_res = 0.0367
Regression-based Egger test for small-study effects
Fixed-effects model
Method: Inverse-variance
H0: beta1 = 0; no small-study effects
            beta1 =      1.63
      SE of beta1 =     0.798
                t =      2.04
       Prob > |t| =    0.0571
```

The traditional version also suggests the presence of small-study effects, but its p-value, 0.0571, is larger than that from example 1.

The results of the above command is identical to the following:

```
. meta regress _meta_se, fixed multiplicative tdistribution

  Effect-size label: Std. mean diff.
        Effect size: stdmdiff
          Std. err.: se

Fixed-effects meta-regression                       Number of obs   =        19
Error: Multiplicative                               Dispersion phi  =      1.69
Method: Inverse-variance                            Model F(1,17)   =      4.17
                                                    Prob > F        =    0.0571

------------------------------------------------------------------------------
    _meta_es | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
    _meta_se |   1.627717   .7975212     2.04   0.057    -.0549052     3.31034
       _cons |  -.1797108    .126835    -1.42   0.175    -.4473093    .0878876
------------------------------------------------------------------------------
Test of residual homogeneity: Q_res = chi2(17) = 28.77   Prob > Q_res = 0.0367
```

The header and coefficient table from `meta bias`'s detailed output is identical to that produced by `meta regress` (see [META] **meta regress**).

◁

▷ Example 3: Harbord's test for small-study effects

In example 1 of [META] **meta funnelplot**, we explored the presence of publication bias in the NSAIDS data, which was described in *Effectiveness of nonsteroidal anti-inflammatory drugs (nsaids.dta)* of [META] **meta**. The contour-enhanced funnel plot from example 5 of [META] **meta funnelplot** revealed that the funnel-plot asymmetry was caused by the absence of small studies in the region where the tests of the log odds-ratios equal to zero were not statistically significant. This may suggest the presence of publication bias. We can explore this more formally by performing a test for small-study effects.

We use the declared version of the NSAIDS dataset.

```
. use https://www.stata-press.com/data/r19/nsaidsset, clear
(Effectiveness of nonsteroidal anti-inflammatory drugs; set with -meta esize-)

. meta query, short
-> meta esize nstreat nftreat nscontrol nfcontrol

  Effect-size label: Log odds-ratio
   Effect-size type: lnoratio
        Effect size: _meta_es
          Std. err.: _meta_se
              Model: Random effects
             Method: REML
```

The declared effect size is log odds-ratio, so we will use the Harbord regression-based test to investigate whether the small-study effects (or funnel-plot asymmetry) is present in these data.

```
. meta bias, harbord

  Effect-size label: Log odds-ratio
        Effect size: _meta_es
          Std. err.: _meta_se

Regression-based Harbord test for small-study effects
Random-effects model
Method: REML

H0: beta1 = 0; no small-study effects
            beta1 =      3.03
      SE of beta1 =     0.741
                z =      4.09
       Prob > |z| =    0.0000
```

The p-value is less than 0.0001, so we reject the null hypothesis of no small-study effects. It is difficult to be certain whether the small-study affects are driven by publication bias because of the presence of substantial heterogeneity in these data (see [META] **meta summarize**). Note that the regression-based test assumed an (REML) random-effects model, which accounts for heterogeneity present among the studies. If we had access to study-level covariates for these data that could explain some of the between-study variability, we could have specified them with `meta bias`.

◁

Stored results

For regression-based tests, `meta bias` stores the following in `r()`:

Scalars

`r(beta1)`	estimate of the main slope coefficient
`r(se)`	standard error for the slope estimate
`r(z)`	z statistic
`r(t)`	t statistic
`r(p)`	two-sided p-value

Macros

`r(testtype)`	type of test: `egger`, `harbord`, or `peters`
`r(model)`	meta-analysis model
`r(method)`	meta-analysis estimation method
`r(moderators)`	moderators used in regression-based tests

Matrices

`r(table)`	regression results

For Begg's test, `meta bias` stores the following in `r()`:

Scalars

`r(score)`	Kendall's score estimate
`r(se_score)`	standard error of Kendall's score
`r(z)`	z test statistic
`r(p)`	two-sided p-value

Macros

`r(testtype)`	`begg`

Methods and formulas

Methods and formulas are presented under the following headings:

Let K be the number of studies for a given meta-analysis. For the jth study, $\widehat{\theta}_j$ denotes the estimated effect size, and $\widehat{\sigma}_j^2$ denotes the effect-size (within-study) variance. The tests are applicable to any type of effect size as long as it is asymptotically normally distributed.

For two-sample binary data, also consider the following 2×2 table for the jth study.

group	event	no event	size
treatment	a_j	b_j	$n_{1j} = a_j + b_j$
control	c_j	d_j	$n_{2j} = c_j + d_j$

The total sample size for the jth study is denoted by $n_j = n_{1j} + n_{2j}$.

Regression-based tests

Regression-based tests use meta-regression to examine a linear relationship between the individual effect sizes and measures of study precision such as the effect-size standard errors, possibly adjusting for moderators that explain some of the between-study variability.

In the subsections below, we provide the traditional versions of the regression-based tests. The extensions of traditional versions include the support of other models such as a random-effects model and the support of moderators.

In the presence of moderators, the test for small-study effects is the test of H_0: $\beta_1 = 0$ in the corresponding meta-regression with the following linear predictor,

$$\mathbf{x}_j\boldsymbol{\beta} = \beta_0 + \beta_1 m_j + \beta_2 x_{2,j} + \cdots + \beta_{p-1} x_{p-1,j}$$

where $x_{2,j}, \ldots, x_{p-1,j}$ represent the *moderators* specified with `meta bias` and $m_j = \widehat{\sigma}_j$ for the Egger test, $m_j = 1/\sqrt{V_j}$ for the Harbord test, and $m_j = 1/n_j$ for the Peters test. See the subsections below for details about these tests. Also see Sterne and Egger (2005).

The computations of regression-based tests are based on the corresponding meta-regression models; see *Methods and formulas* of [META] **meta regress**.

The formulas below are based on Harbord, Harris, and Sterne (2016), Sterne and Egger (2005), and Peters et al. (2010).

Egger's linear regression test

The formulas and discussion in this subsection are based on Sterne and Egger (2005).

The test proposed by Egger, Davey Smith, Schneider, and Minder (1997) is based on a simple linear regression of the standard normal variate, which is defined as the individual effect-size estimate divided by its standard error, against the study precision, which is defined as the reciprocal of the standard error:

$$E\left(\frac{\hat{\theta}_j}{\hat{\sigma}_j}\right) = b_0 + b_1 \frac{1}{\hat{\sigma}_j} \tag{1}$$

The Egger test of no small-study effects is the test of H_0: $b_0 = 0$.

Linear regression model (1) is equivalent to the weighted linear regression of the effect sizes $\hat{\theta}_j$'s on their standard errors $\hat{\sigma}_j$'s,

$$E\left(\hat{\theta}_j\right) = b_1 + b_0 \hat{\sigma}_j \tag{2}$$

with weights inversely proportional to the variances of the effect sizes, $w_j = 1/\hat{\sigma}_j^2$. Note that the intercept b_0 in regression (1) corresponds to the slope in the weighted regression (2). Therefore, Egger test for small-study effects corresponds to a test of a linear trend in a funnel plot (see [META] **meta funnelplot**) of effect sizes against their standard errors.

Let's denote $\beta_0 = b_1$ and $\beta_1 = b_0$. The statistical model for the traditional Egger's test, as it originally appeared in the literature (Egger et al. 1997), is given by

$$\hat{\theta}_j = \beta_0 + \beta_1 \hat{\sigma}_j + \epsilon_j \quad \text{weighted by } w_j = 1/\hat{\sigma}_j^2, \text{ where } \epsilon_j \sim N\left(0,\, \hat{\sigma}_j^2 \phi\right)$$

and ϕ is the overdispersion parameter as defined in multiplicative meta-regression; see *Introduction* of [META] **meta regress**.

Egger's test for small-study effects is the test of H_0: $\beta_1 = 0$, and the null hypothesis is rejected if

$$t_{\text{egger}} = \left| \frac{\hat{\beta}_1}{\widehat{\text{SE}}\left(\hat{\beta}_1\right)} \right| > t_{K-2,1-\alpha/2}$$

where $t_{K-2,1-\alpha/2}$ is the $(1-\alpha/2)$th quantile of the Student's t distribution with $K-2$ degrees of freedom. The above test is performed when you specify options `egger` and `traditional`.

❑ Technical note

Sterne and Egger (2005) point out that, originally, Egger et al. (1997) used a weighted version of (1) with weights equal to the inverse of the variances of effect sizes ($1/\hat{\sigma}_j^2$'s). The authors strongly recommend that this version of the test not be used because it does not have a theoretical justification.

❑

Harbord's test for log odds-ratios or log risk-ratios

Consider the fixed-effects model $\hat{\theta}_j \sim N(\theta, \hat{\sigma}_j^2)$. For a study j, let Z_j be the first derivative (score) and V_j be the negative second derivative (Fisher's information) of the model log likelihood with respect to θ evaluated at $\theta = 0$ (Whitehead and Whitehead 1991; Whitehead 1997).

For two-sample binary data, Harbord, Egger, and Sterne (2006) proposed a modification of the Egger test based on the intercept in an unweighted regression of $Z_j/\sqrt{V_j}$ against $\sqrt{V_j}$:

$$E\left(\frac{Z_j}{\sqrt{V_j}}\right) = b_0 + b_1\sqrt{V_j} \tag{3}$$

When the effect of interest is the log odds-ratio,

$$Z_j = \frac{a_j - (a_j + c_j)\, n_{1j}}{n_j} \text{ and } V_j = \frac{n_{1j} n_{2j} (a_j + c_j)(b_j + d_j)}{n_j^2 (n_j - 1)}$$

Note that Z_j and V_j are the numerator and denominator of the log Peto's odds-ratio as defined in *Methods and formulas* of [META] **meta esize**.

When the effect of interest is the log risk-ratio,

$$Z_j = \frac{a_j n_j - (a_j + c_j)\, n_{1j}}{b_j + d_j} \text{ and } V_j = \frac{n_{1j} n_{2j} (a_j + c_j)}{n_j (b_j + d_j)}$$

Whitehead (1997) showed that when θ_j is small and n_j is large, $\widehat{\theta}_j \approx Z_j/V_j$ and $\widehat{\sigma}_j^2 \approx 1/V_j$. In this case, the Harbord regression model (3) is equivalent to Egger's regression model (1). Thus, Harbord's test becomes equivalent to Egger's test when all studies are large and have small effect sizes (Harbord, Harris, and Sterne 2016).

As with Egger's test, if we use the weighted version of regression model (3) and denote $\beta_0 = b_1$ and $\beta_1 = b_0$ in that model, the statistical model for the Harbord test, as it originally appeared in the literature, is given by

$$\frac{Z_j}{V_j} = \beta_0 + \beta_1 \frac{1}{\sqrt{V_j}} + \epsilon_j \quad \text{weighted by } w_j = V_j, \text{ where } \epsilon_j \sim N\left(0, \frac{\phi}{V_j}\right)$$

where ϕ is the overdispersion parameter as defined in multiplicative meta-regression; see *Introduction* of [META] **meta regress**.

Then, the traditional Harbord test is the test of $H_0\colon \beta_1 = 0$, and its null hypothesis is rejected if $t_{\text{harbord}} = \left|\widehat{\beta}_1/\text{SE}(\widehat{\beta}_1)\right| > t_{K-2,1-\alpha/2}$. This test can be performed when you specify options `harbord` and `traditional`.

Peters's test for log odds-ratios

Peters et al. (2006) provide a test based on the following model:

$$\widehat{\theta}_j = \beta_0 + \beta_1 \frac{1}{n_j} + \epsilon_j \quad \text{weighted by } w_j = (a_j + c_j)(b_j + d_j)/n_j, \text{ where } \epsilon_j \sim N(0, \widehat{\sigma}_j^2 \phi)$$

$\widehat{\theta}_j = \ln\left(\widehat{\text{OR}}_j\right)$, and ϕ is the overdispersion parameter as defined in multiplicative meta-regression; see *Introduction* of [META] **meta regress**.

The traditional Peters test is the test of $H_0 : \beta_1 = 0$, and its null hypothesis is rejected if $t_{\text{peters}} = \left|\widehat{\beta}_1/\text{SE}(\widehat{\beta}_1)\right| > t_{K-2,1-\alpha/2}$. This test can be performed when you specify options `peters` and `traditional`.

When the test is based on the random-effects model, the weights are given by $w_j = 1/(\widehat{\sigma}_j^2 + \widehat{\tau}^2)$.

Begg's rank correlation test

Consider the standardized effect sizes

$$\widehat{\theta}_j^{\text{s}} = \frac{\widehat{\theta}_j - \widehat{\theta}_{\text{IV}}}{\sqrt{v_j^{\text{s}}}}$$

where

$$\widehat{\theta}_{\text{IV}} = \frac{\sum_{j=1}^K \widehat{\theta}_j/\widehat{\sigma}_j^2}{\sum_{j=1}^K 1/\widehat{\sigma}_j^2}$$

and

$$v_j^{\text{s}} = \text{Var}\left(\widehat{\theta}_j - \widehat{\theta}_{\text{IV}}\right) = \widehat{\sigma}_j^2 - \left(\sum_{j=1}^K \widehat{\sigma}_j^{-2}\right)^{-1}$$

The Begg test (Begg and Mazumdar 1994) is Kendall's rank correlation test of independence between $\widehat{\theta}_j^{\text{s}}$'s and $\widehat{\sigma}_j^2$'s; see *Methods and formulas* of [R] **spearman**.

References

Begg, C. B., and M. Mazumdar. 1994. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 50: 1088–1101. https://doi.org/10.2307/2533446.

Egger, M., G. Davey Smith, and A. N. Phillips. 1997. Meta-analysis: Principles and procedures. *BMJ* 315: 1533–1537. https://doi.org/10.1136/bmj.315.7121.1533.

Egger, M., G. Davey Smith, M. Schneider, and C. Minder. 1997. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 315: 629–634. https://doi.org/10.1136/bmj.315.7109.629.

Harbord, R. M., M. Egger, and J. A. C. Sterne. 2006. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Statistics in Medicine* 25: 3443–3457. https://doi.org/10.1002/sim.2380.

Harbord, R. M., R. J. Harris, and J. A. C. Sterne. 2016. "Updated tests for small-study effects in meta-analyses". In *Meta-Analysis in Stata: An Updated Collection from the Stata Journal*, edited by T. M. Palmer and J. A. C. Sterne, 153–165. 2nd ed. College Station, TX: Stata Press.

Hartung, J., and G. Knapp. 2001a. On tests of the overall treatment effect in meta-analysis with normally distributed responses. *Statistics in Medicine* 20: 1771–1782. https://doi.org/10.1002/sim.791.

———. 2001b. A refined method for the meta-analysis of controlled clinical trials with binary outcome. *Statistics in Medicine* 20: 3875–3889. https://doi.org/10.1002/sim.1009.

Ioannidis, J. P. A., and T. A. Trikalinos. 2007. The appropriateness of asymmetry tests for publication bias in meta-analyses: A large survey. *Canadian Medical Association Journal* 176: 1091–1096. https://doi.org/10.1503/cmaj.060410.

Knapp, G., and J. Hartung. 2003. Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine* 22: 2693–2710. https://doi.org/10.1002/sim.1482.

Peters, J. L., A. J. Sutton, D. R. Jones, K. R. Abrams, and L. Rushton. 2006. Comparison of two methods to detect publication bias in meta-analysis. *Journal of the American Medical Association* 295: 676–680. https://doi.org/10.1001/jama.295.6.676.

Peters, J. L., A. J. Sutton, D. R. Jones, K. R. Abrams, L. Rushton, and S. G. Moreno. 2010. Assessing publication bias in meta-analyses in the presence of between-study heterogeneity. *Journal of the Royal Statistical Society*, A ser., 173: 575–591. https://doi.org/10.1111/j.1467-985X.2009.00629.x.

Raudenbush, S. W. 1984. Magnitude of teacher expectancy effects on pupil IQ as a function of the credibility of expectancy induction: A synthesis of findings from 18 experiments. *Journal of Educational Psychology* 76: 85–97. https://doi.org/10.1037/0022-0663.76.1.85.

Raudenbush, S. W., and A. S. Bryk. 1985. Empirical Bayes meta-analysis. *Journal of Educational Statistics* 10: 75–98. https://doi.org/10.2307/1164836.

Sidik, K., and J. N. Jonkman. 2002. A simple confidence interval for meta-analysis. *Statistics in Medicine* 21: 3153–3159. https://doi.org/10.1002/sim.1262.

Steichen, T. J. 2016. "Tests for publication bias in meta-analysis". In *Meta-Analysis in Stata: An Updated Collection from the Stata Journal*, edited by T. M. Palmer and J. A. C. Sterne, 166–176. 2nd ed. College Station, TX: Stata Press.

Sterne, J. A. C., and M. Egger. 2005. "Regression methods to detect publication and other bias in meta-analysis". In *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*, edited by H. R. Rothstein, A. J. Sutton, and M. Borenstein, 99–110. Chichester, UK: Wiley. https://doi.org/10.1002/0470870168.ch6.

Sterne, J. A. C., D. Gavaghan, and M. Egger. 2000. Publication and related bias in meta-analysis: Power of statistical tests and prevalence in the literature. *Journal of Clinical Epidemiology* 53: 1119–1129. https://doi.org/10.1016/S0895-4356(00)00242-0.

Sterne, J. A. C., A. J. Sutton, J. P. A. Ioannidis, N. Terrin, D. R. Jones, J. Lau, J. R. Carpenter, G. Rücker, R. M. Harbord, C. H. Schmid, J. Tetzlaff, J. J. Deeks, J. L. Peters, P. Macaskill, G. Schwarzer, S. Duval, D. G. Altman, D. Moher, and J. P. T. Higgins. 2011. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 343: d4002. https://doi.org/10.1136/bmj.d4002.

Sutton, A. J. 2009. "Publication bias". In *The Handbook of Research Synthesis and Meta-Analysis*, edited by H. Cooper, L. V. Hedges, and J. C. Valentine, 435–452. 2nd ed. New York: Russell Sage Foundation.

Whitehead, A., and J. Whitehead. 1991. A general parametric approach to the meta-analysis of randomized clinical trials. *Statistics in Medicine* 10: 1665–1677. https://doi.org/10.1002/sim.4780101105.

Whitehead, J. 1997. *The Design and Analysis of Sequential Clinical Trials*. Rev. 2nd ed. Chichester, UK: Wiley.

Also see

[META] **meta data** — Declare meta-analysis data

[META] **meta funnelplot** — Funnel plots

[META] **meta regress** — Meta-analysis regression

[META] **meta summarize** — Summarize meta-analysis data

[META] **meta trimfill** — Nonparametric trim-and-fill analysis of publication bias

[META] **meta** — Introduction to meta

[META] **Glossary**

[META] **Intro** — Introduction to meta-analysis