meta mvregress — Multivariate meta-regression

Description

`meta mvregress` performs multivariate meta-regression. You can think of multivariate meta-regression as an extension of meta-regression, where multiple potentially dependent effect sizes are available for each study. `meta mvregress` performs both random-effects and fixed-effects multivariate meta-regression with various covariance structures and estimation methods for the random effects. `meta mvregress` is a standalone command in the sense that it does not require you to declare your data as `meta` data using `meta set` or `meta esize`.

Quick start

Perform random-effects multivariate meta-analysis of the effect-size variables `y1` and `y2` with within-study covariance structure defined by variables `v11`, `v12`, and `v22`

```
meta mvregress y1 y2, wcovvariables(v11 v12 v22)
```

Same as above, but perform random-effects multivariate meta-regression on continuous variable `x1` and factor variable `x2`

```
meta mvregress y1 y2 = x1 i.x2, wcovvariables(v11 v12 v22)
```

Same as above, but estimate random-effects using ML instead of the default REML

```
meta mvregress y1 y2 = x1 i.x2, wcovvariables(v11 v12 v22) random(mle)
```

Same as above, but specify an independent random-effects covariance structure instead of the default unstructured covariance matrix

```
meta mvregress y1 y2 = x1 i.x2, wcovvariables(v11 v12 v22) ///
random(mle, covariance(independent))
```

Same as above, but use a truncated Jackson–Riley adjustment to the standard errors of coefficients

```
meta mvregress y1 y2 = x1 i.x2, wcovvariables(v11 v12 v22) ///
random(mle, covariance(independent) se(truncjriley))
```

Perform a fixed-effects multivariate meta-analysis of variables `y1` and `y2` with standard error variables `s1` and `s2`, and assume a within-study correlation value of 0

```
meta mvregress y1 y2, fixed wsevariables(s1 s2) wcorrelations(0)
```

Perform multivariate meta-analysis of three effect-size variables `y1`, `y2`, and `y3` with six within-study variance–covariance variables `v11`, `v12`, `v13`, `v22`, `v23`, and `v33`

```
meta mvregress y1 y2 y3, wcovvariables(v11 v12 v13 v22 v23 v33)
```

Same as above, but using *varlist* shortcut notations and assuming the variables appear in the dataset in the order shown above

```
meta mvregress y1-y3, wcovvariables(v11-v33)
meta mvregress y*, wcovvariables(v*)
```

Menu

Statistics > Meta-analysis

Syntax

Random-effects multivariate meta-regression

meta mvregress *depvars* = *moderators* [*if*] [*in*] , *wcovspec* [random(*randomspec*)
options]

Fixed-effects multivariate meta-regression

meta mvregress *depvars* = *moderators* [*if*] [*in*] , *wcovspec* fixed [*options*]

Multivariate meta-analysis (constant-only model)

meta mvregress *depvars* [*if*] [*in*] , *wcovspec* [*modelopts*]

wcovspec	Description
Model	
* wcovvariables(*varlist*)	specify within-study variance and covariance variables
* wsevariables(*varlist*)	specify within-study standard-error variables
* wcorrelations(#|*numlist*)	specify within-study correlation values

Either wcovvariables() or both wsevariables() and wcorrelations() are required.

For random(*randomspec*), the syntax of *randomspec* is

remethod [, covariance(*recov*) se(*seadj*)]

remethod	Description
reml	restricted maximum likelihood; the default
mle	maximum likelihood
jwriley	Jackson–White–Riley

recov	Description
unstructured	all variances and covariances to be distinctly estimated; the default
independent	one unique variance parameter per random effect; all covariances 0
exchangeable	equal variances for random effects and one common pairwise covariance
identity	equal variances for random effects; all covariances 0
fixed(*matname*)	fixed random-effects covariance matrix *matname*

seadj	Description
jriley	Jackson–Riley standard-error adjustment
truncjriley	truncated Jackson–Riley standard-error adjustment

options	Description
Model	
noconstant	suppress constant term
tdistribution(#)	compute t tests instead of z tests for regression coefficients
Reporting	
level(#)	set confidence level; default is level(95)
stddeviations	show random-effects parameter estimates as standard deviations and correlations; the default
variance	show random-effects parameter estimates as variances and covariances
nohomtest	suppress output for homogeneity test
noretable	suppress random-effects table
nofetable	suppress fixed-effects table
estmetric	show parameter estimates as stored in e(b)
noheader	suppress output header
display_options	control columns and column formats, row spacing, line width, display of omitted variables and base and empty cells, and factor-variable labeling
Maximization	
maximize_options	control the maximization process; seldom used
coeflegend	display legend instead of statistics

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

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

coeflegend does not appear in the dialog box.

See **[U] 20 Estimation and postestimation commands** for more capabilities of estimation commands.

modelopts is any of *options* except noconstant.

Options

Model

wcovvariables(*varlist*) or wsevariables(*varlist*) and wcorrelations(#|*numlist*) specify information about the within-study covariance matrices $\mathbf{\Lambda}_j$, which are required for multivariate meta-regression.

wcovvariables(*varlist*) specifies variables that define the within-study covariance matrices $\mathbf{\Lambda}_j$. If d is the number of *depvars*, then $d(d+1)/2$ variables must be provided. The order in which the variables are specified is important. For example, if we have $d = 3$ dependent variables y1, y2, and y3, then 6 variables must be provided within wcovvariables() in the following order: Var(y1), Cov(y1, y2), Cov(y1, y3), Var(y2), Cov(y2, y3), and Var(y3). This option may not be combined with options wsevariables() and wcorrelations().

wsevariables(*varlist*) specifies variables that define the within-study standard errors of *depvars*. This option is useful, in combination with wcorrelations(), when the within-study covariances are not reported but only standard errors are available for *depvars*. If d is the number of *depvars*, then d variables must be specified, which represent the within-study standard errors of each variable in *depvars*. The order of the variables must follow the order in which *depvars* were specified. This option must be specified in combination with option `wcorrelations()`, which together define the within-study covariance matrices. `wsevariables()` may not be combined with `wcovvariables()`.

`wcorrelations(#|`*numlist*`)` specifies values for the within-study correlations between *depvars*. This option is also used to specify assumed correlations when only within-study standard errors are available, which are specified in option `wsevariables()`. If `wcorrelations(#)` is specified, # is assumed to be the common within-study correlation value between all *depvars*. If *numlist* is specified, then $d(d-1)/2$ values must be provided, where d is the number of *depvars*. The order in which the correlation values are specified is important. For example, if we have $d = 3$ dependent variables `y1`, `y2`, and `y3`, then 3 values must be provided in the following order: $\text{Corr}(\texttt{y1},\texttt{y2})$, $\text{Corr}(\texttt{y1},\texttt{y3})$, and $\text{Corr}(\texttt{y2},\texttt{y3})$. This option must be specified in combination with option `wsevariables()`, which together define the within-study covariance matrices. `wcorrelations()` may not be combined with `wcovvariables()`.

`random` and `random(`*randomspec*`)` specify that a random-effects model be assumed for the multivariate meta-regression. The syntax for *randomspec* is *remethod* [, `covariance(`*recov*`)` `se(`*seadj*`)`].

remethod specifies the type of estimator for the between-study covariance matrix $\mathbf{\Sigma}$. *remethod* is one of `reml`, `mle`, or `jwriley`. `random` is a synonym for `random(reml)`.

`reml`, the default, specifies that the REML method (Jackson, Riley, and White 2011) be used to estimate $\mathbf{\Sigma}$. This method produces an unbiased positive semidefinite estimate of the between-study covariance matrix and is commonly used in practice. The `reml` method requires iteration.

`mle` specifies that the ML method (Jackson, Riley, and White 2011) be used to estimate $\mathbf{\Sigma}$. It produces a positive semidefinite estimate of the between-study covariance matrix. With a few studies or small studies, this method may produce biased estimates. With many studies, the ML method is more efficient than the REML method. Method `mle` requires iteration.

`jwriley` specifies that the Jackson–White–Riley method (Jackson, White, and Riley 2013) be used to estimate $\mathbf{\Sigma}$. This method is a multivariate generalization of the popular DerSimonian–Laird method in univariate meta-analysis. The method does not make any assumptions about the distribution of random effects and does not require iteration. But it may produce an estimate of $\mathbf{\Sigma}$ that is not positive semidefinite and is thus "truncated" (via spectral decomposition) in that case.

`covariance(`*recov*`)` specifies the structure of the covariance matrix for the random effects. *recov* is one of the following: `unstructured`, `independent`, `exchangeable`, `identity`, or `fixed(`*matname*`)`.

`unstructured` allows for all variances and covariances to be distinct. If there are d random-effects terms (corresponding to the d *depvars*), the unstructured covariance matrix will have $d(d+1)/2$ unique parameters. This is the default covariance structure.

`independent` allows for a distinct variance for each random effect corresponding to a dependent variable and assumes that all covariances are 0.

`exchangeable` specifies one common variance for all random effects and one common pairwise covariance.

`identity` is short for "multiple of the identity"; that is, all variances are equal and all covariances are 0.

`fixed(`*matname*`)` specifies a fixed (known) $\mathbf{\Sigma} = matname$. This covariance structure requires no iteration.

se(*seadj*) specifies that the adjustment *seadj* be applied to the standard errors of the regression coefficients. Additionally, the tests of significance of the regression coefficients are based on a Student's t distribution instead of the normal distribution. The Jackson–Riley adjustments are multivariate generalizations of the Knapp–Hartung standard-error adjustments in univariate meta-regression. *seadj* is one of jriley or truncjriley.

jriley specifies that the Jackson–Riley adjustment (Jackson and Riley 2014) be applied to the standard errors of the coefficients.

truncjriley specifies that the truncated Jackson–Riley adjustment (Jackson and Riley 2014) be applied to the standard errors of the coefficients.

fixed specifies that a fixed-effects model be assumed for the multivariate meta-regression. In this case, $\mathbf{\Sigma} = \mathbf{0}$, and no iteration is performed to estimate the random-effects parameters.

noconstant; see [R] **Estimation options**. This option is not allowed with constant-only multivariate meta-regression.

tdistribution(#) computes t tests instead of z tests for the regression coefficients. The t tests are based on # degrees of freedom, which does not have to be an integer.

Reporting

level(#); see [R] **Estimation options**.

stddeviations, variance; see [ME] **mixed**.

nohomtest suppresses the homogeneity test based on the Q statistic from the output.

noretable, nofetable, estmetric, noheader; see [ME] **mixed**.

display_options: noci, nopvalues, noomitted, vsquish, noemptycells, baselevels, allbaselevels, nofvlabel, fvwrap(#), fvwrapon(*style*), cformat(%*fmt*), pformat(%*fmt*), sformat(%*fmt*), and nolstretch; see [R] **Estimation options**.

Maximization

maximize_options: difficult, technique(*algorithm_spec*), iterate(#), [no]log, trace, gradient, showstep, hessian, showtolerance, tolerance(#), ltolerance(#), nrtolerance(#), and nonrtolerance; see [R] **Maximize**. Those that require special mention for meta mvregress are listed below.

For the technique() option, the default is technique(nr). The bhhh algorithm is not available.

matsqrt, the default, and matlog; see [ME] **mixed**, except meta mvregress implies a single model level.

maximize_options are not available with fixed-effects multivariate meta-regression.

The following option is available with meta mvregress but is not shown in the dialog box:

coeflegend; see [R] **Estimation options**.

Remarks and examples

Remarks are presented under the following headings:

Introduction

Multivariate meta-regression is a technique used to study the relationship between multiple, usually dependent, effect sizes reported for each study and covariates. Multivariate meta-regression is analogous to multivariate regression that is used when individual data are available, but in multivariate meta-regression, the observations are the studies, the outcomes of interest are effect sizes, and the covariates are recorded at the study level. The study-level covariates in meta-regression are known as moderators. Examples of moderators include study publication year, study test environment, and drug administration method. For a comprehensive introduction to multivariate meta-regression, see Gleser and Olkin (2009) and Jackson, Riley, and White (2011).

A study may report multiple effect sizes in two different scenarios. In the first scenario, a study may compare various treatment groups against a common control group. For example, in a study that investigates the effect of multiple dietary regimens on weight loss, independent groups of individuals may be assigned to one of several diets: Keto diet, vegan diet, high-protein diet, or intermittent fasting. Multiple effect sizes that compare each of these diets with a control group (not following an assigned diet) can be computed. These effect sizes are usually correlated because they share a common control group. Studies falling under this category are called "multiple-treatment studies" or "mixed-treatment studies" in the multivariate meta-analysis literature.

In the second scenario, subjects are allocated to a treatment group or a control group as in the case of univariate meta-analysis, but multiple outcomes (endpoints) are compared across the two groups. For example, consider a study that explores the impact of a new teaching technique on math (outcome 1), physics (outcome 2), and chemistry (outcome 3) testing scores. Students are randomly assigned to one of two groups: those who were taught using the new technique (treatment group) and those who were not (control group). Three effect sizes that compare the three testing scores across the two groups are computed. These effect sizes are dependent because they were reported on the same set of students. Studies of this kind are referred to as "multiple-endpoint studies" in the literature.

Traditionally, the standard approach for handling multiple effect sizes reported per study was to perform separate univariate meta-analysis for each effect size. This approach ignores the dependence between the effect sizes and usually leads to biased pooled effects with overestimated variances. Another approach (Rosenthal and Rubin 1986) is to summarize the multiple effects by a single value for each study and then combine these values via standard univariate meta-analysis. This approach will result in information loss because of data reduction and may yield univariate summaries that are difficult to interpret in light of the original dependent effect sizes.

By properly accounting for the dependence between the effect sizes, multivariate meta-regression often provides parameter estimators with more optimal properties when compared with the previous two approaches. This is because it exploits the correlation between the multiple effect sizes, and thus the dependent effect sizes may borrow strength from each other to produce pooled effect sizes with smaller variances (Jackson, Riley, and White 2011).

As is the case with meta-regression, the goal of multivariate meta-regression is also to explore and explain the between-study heterogeneity as a function of moderators. Two types of multivariate regression models, fixed-effects and random-effects, are available. A fixed-effects multivariate meta-regression assumes that all heterogeneity between study effect sizes can be accounted for by the included moderators. A random-effects multivariate meta-regression accounts for potential additional variability unexplained by the included moderators, also known as residual heterogeneity.

`meta mvregress` fits multivariate meta-regression. The default model assumed by `meta mvregress` is a random-effects model using the REML method with an unstructured between-study covariance matrix. Use the `random()` option to specify other random-effects methods such as the MLE or a noniterative Jackson–White–Riley method, which can be viewed as an extension of the univariate DerSimonian–Laird method to the multivariate setting. You may also use the `random()` option to specify an alternative covariance structure such as `exchangeable`, `independent`, `identity`, or `fixed()` in the `covariance()` suboption.

Covariance structure `fixed()` specifies a fixed between-study covariance matrix and thus can be used to perform sensitivity analysis similarly to option `tau2()` in [META] **meta regress**. Specifying a covariance structure other than the default `unstructured` is particularly useful when the number of observations, n, is small relative to the number of estimated fixed-effects parameters and variance components.

Jackson and Riley (2014) proposed an adjustment to the standard errors of the fixed-effects parameters that provides more accurate inference when the number of studies is relatively small. This adjustment is available with the `se()` option. The Jackson–Riley adjustment can be seen as a multivariate extension of the Knapp–Hartung adjustment (Knapp and Hartung 2003) in univariate meta-regression, and the two adjustments are identical when there is only one effect-size variable.

Consider data from K independent studies and d outcomes (effect sizes). Let $\widehat{\theta}_{ij}$ be the estimated effect size reported by study j for outcome i, and let a $d \times 1$ vector $\widehat{\boldsymbol{\theta}}_j = (\widehat{\theta}_{1j}, \widehat{\theta}_{2j}, \ldots, \widehat{\theta}_{dj})'$ be an estimate of the true population effect size $\boldsymbol{\theta}_j$ for study j.

A model for the fixed-effects multivariate meta-regression (Raudenbush, Becker, and Kalaian 1988) can be expressed as

$$\widehat{\theta}_{ij} = \beta_{i0} + \beta_{i1}x_{1j} + \cdots + \beta_{i,p-1}x_{p-1,j} + \epsilon_{ij} = \mathbf{x}_j\boldsymbol{\beta}_i + \epsilon_{ij}$$

for outcome $i = 1, \ldots, d$ and study $j = 1, \ldots, K$. Here $\mathbf{x}_j = (1, x_{1j}, \ldots, x_{p-1,j})$ is a $1 \times p$ vector of categorical and continuous moderators (covariates), $\boldsymbol{\beta}_i$ is an outcome-specific $p \times 1$ vector of unknown regression coefficients, and $\boldsymbol{\epsilon}_j = (\epsilon_{1j}, \epsilon_{2j}, \ldots, \epsilon_{dj})'$ is a $d \times 1$ vector of within-study errors that have a d-variate normal distribution with zero mean vector and a $d \times d$ covariance matrix $\text{Var}(\boldsymbol{\epsilon}_j) = \boldsymbol{\Lambda}_j$. The within-study covariance matrices $\boldsymbol{\Lambda}_j$'s are treated as known and do not require estimation. The values of these matrices are specified as variables in the `wcovvariables()` option or in a combination of the `wsevariables()` and `wcorrelations()` options.

In a matrix notation, the above fixed-effects model can be defined as

$$\widehat{\boldsymbol{\theta}}_j = \mathbf{X}_j\boldsymbol{\beta} + \boldsymbol{\epsilon}_j, \quad \boldsymbol{\epsilon}_j \sim N_d(\mathbf{0}, \boldsymbol{\Lambda}_j)$$

where $\mathbf{X}_j = \mathbf{x}_j \otimes I_d$ ($\otimes$ is the Kronecker product) is a $d \times dp$ matrix and $\boldsymbol{\beta} = (\boldsymbol{\beta}_1', \boldsymbol{\beta}_2', \ldots, \boldsymbol{\beta}_d')'$ is a $dp \times 1$ vector of all unknown regression coefficients.

Residual heterogeneity may be accounted for by including an additive between-study covariance component, $\boldsymbol{\Sigma}$, that leads to a random-effects multivariate meta-regression (Berkey et al. 1998):

$$\widehat{\boldsymbol{\theta}}_j = \mathbf{X}_j\boldsymbol{\beta} + \boldsymbol{\epsilon}_j^* = \mathbf{X}_j\boldsymbol{\beta} + \mathbf{u}_j + \boldsymbol{\epsilon}_j, \quad \text{where } \boldsymbol{\epsilon}_j^* \sim N_d\left(\mathbf{0}, \boldsymbol{\Lambda}_j + \boldsymbol{\Sigma}\right)$$

As we mentioned earlier, a random-effects multivariate meta-regression assumes that the moderators explain only part of heterogeneity, and random effects $\mathbf{u}_j = (u_{1j}, u_{2j}, \ldots, u_{dj})' \sim N_d(\mathbf{0}, \boldsymbol{\Sigma})$ ($j = 1, \ldots, K$) account for the remainder.

Harbord and Higgins (2016) point out that some authors (Thompson and Sharp 1999; Higgins and Thompson 2004) argue that a fixed-effects meta-regression should not be used because, in practice, the included moderators rarely capture all the between-study heterogeneity and that the failure of the fixed-effects regression to capture the extra between-study heterogeneity can lead to excessive type I errors. This observation is also echoed by Jackson, Riley, and White (2011) in the multivariate setting.

Examples of using meta mvregress

Examples are presented under the following headings:

▷ Example 1: Univariate versus multivariate meta-analysis

Consider a dataset from Antczak-Bouckoms et al. (1993) of five randomized controlled trials that explored the impact of two procedures (surgical and nonsurgical) for treating periodontal disease. This dataset was also analyzed by Berkey et al. (1998).

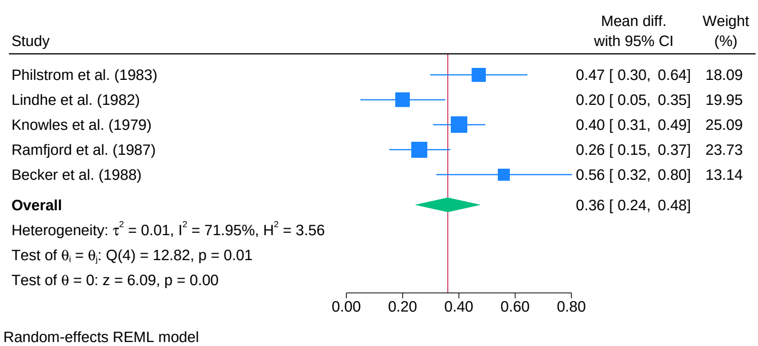
In these trials, subjects' mouths were split into sections. These sections were randomly allocated to the two treatment procedures. At least one section was treated surgically and at least one other section was treated nonsurgically for each patient. The main objectives of the periodontal treatment were to reduce probing depths and increase attachment levels (Berkey et al. 1998).

Two outcomes of interest are improvements from baseline (pretreatment) in probing depth (`y1`) and attachment level (`y2`) around the teeth. Because the two outcomes `y1` and `y2` are measured on the same subject, they should not be treated as independent. This is an example of multiple-endpoint studies where multiple outcomes (two in this case) are compared across two groups (surgical versus nonsurgical). We first describe our dataset.

```
. use https://www.stata-press.com/data/r19/periodontal
(Treatment of moderate periodontal disease)

. describe

Contains data from https://www.stata-press.com/data/r19/periodontal.dta
 Observations:             5                  Treatment of moderate
                                                periodontal disease
    Variables:             9                  13 Jan 2025 18:11
                                              (_dta has notes)
────────────────────────────────────────────────────────────────────────────────
Variable      Storage   Display    Value
    name         type    format    label      Variable label
────────────────────────────────────────────────────────────────────────────────
trial           str23   %23s                  Trial label
pubyear         byte    %9.0g                 Publication year centered at 1983
y1              float   %6.2f                 Mean improvement in probing depth
                                                (mm)
y2              float   %6.2f                 Mean improvement in attachment
                                                level (mm)
v11             float   %6.4f                 Variance of y1
v12             float   %6.4f                 Covariance of y1 and y2
v22             float   %6.4f                 Variance of y2
s1              double  %10.0g                Standard error of y1
s2              double  %10.0g                Standard error of y2
────────────────────────────────────────────────────────────────────────────────
Sorted by:
```

We will start by performing a separate meta-analysis for each outcome. We declare our data as meta data using the `meta set` command and then construct a forest plot for each outcome; see [META] **meta set** and [META] **meta forestplot**, respectively.

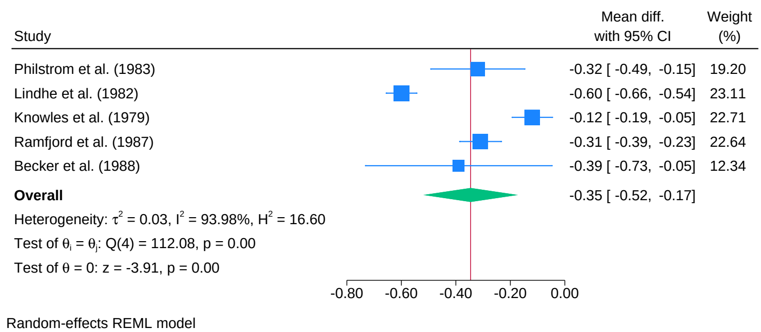
```
. quietly meta set y1 s1, studylabel(trial) eslabel("Mean diff.")

. meta forestplot, esrefline

  Effect-size label: Mean diff.
        Effect size: y1
          Std. err.: s1
        Study label: trial
```



Positive `y1` values indicate that the mean improvement (reduction) in probing depth for the surgical group is larger than that for the nonsurgical group. It appears that the surgical treatment performs consistently better ($\texttt{y1} > 0$) across all studies. The overall mean difference is 0.36 with a 95% CI of $[0.24, 0.48]$, which means that, on average, the reduction in probing depth was 0.36 mm higher than that for the nonsurgical group.

Similarly, we will construct a forest plot for variable y2.

```
. quietly meta set y2 s2, studylabel(trial) eslabel("Mean diff.")
. meta forestplot, esrefline
  Effect-size label: Mean diff.
        Effect size: y2
          Std. err.: s2
        Study label: trial
```



Negative y2 values indicate that the mean improvement (increase) in attachment level for the surgical group is smaller than that for the nonsurgical group. Because y2 < 0 across all studies, the nonsurgical treatment performs consistently better in terms of attachment level. It appears that there is considerable heterogeneity in attachment levels (y2) based on the nonoverlapping CIs in the forest plot and a large value of the I^2 statistic (93.98%).

Notice that the obtained heterogeneity statistics are from univariate meta-analyses conducted separately. In example 4, we show how to assess heterogeneity from a multivariate analysis by using the `estat heterogeneity` command.

The two separate meta-analyses do not account for the dependence between y1 and y2. Let's fit a bivariate meta-analysis (constant-only bivariate meta-regression) using the meta mvregress command.

```
. meta mvregress y1 y2, wcovvariables(v11 v12 v22)
Performing EM optimization ...
Performing gradient-based optimization:
Iteration 0:  Log restricted-likelihood =  2.0594015
Iteration 1:  Log restricted-likelihood =  2.0822862
Iteration 2:  Log restricted-likelihood =  2.0823276
Iteration 3:  Log restricted-likelihood =  2.0823276
Multivariate random-effects meta-analysis      Number of obs     =         10
Method: REML                                   Number of studies =          5
                                               Obs per study:
                                                             min =          2
                                                             avg =        2.0
                                                             max =          2
                                               Wald chi2(0)      =          .
Log restricted-likelihood = 2.0823276          Prob > chi2       =          .

------------------------------------------------------------------------------
             | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
y1           |
       _cons |   .3534282   .0588486     6.01   0.000      .238087    .4687694
-------------+----------------------------------------------------------------
y2           |
       _cons |  -.3392152   .0879051    -3.86   0.000    -.5115061   -.1669243
------------------------------------------------------------------------------
Test of homogeneity: Q_M = chi2(8) = 128.23               Prob > Q_M = 0.0000

------------------------------------------
  Random-effects parameters |   Estimate
----------------------------+-------------
Unstructured:               |
                      sd(y1)|   .1083191
                      sd(y2)|   .1806968
                 corr(y1,y2)|   .6087987
------------------------------------------
```

The output shows information about the optimization algorithm, the iteration log, and the model (random-effects) and method (REML) used for estimation. It also displays the number of studies, $K = 5$, and the total number of observations on the outcomes, $n = 10$, which is equal to Kd because there are no missing observations. The minimum, maximum, and average numbers of observations per study are also reported. Because there were no missing observations, all of these numbers are identical and equal to 2.

The first table displays the regression (fixed-effects) coefficient estimates from the bivariate meta-analysis. These estimates correspond to the overall bivariate effect size $\widehat{\boldsymbol{\theta}} = (\widehat{\theta}_1, \widehat{\theta}_2)'$. The estimates are close to the univariate ones reported on the forest plots. But from a bivariate analysis, we obtained slightly narrower 95% CIs for the overall effect sizes. The multivariate homogeneity test, which tests whether $\boldsymbol{\theta}_j = (\theta_{1j}, \theta_{2j})'$ is constant across studies, is rejected ($p < 0.0001$). This agrees with earlier univariate results, particularly from the second forest plot, which exhibited considerable heterogeneity.

The second table displays the random-effects parameters, traditionally known as variance components in the context of multilevel or mixed-effects models. By default, similar to the mixed command, meta mvregress reports standard deviations of y1 and y2 and their correlation: sd(y1), sd(y2), and corr(y1,y2), respectively. But you can instead specify the variance option to report variances and the covariance.

◁

▷ Example 2: Random-effects multivariate meta-regression

Berkey et al. (1998) noted that although the meta-analysis of Antczak-Bouckoms et al. (1993) accounted for many factors that could potentially lead to heterogeneity, a substantive variability was still present, as we highlighted in example 1. They suggested to use the year of publication centered at 1983 (pubyear), a surrogate for the time when the trial was performed, as a moderator to explain a portion of this heterogeneity. They reasoned that as the surgical experience accumulates, the surgical procedure will become more efficient so the most recent studies may show greater surgical benefits.

Let's first perform separate univariate meta-regressions for outcomes y1 and y2 with pubyear as a moderator. We can do this by specifying only one dependent variable with meta mvregress or by using meta regress. We will use meta mvregress because it does not require setting the data.

```
. meta mvregress y1 = pubyear, wsevariables(s1)
Performing EM optimization ...
Performing gradient-based optimization:
Iteration 0:  Log restricted-likelihood = -1.6637351
Iteration 1:  Log restricted-likelihood = -1.6426005
Iteration 2:  Log restricted-likelihood = -1.6414308
Iteration 3:  Log restricted-likelihood = -1.6414292
Iteration 4:  Log restricted-likelihood = -1.6414292

Multivariate random-effects meta-regression     Number of obs     =          5
Method: REML                                    Number of studies =          5
                                                Obs per study:
                                                              min =          1
                                                              avg =        1.0
                                                              max =          1
                                                Wald chi2(1)      =       0.04
Log restricted-likelihood = -1.6414292          Prob > chi2       =     0.8332

------------------------------------------------------------------------------
          y1 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
     pubyear |    .004542    .021569     0.21   0.833    -.0377325    .0468165
       _cons |    .362598   .0725013     5.00   0.000     .2204981     .504698
------------------------------------------------------------------------------
Test of homogeneity: Q_M = chi2(3) = 11.80                Prob > Q_M = 0.0081

------------------------------------------
  Random-effects parameters  |   Estimate
-----------------------------+------------
Identity:                    |
                       sd(y1) |   .1406077
------------------------------------------
```

```
. meta mvregress y2 = pubyear, wsevariables(s2)
Performing EM optimization ...
Performing gradient-based optimization:
Iteration 0:  Log restricted-likelihood = -2.4661957
Iteration 1:  Log restricted-likelihood = -2.3230318
Iteration 2:  Log restricted-likelihood = -2.3229928
Iteration 3:  Log restricted-likelihood = -2.3229928
Multivariate random-effects meta-regression     Number of obs     =          5
Method: REML                                    Number of studies =          5
                                                Obs per study:
                                                              min =          1
                                                              avg =        1.0
                                                              max =          1
                                                Wald chi2(1)      =       0.20
Log restricted-likelihood = -2.3229928          Prob > chi2       =     0.6524

------------------------------------------------------------------------------
          y2 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
     pubyear |  -.0134909   .0299534    -0.45   0.652    -.0721985    .0452167
       _cons |  -.3399793   .0978864    -3.47   0.001    -.5318331   -.1481256
------------------------------------------------------------------------------
Test of homogeneity: Q_M = chi2(3) = 108.29                 Prob > Q_M = 0.0000

------------------------------------------
  Random-effects parameters |   Estimate
----------------------------+-------------
Identity:                   |
                     sd(y2) |    .201787
------------------------------------------
```

Here we specified the standard error variables `s1` and `s2` in the `wsevariables()` options to match the univariate setup more closely, but we could have used `wcovvariables(v11)` and `wcovvariables(v22)`, following example 1.

Results from the univariate meta-regressions suggest that variable `pubyear` does not seem to explain the between-study heterogeneity between the effect sizes `y1` and `y2`; the p-values for testing the `pubyear` coefficients to be 0 are $p = 0.833$ and $p = 0.652$, respectively.

The two separate meta-regressions do not account for the dependence between y1 and y2. Below, we fit a bivariate meta-regression that accounts for this dependence.

```
. meta mvregress y1 y2 = pubyear, wcovvariables(v*)
Performing EM optimization ...
Performing gradient-based optimization:
Iteration 0:  Log restricted-likelihood = -3.5544446
Iteration 1:  Log restricted-likelihood = -3.5402086
Iteration 2:  Log restricted-likelihood = -3.5399568
Iteration 3:  Log restricted-likelihood = -3.5399567
Multivariate random-effects meta-regression     Number of obs     =         10
Method: REML                                    Number of studies =          5
                                                Obs per study:
                                                              min =          2
                                                              avg =        2.0
                                                              max =          2
                                                Wald chi2(2)      =       0.40
Log restricted-likelihood = -3.5399567          Prob > chi2       =     0.8197

             | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
y1           |
     pubyear |   .0048615   .0218511     0.22   0.824    -.0379658    .0476888
       _cons |   .3587569     .07345     4.88   0.000     .2147975    .5027163
-------------+----------------------------------------------------------------
y2           |
     pubyear |  -.0115367   .0299635    -0.39   0.700     -.070264    .0471907
       _cons |  -.3357368   .0979979    -3.43   0.001    -.5278091   -.1436645
------------------------------------------------------------------------------
Test of homogeneity: Q_M = chi2(6) = 125.76               Prob > Q_M = 0.0000

  Random-effects parameters |   Estimate
----------------------------+------------
Unstructured:               |
                      sd(y1)|   .1429917
                      sd(y2)|   .2021314
                 corr(y1,y2)|    .561385
----------------------------+------------
```

Instead of listing all the variance–covariance variables v11, v12, and v22 in the wcovvariables() option, we used the stub notation v* to refer to all of them. This notation is especially convenient for models with more dependent variables. You just need to make sure that these are the only variables starting with v in the dataset and that the variables are properly ordered (think of a vectorized upper triangle of the variance–covariance matrix) before using the stub notation; see the description of wcovvariables().

The estimates of the regression coefficients of variable pubyear are 0.0049 with a 95% CI of $[-0.0380, 0.0477]$ for outcome y1 and -0.0115 with a 95% CI of $[-0.0703, 0.0472]$) for outcome y2. The coefficients are not significant according to the z tests with the respective p-values $p = 0.824$ and $p = 0.7$.

Although pubyear did not explain the between-study heterogeneity, we continue to include it as a moderator in our subsequent examples (example 3–example 6) for illustration purposes.

◁

▷ Example 3: Identical results from univariate and multivariate analyses

At this point, it may be interesting to explore when the results from a multivariate meta-regression can match the results of separate univariate meta-analyses. Theoretically, if the within-study covariances (and thus correlations) in $\mathbf{\Lambda}_j$ are equal to 0 and the between-study covariances in $\mathbf{\Sigma}$ are also equal to 0, then performing a multivariate meta-regression is equivalent to performing separate univariate meta-regressions for each outcome.

Continuing with example 2, we specify the `wsevariables(s1 s2)` and `wcorrelations(0)` options to assume there is no within-study correlation between `y1` and `y2`. We also assume that the between-study covariances are 0 by specifying an independent covariance structure for the random effects with the `covariance(independent)` suboption of the `random()` option.

```
. meta mvregress y1 y2 = pubyear, wsevariables(s1 s2) wcorrelations(0)
> random(reml, covariance(independent))
Performing EM optimization ...
Performing gradient-based optimization:
Iteration 0:  Log restricted-likelihood = -3.9946242
Iteration 1:  Log restricted-likelihood = -3.9656463
Iteration 2:  Log restricted-likelihood = -3.9644233
Iteration 3:  Log restricted-likelihood =  -3.964422
Iteration 4:  Log restricted-likelihood =  -3.964422
Multivariate random-effects meta-regression     Number of obs     =         10
Method: REML                                    Number of studies =          5
                                                Obs per study:
                                                              min =          2
                                                              avg =        2.0
                                                              max =          2
                                                Wald chi2(2)      =       0.25
Log restricted-likelihood = -3.964422           Prob > chi2       =     0.8837
```

	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
y1						
pubyear	.004542	.021569	0.21	0.833	-.0377325	.0468165
_cons	.362598	.0725013	5.00	0.000	.2204981	.504698
y2						
pubyear	-.0134909	.0299534	-0.45	0.652	-.0721985	.0452167
_cons	-.3399793	.0978864	-3.47	0.001	-.5318331	-.1481256

```
Test of homogeneity: Q_M = chi2(6) = 120.10              Prob > Q_M = 0.0000
```

Random-effects parameters	Estimate
Independent:	
sd(y1)	.1406077
sd(y2)	.201787

The results for regression coefficients and variance components are identical to those from separate univariate meta-regressions in example 2. Note that the multivariate homogeneity statistic $Q_{\rm M} = 120.10$ is the sum of the univariate statistics $Q_{\rm M} = Q_{\rm res} = 11.8$ and $Q_{\rm M} = Q_{\rm res} = 108.3$, where $Q_{\rm res}$ is the univariate version of $Q_{\rm M}$ defined in [META] **meta regress**.

◁

▷ Example 4: Heterogeneity statistics

Continuing with example 2, let's refit the model and use the postestimation command `estat heterogeneity` to quantify heterogeneity after the bivariate meta-regression. Assessing the residual between-study variability is important in the context of random-effects multivariate meta-regression, so we will discuss various heterogeneity measures in detail in this example.

```
. quietly meta mvregress y1 y2 = pubyear, wcovvariables(v*)
. estat heterogeneity
Method: Cochran
Joint:
  I2 (%) = 95.23
      H2 = 20.96
Method: Jackson-White-Riley
y1:
  I2 (%) = 85.26
       R =  2.60
y2:
  I2 (%) = 95.85
       R =  4.91
Joint:
  I2 (%) = 91.57
       R =  3.44
```

By default, the Cochran and Jackson–White–Riley heterogeneity statistics are reported, but the White heterogeneity statistic is also available, as we demonstrate later in this example.

Cochran $I_{\rm Q}^2$ and $H_{\rm Q}^2$ are direct extensions to the multivariate setting of the univariate I^2 and H^2 statistics based on the DerSimonian–Laird method and thus have the same interpretations; see *Heterogeneity measures* in *Methods and formulas* in [META] **meta summarize** and *Residual heterogeneity measures* in *Methods and formulas* in [META] **meta regress**. For instance, $I_{\rm Q}^2 = 95.23\%$ means that 95.23% of the residual heterogeneity, heterogeneity not accounted for by the moderator `pubyear`, is due to true heterogeneity between the studies as opposed to the sampling variability. The high value for this statistic is not surprising because, as we saw in example 2, `pubyear` did not explain much heterogeneity between the studies.

The values of Cochran statistics are the same for all random-effects methods because they are based on the Cochran multivariate Q statistic, which is calculated based on the fixed-effects model; see *Cochran heterogeneity statistics* in *Methods and formulas* in [META] **estat heterogeneity (mv)** for details. One potential shortcoming of the Cochran statistics is that they quantify the amount of heterogeneity jointly for all outcomes. The Jackson–White–Riley statistics (Jackson, White, and Riley 2012) provide ways to assess the contribution of each outcome to the total heterogeneity, in addition to their joint contribution.

You can also investigate the impact of any subset of outcomes on heterogeneity by specifying the subset of outcomes in the `jwriley()` option of `estat heterogeneity`; see example 1 of [META] **estat heterogeneity (mv)**. These statistics are also the only truly multivariate heterogeneity statistics in the sense that their definitions stem from purely multivariate concepts rather than from univariate concepts applied to the multivariate setting.

The Jackson–White–Riley statistics measure the variability of the random-effects estimator relative to the fixed-effects estimator. The larger the values, the more between-study heterogeneity is left unexplained after accounting for moderators. The $R_{\rm JWR}$ statistic is an absolute measure ($R_{\rm JWR} \geq 1$), and

I^2_{JWR} is defined based on R_{JWR} as a percentage increase in the variability of the random-effects estimates relative to the fixed-effects estimates; see *Jackson–White–Riley heterogeneity statistics* in *Methods and formulas* in [META] **estat heterogeneity (mv)** for technical details.

$R_{\text{JWR}} = 1$, and consequently $I^2_{\text{JWR}} = 0\%$, means that the moderators have accounted for all the heterogeneity between the effect sizes, and therefore there is no difference between the random-effects and fixed-effects models. Values of I^2_{JWR} that are close to 100% mean that considerable residual heterogeneity is still present in the model so that the random-effects model is more appropriate. In our example, for instance, for outcome `y1`, $R_{\text{JWR}} = 2.6$, and the corresponding $I^2_{\text{JWR}} = 85.26\% > 75\%$, which suggests "large heterogeneity" according to Higgins et al. (2003).

Other multivariate extensions of the I^2 heterogeneity statistic have also been used in practice. For example, the White I^2 statistic (White 2011) can be computed by using the `white` option.

```
. estat heterogeneity, white
Method: White
y1:
  I2 (%) = 77.26

y2:
  I2 (%) = 94.32
```

The White I^2 statistic is a direct extension of the univariate I^2 statistic (*Residual heterogeneity measures* in *Methods and formulas* in [META] **meta regress**), except the estimated between-study variance $\hat{\tau}^2$ is replaced by a diagonal of the estimated between-study covariance matrix, $\widehat{\boldsymbol{\Sigma}}$. It has the same interpretation as the univariate I^2 and reduces to it when there is only one dependent variable.

Unlike the Cochran and Jackson–White–Riley statistics that can assess heterogeneity jointly for all outcomes, the White statistic can only quantify heterogeneity separately for each outcome; see table 1 in [META] **estat heterogeneity (mv)**. In our example, continuing with outcome `y1`, we see that $I^2_{\text{W}} = 77.26\% > 75\%$ also reports the presence of a large between-study variability for that outcome even after accounting for `pubyear`.

◁

❑ Technical note

The actual definition for the Jackson–White–Riley R_{JWR} statistic is somewhat technical. It is easier to think about it first in the univariate setting, where it is defined as the ratio of the widths of the CIs of the random-effects estimator for the regression coefficient vector to the corresponding fixed-effects estimator raised to the power of $1/2p$. In the multivariate setting, the widths of confidence intervals become areas or volumes of confidence regions, and the power becomes $1/2pd$.

For example, for outcome `y1`, $d = 1$, $p = 2$, and $\hat{\beta}_{01}$ and $\hat{\beta}_{11}$ are the estimates of the constant and the regression coefficient for `pubyear`. Then, $R_{\text{JWR}} = 2.6$ is the ratio, raised to the power of $1/4$, of the areas of the confidence regions (ellipses) for estimates $\hat{\beta}_{01}$ and $\hat{\beta}_{11}$ under the random-effects and fixed-effects multivariate meta-regressions. This ratio is greater than 1 because the area of the confidence region under the random-effects model is larger.

The $I^2_{\text{JWR}} = 85.26\%$ for outcome `y1` is interpreted as roughly an 85% increase in the area of the confidence regions for the random-effects estimator of β_{01} and β_{11} relative to the fixed-effects estimator. See Jackson, White, and Riley (2012) for more ways of interpreting the I^2_{JWR} statistic in terms of generalized variances and geometric means.

Note that with three- and higher-dimensional models, the areas of confidence regions become volumes, and the shapes of confidence regions become ellipsoids.

❑

▷ Example 5: Jackson–White–Riley random-effects method

Continuing with example 2, we demonstrate the use of an alternative random-effects estimation method, the Jackson–White–Riley method, instead of the default REML method. This method is a multivariate extension of the popular univariate DerSimonian–Laird method.

```
. meta mvregress y1 y2 = pubyear, wcovvariables(v*) random(jwriley)
Multivariate random-effects meta-regression      Number of obs      =         10
Method: Jackson-White-Riley                      Number of studies =          5
                                                 Obs per study:
                                                               min =          2
                                                               avg =        2.0
                                                               max =          2
                                                 Wald chi2(2)       =       0.30
                                                 Prob > chi2        =     0.8621

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

y1
     pubyear    .0046544    .023268     0.20   0.841      -.04095    .0502588
       _cons     .358993   .0783252     4.58   0.000     .2054784    .5125075

y2
     pubyear   -.0117463   .0419197    -0.28   0.779    -.0939074    .0704147
       _cons    -.335579   .1393286    -2.41   0.016     -.608658   -.0624999

Test of homogeneity: Q_M = chi2(6) = 125.76              Prob > Q_M = 0.0000

  Random-effects parameters     Estimate

Unstructured:
                       sd(y1)    .1547229
                       sd(y2)    .2947281
                  corr(y1,y2)    .6518347
```

The estimates of the regression coefficients are very similar to those from example 2 using the REML method. For instance, the coefficient of pubyear for outcome y1 is 0.0049 and is similar to the REML estimate of 0.0047. The standard errors and estimates of variance components are larger than those obtained from the REML estimation. This is because REML assumes normality and, when this assumption is satisfied, it is likely to produce more efficient estimates than a method of moments estimator such as the Jackson–White–Riley.

◁

▷ Example 6: Jackson–Riley standard-error adjustment

Jackson and Riley (2014) proposed a multivariate extension of the univariate Knapp and Hartung (2003) standard-error adjustment that provides more accurate inference for the regression coefficients when the number of studies is small (as is the case in our example where $K = 5$).

Continuing with example 2, we compute the Jackson–Riley standard-error adjustment by specifying the `se(jriley)` suboption within `random()`.

```
. meta mvregress y1 y2 = pubyear, wcovvariables(v*) random(reml, se(jriley))
Performing EM optimization ...
Performing gradient-based optimization:
Iteration 0:  Log restricted-likelihood = -3.5544446
Iteration 1:  Log restricted-likelihood = -3.5402086
Iteration 2:  Log restricted-likelihood = -3.5399568
Iteration 3:  Log restricted-likelihood = -3.5399567
Multivariate random-effects meta-regression      Number of obs     =         10
Method: REML                                     Number of studies =          5
SE adjustment: Jackson-Riley                     Obs per study:
                                                               min =          2
                                                               avg =        2.0
                                                               max =          2
                                                 F(2,      6.00)   =       0.20
Log restricted-likelihood = -3.5399567           Prob > F          =     0.8249

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

y1
         pubyear      .0048615    .021313     0.23   0.827    -.0472895    .0570124
           _cons      .3587569   .0716413     5.01   0.002      .183457    .5340569

y2
         pubyear     -.0115367   .0292256    -0.39   0.707    -.0830492    .0599758
           _cons     -.3357368   .0955846    -3.51   0.013     -.569624   -.1018496

Test of homogeneity: Q_M = chi2(6) = 125.76               Prob > Q_M = 0.0000

   Random-effects parameters     Estimate

Unstructured:
                        sd(y1)    .1429917
                        sd(y2)    .2021314
                   corr(y1,y2)     .561385
```

The regression coefficients and variance components are identical to those in example 2. But the standard errors of the regression coefficients have been adjusted; see *Jackson–Riley standard-error adjustment* in *Methods and formulas* below. The tests of the regression coefficients and the model test now use the Student's t and F distributions, respectively, instead of the default normal and χ^2 distributions.

Another standard error adjustment that is used in practice is the truncated Jackson–Riley adjustment, which may be obtained by specifying the `se(truncjriley)` suboption. The Jackson–Riley standard-error adjustment reduces to the Knapp–Hartung adjustment when there is only one dependent variable.

◁

▷ Example 7: When within-study covariances are not available

Glas et al. (2003, table 3) reported a dataset of 10 studies to investigate the sensitivity and specificity of the tumor marker telomerase to diagnose primary bladder cancer. This dataset was also analyzed by Riley et al. (2007) and White (2016). Let's describe our dataset.

```
. use https://www.stata-press.com/data/r19/telomerase
(Telomerase for diagnosing primary bladder cancer)

. describe

Contains data from https://www.stata-press.com/data/r19/telomerase.dta
 Observations:            10                  Telomerase for diagnosing
                                                primary bladder cancer
    Variables:             8                  4 Feb 2025 04:09
                                              (_dta has notes)
--------------------------------------------------------------------------------
Variable      Storage   Display    Value
    name         type    format    label      Variable label
--------------------------------------------------------------------------------
trial           str22   %22s                  Trial label
trialnum        byte    %9.0g                 Trial ID
y1              float   %9.0g                 Logit sensitivity
y2              float   %9.0g                 Logit specificity
s1              float   %9.0g                 Standard error of logit
                                                sensitivity
s2              float   %9.0g                 Standard error of logit
                                                specificity
v1              double  %10.0g                Variance of logit sensitivity
v2              double  %10.0g                Variance of logit specificity
--------------------------------------------------------------------------------
Sorted by:
```

Variables y1 and y2 are logit-transformed sensitivity and specificity for telomerase, and s1 and s2 are the corresponding standard errors.

No within-study covariances are reported for this dataset. When this occurs, one possible approach is to perform a sensitivity analysis (see example 10), where we assess the impact of different magnitudes of correlations on our bivariate meta-analysis results. In our case, sensitivity and specificity are typically measured on independent groups of individuals, so it is reasonable to assume that the within-study correlation is zero between y1 and y2.

We specify the variance option to report variances and covariances of the random effects instead of the default standard deviations and correlations to replicate the results of Riley et al. (2007, table 3), who reported variances of the random effects.

```
. meta mvregress y*, wsevariables(s*) wcorrelation(0) variance
Performing EM optimization ...
Performing gradient-based optimization:
Iteration 0:  Log restricted-likelihood = -28.449202  (not concave)
Iteration 1:  Log restricted-likelihood =  -25.18825
Iteration 2:  Log restricted-likelihood = -24.713278
Iteration 3:  Log restricted-likelihood = -24.609916
Iteration 4:  Log restricted-likelihood = -24.418125
Iteration 5:  Log restricted-likelihood = -24.415969
Iteration 6:  Log restricted-likelihood = -24.415967
Multivariate random-effects meta-analysis        Number of obs     =         20
Method: REML                                     Number of studies =         10
                                                 Obs per study:
                                                              min =          2
                                                              avg =        2.0
                                                              max =          2
                                                 Wald chi2(0)      =          .
Log restricted-likelihood = -24.415967           Prob > chi2       =          .
```

	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
y1						
_cons	1.166189	.1861349	6.27	0.000	.801371	1.531006
y2						
_cons	2.057721	.5534499	3.72	0.000	.9729789	3.142462

```
Test of homogeneity: Q_M = chi2(18) = 90.87             Prob > Q_M = 0.0000
```

Random-effects parameters	Estimate
Unstructured:	
var(y1)	.2022309
var(y2)	2.583339
cov(y1,y2)	-.7227936

Our results match those reported by Riley et al. (2007). The estimated overall sensitivity for `y1` is `invlogit(1.166)` $=$ 76.24 or roughly 76%, and the estimated overall specificity for `y2` is `invlogit(2.058)` $=$ 88.68 or roughly 89%. Glas et al. (2003) noted that the sensitivity of telomerase may not be large enough for clinical use in diagnosing bladder cancer.

Had we not specified the `variance` option and reported the default standard deviations and correlations of the random-effects, we would get `corr(y1,y2)` $= -1$. We can verify this either by typing `meta mvregress` to replace the results or by using the postestimation command `estat sd`. We demonstrate the latter.

```
. estat sd
```

Random-effects parameters	Estimate
Unstructured:	
sd(y1)	.4497009
sd(y2)	1.607277
corr(y1,y2)	-.9999998

Riley et al. (2007) noted that having a between-study correlation of 1 or -1 is common in multivariate meta-analysis when the number of studies is small, especially when the within-study variances are similar to or larger than the corresponding between-study variances. This is the case in our data where, for example, the mean within-study variance for `y1` is 0.18 (for instance, type `summarize v1`), which is comparable with the estimated between-study variance `var(y1)` $= 0.20$. Other random-effects covariance structures should be explored to address correlations of 1 and -1; see example 1 of [META] **meta mvregress postestimation**.

◁

▷ Example 8: Missing outcome data

Fiore et al. (1996) reported a dataset of 24 studies investigating the impact of 4 intervention types to promote smoking cessation. This dataset was also analyzed by Lu and Ades (2006).

The four intervention types are (a) no contact, (b) self-help, (c) individual counseling, and (d) group counseling. The goal is to compare types (b), (c), and (d) with (a). Variables `yb`, `yc`, and `yd` represent the log odds-ratios for types (b), (c), and (d) relative to group (a). The corresponding within-study variances and covariances are reported by the six variables `vbb`, `vbc`, `vbd`, `vcc`, `vcd`, and `vdd`.

An odds ratio greater than 1 (or, equivalently, positive log odds-ratio) means that the odds of quitting smoking are larger in the corresponding group compared with the odds in type (a). This dataset is an example of multiple-treatment studies.

```
. use https://www.stata-press.com/data/r19/smokecess
(Smoking cessation interventions)

. describe y* v*

Variable      Storage   Display    Value
    name         type    format    label      Variable label
-------------------------------------------------------------------------------

yb              double   %9.0g                Log-odds ratio (b vs a)
yc              double   %9.0g                Log-odds ratio (c vs a)
yd              double   %9.0g                Log-odds ratio (d vs a)
vbb             double   %9.0g                Variance of yb
vbc             double   %9.0g                Covariance of yb and yc
vbd             double   %9.0g                Covariance of yb and yd
vcc             double   %9.0g                Variance of yc
vcd             double   %9.0g                Covariance of yc and yd
vdd             double   %9.0g                Variance of yd
```

Let's explore the missing-value structure of this dataset.

```
. misstable pattern y*, frequency

  Missing-value patterns
    (1 means complete)

              |   Pattern
    Frequency |  1  2  3
  ------------+-------------
            1 |  1  1  1
              |
           14 |  1  0  0
            3 |  0  0  1
            3 |  1  1  0
            1 |  0  1  0
            1 |  0  1  1
            1 |  1  0  1
  ------------+-------------
           24 |

  Variables are  (1) yc  (2) yd  (3) yb
```

There are 24 observations, and only 1 contains values for all 3 variables. There is only one observation when both yd and yb and both yc and yb are observed. And variables yd and yb have only six nonmissing values. So, among all variables, there are a total of $72 = 3 \times 24$ values, and only $31 = 72 - (14 \times 2 + 3 \times 2 + 3 + 2 + 1 + 1)$ of them are not missing. Given how small and sparse these data are, we can anticipate that the joint estimation of these variables will be challenging without additional, potentially strong, assumptions about the data.

In fact, if we try to run the following model, where for demonstration we use the ML method,

```
. meta mvregress yb yc yd, wcovvariables(vbb vbc vbd vcc vcd vdd) random(mle)
  (output omitted)
```

we will obtain a correlation between the random effects associated with outcomes yb and yd, `corr(yb,yd)`, close to 1. This is because only 2 out of the 24 studies have observations on both of the outcomes (type `misstable pattern yb yd, frequency`), which makes the estimation of `corr(yb,yd)` unstable and inaccurate. Also, the between-study covariance structure may be overparameterized given how sparse the data are.

Note that `meta mvregress` uses all available data (all 31 nonmissing values in our example) and not just complete observations. It produces valid results under the assumption that the missing observations are missing at random.

The first model we ran assumed an unrestricted (`unstructured`) between-study covariance for yb, yc, and yd. Let's simplify this assumption and assume an independent covariance structure to reduce the number of estimated variance components. Also, whenever a large portion of the observations is missing, as in our example, parameter estimates tend to be less accurate. We thus specify the `cformat(%9.3f)` option to display results up to three decimal points.

```
. meta mvregress y*, wcovvariables(v*) random(mle, covariance(independent))
> cformat(%9.3f)
Performing EM optimization ...
Performing gradient-based optimization:
Iteration 0:  Log likelihood = -71.117927  (not concave)
Iteration 1:  Log likelihood =  -57.19315  (not concave)
Iteration 2:  Log likelihood = -53.591501
Iteration 3:  Log likelihood = -52.323504
Iteration 4:  Log likelihood = -52.108529
Iteration 5:  Log likelihood = -52.106793
Iteration 6:  Log likelihood = -52.106792
Multivariate random-effects meta-analysis        Number of obs     =         31
Method: ML                                       Number of studies =         24
                                                 Obs per study:
                                                              min =          1
                                                              avg =        1.3
                                                              max =          3
                                                 Wald chi2(0)      =          .
Log likelihood = -52.106792                      Prob > chi2       =          .

             |  Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
yb           |
       _cons |      0.147      0.135     1.09   0.274      -0.116       0.411
-------------+----------------------------------------------------------------
yc           |
       _cons |      0.649      0.193     3.36   0.001       0.270       1.027
-------------+----------------------------------------------------------------
yd           |
       _cons |      0.663      0.243     2.72   0.006       0.186       1.140
------------------------------------------------------------------------------
Test of homogeneity: Q_M = chi2(28) = 204.22              Prob > Q_M = 0.0000

------------------------------------------
  Random-effects parameters |   Estimate
----------------------------+-------------
Independent:                |
                     sd(yb) |      0.000
                     sd(yc) |      0.694
                     sd(yd) |      0.092
------------------------------------------
```

All the regression coefficient estimates are positive, which means that all interventions are better than intervention (a), although without statistical significance for outcome yb. Parameter sd(yb) is close to 0, which means that the between-study covariance may still be overparameterized. In example 9 below, we will demonstrate alternative random-effects covariance structures that further restrict the between-study covariance structure.

◁

▷ Example 9: Between-study covariance structures

Continuing with example 8, we further reduce the number of variance components to be estimated by specifying a more restrictive between-study covariance structure than covariance(independent). One such structure is identity, where we assume that all random effects are uncorrelated and have one common variance, which is to be estimated.

```
. meta mvregress y*, wcovvariables(v*) random(mle, covariance(identity))
> cformat(%9.3f)

Performing EM optimization ...

Performing gradient-based optimization:
Iteration 0:  Log likelihood = -62.707676  (not concave)
Iteration 1:  Log likelihood = -54.538092
Iteration 2:  Log likelihood = -54.501914
Iteration 3:  Log likelihood = -54.501897
Iteration 4:  Log likelihood = -54.501897

Multivariate random-effects meta-analysis        Number of obs     =        31
Method: ML                                       Number of studies =        24
                                                 Obs per study:
                                                               min =         1
                                                               avg =       1.3
                                                               max =         3
                                                 Wald chi2(0)      =         .
Log likelihood = -54.501897                      Prob > chi2       =         .

------------------------------------------------------------------------------
             |  Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
yb           |
       _cons |      0.367      0.317     1.16   0.247       -0.254       0.988
-------------+----------------------------------------------------------------
yc           |
       _cons |      0.674      0.176     3.83   0.000        0.329       1.019
-------------+----------------------------------------------------------------
yd           |
       _cons |      0.864      0.396     2.18   0.029        0.087       1.641
------------------------------------------------------------------------------
Test of homogeneity: Q_M = chi2(28) = 204.22              Prob > Q_M = 0.0000

-----------------------------------------
  Random-effects parameters  |   Estimate
-----------------------------+-----------
Identity:                    |
                sd(yb yc yd) |      0.580
-----------------------------------------
```

The random-effects (or between-study) covariance structure is now labeled `Identity:`, and the common standard deviation is labeled as `sd(yb yc yd)` and is equal to 0.580. Notice how sensitive the regression coefficient estimates are to the choice of the between-study covariance structure. This phenomenon is a consequence of many missing values in the data. In this case, it is important to also explore univariate results by performing meta-analysis separately for each outcome.

We can also assume that all random effects have the same correlation and the same variance by specifying the `exchangeable` covariance structure.

```
. meta mvregress y*, wcovvariables(v*) random(mle, covariance(exchangeable))
> cformat(%9.3f)
Performing EM optimization ...
Performing gradient-based optimization:
Iteration 0:  Log likelihood = -65.135877  (not concave)
Iteration 1:  Log likelihood = -54.442273  (not concave)
Iteration 2:  Log likelihood = -53.488791
Iteration 3:  Log likelihood = -53.376428
Iteration 4:  Log likelihood =  -53.35636
Iteration 5:  Log likelihood = -53.356319
Iteration 6:  Log likelihood = -53.356319
Multivariate random-effects meta-analysis        Number of obs     =         31
Method: ML                                       Number of studies =         24
                                                 Obs per study:
                                                               min =          1
                                                               avg =        1.3
                                                               max =          3
                                                 Wald chi2(0)      =          .
Log likelihood = -53.356319                      Prob > chi2       =          .
```

	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
yb						
_cons	0.413	0.296	1.40	0.162	-0.166	0.992
yc						
_cons	0.705	0.193	3.66	0.000	0.327	1.082
yd						
_cons	0.837	0.308	2.71	0.007	0.232	1.441

```
Test of homogeneity: Q_M = chi2(28) = 204.22            Prob > Q_M = 0.0000
```

Random-effects parameters	Estimate
Exchangeable:	
sd(yb yc yd)	0.672
corr(yb yc yd)	0.817

The common correlation is labeled as `corr(yb yc yd)` with an estimated value of 0.817, and the common standard deviation, `sd(yb yc yd)`, is estimated to be 0.672.

`meta mvregress` lists only the estimated variance components. If you would like to see the full between-study covariance matrix, you can use the `estat recovariance` command.

```
. estat recovariance
Between-study covariance matrix
```

	yb	yc	yd
yb	.451656		
yc	.3690338	.451656	
yd	.3690338	.3690338	.451656

To see the corresponding correlation matrix, you can specify the `correlation` option.

◁

▷ Example 10: Sensitivity meta-analysis

It is quite common in multivariate meta-regression to produce unstable estimates, especially when the number of observations is small relative to the number of parameters to be estimated or when a relatively large portion of the observations is missing. In this case, our goal may shift toward assessing the impact of different magnitudes of between-study variances and covariances on the estimates of regression coefficients.

Continuing with the dataset in example 8, we can investigate the effect of no correlation, moderate correlation (0.4), and high correlation (0.8) between the random-effects associated with variables yb and yc on the regression coefficients estimates. For simplicity, we will assume that the random effect associated with yd is uncorrelated with the random-effects of yb and yc and that all random-effects have unit variance (so covariances and correlations are identical). Thus, our fixed between-study covariance matrices for the three scenarios are

```
. matrix Sigma1 = (1,0,0\0,1,0\0,0,1)
. matrix Sigma2 = (1,0.4,0\0.4,1,0\0,0,1)
. matrix Sigma3 = (1,0.8,0\0.8,1,0\0,0,1)
```

We fit the first model using the correlations of 0 and store the estimation results as `corr0`.

```
. meta mvregress y*, wcovvariables(v*) random(mle, covariance(fixed(Sigma1)))
Multivariate random-effects meta-analysis        Number of obs     =         31
Method: User-specified Sigma = Sigma1            Number of studies =         24
                                                 Obs per study:
                                                              min =          1
                                                              avg =        1.3
                                                              max =          3
                                                 Wald chi2(0)      =          .
                                                 Prob > chi2       =          .

------------------------------------------------------------------------------
             | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
yb           |
       _cons |   .4293913    .502528     0.85   0.393    -.5555455    1.414328
-------------+----------------------------------------------------------------
yc           |
       _cons |   .7629462   .2739889     2.78   0.005     .2259379    1.299955
-------------+----------------------------------------------------------------
yd           |
       _cons |   1.028532   .5979445     1.72   0.085    -.1434175    2.200482
------------------------------------------------------------------------------
Test of homogeneity: Q_M = chi2(28) = 204.22               Prob > Q_M = 0.0000

------------------------------------------------
  Random-effects parameters |   Estimate
----------------------------+-------------------
User-specified Sigma1:      |
                      sd(yb)|          1
                      sd(yc)|          1
                      sd(yd)|          1
                 corr(yb,yc)|          0
                 corr(yb,yd)|          0
                 corr(yc,yd)|          0
------------------------------------------------
. estimates store corr0
```

Next, we fit the model with correlations of 0.4 and store results as `corr4` and the model with correlations of 0.8 and store results as `corr8`. For brevity, we suppress the output from both commands.

```
. quietly meta mvregress y*, wcovvariables(v*) random(mle, covariance(fixed(Sigma2)))
. estimates store corr4
. quietly meta mvregress y*, wcovvariables(v*) random(mle, covariance(fixed(Sigma3)))
. estimates store corr8
```

We compare the estimates side by side by using estimates table:

```
. estimates table corr0 corr4 corr8,
> keep(yb:_cons yc:_cons yd:_cons) b(%8.3f) se(%8.3f)
```

Variable	corr0	corr4	corr8
yb			
_cons	0.429	0.472	0.566
	0.503	0.478	0.418
yc			
_cons	0.763	0.752	0.730
	0.274	0.271	0.266
yd			
_cons	1.029	1.039	1.057
	0.598	0.603	0.607

Legend: b/se

As the correlation between the random effects associated with yb and yc increases, the coefficient estimate for yb increases, whereas that for yc decreases. Also, the two estimates become more precise (have smaller standard errors) as the correlation increases. This is expected because estimation borrows information from one outcome to estimate the coefficient of the other correlated outcome. This phenomenon is referred to as "strength borrowing" in the multivariate meta-analysis literature. Notice also how the various magnitudes of correlations had little to no impact on the estimation of yd because of the assumption of zero correlation between the random effect of yd and those of yb and of yc.

◁

▷ Example 11: Fixed-effects multivariate meta-regression

Gleser and Olkin (2009) reported six studies that compare the effects of five types of exercise with a control group (no exercise) on systolic blood pressure. This dataset was also analyzed by Hartung, Knapp, and Sinha (2008). Variables y1 to y5 are standard mean differences between each type of exercise and the control group. Ten variables, v11, v12, ..., v55, define the corresponding within-study variances and covariances.

The goal of this example is to demonstrate a potential problem that you may encounter in practice when there are missing observations in the data. And we also demonstrate how to perform a fixed-effects multivariate meta-analysis.

If we run the default random-effects model, we will get the following error message:

```
. use https://www.stata-press.com/data/r19/systolicbp
(Effect of exercise on systolic blood pressure)
. meta mvregress y*, wcovvariables(v*)
cannot estimate unstructured between-study covariance
    Variables y1 and y4 have 1 jointly observed value. With recov
    unstructured, at least 2 jointly observed values are required to estimate
    the between-study covariance. You may try specifying a different recov in
    option random(), such as random(, covariance(independent)).
r(459);
```

We list the observations on variables y1 and y4:

```
. list y1 y4, sep(0) noobs

  +---------------+
  |    y1      y4 |
  |---------------|
  |  .808       . |
  |     .   1.962 |
  |     .   2.568 |
  |     .       . |
  | 1.171   3.159 |
  |  .681       . |
  +---------------+
```

As the error message suggests, the estimation of the between-study covariance matrix, especially the element `cov(y1,y4)`, is not possible, because there is only one joint observation $(1.171, 3.159)$ on variables y1 and y4.

We may try a different random-effects covariance structure (see example 9 and example 10). Alternatively, we will follow Gleser and Olkin (2009) and perform a fixed-effects multivariate meta-analysis by specifying the `fixed` option.

```
. meta mvregress y*, wcovvariables(v*) fixed

Multivariate fixed-effects meta-analysis        Number of obs     =         15
                                                Number of studies =          6
                                                Obs per study:
                                                              min =          1
                                                              avg =        2.5
                                                              max =          4
                                                Wald chi2(0)      =          .
                                                Prob > chi2       =          .

------------------------------------------------------------------------------
             | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
y1           |
       _cons |   .7560005   .1144556     6.61   0.000     .5316716    .9803294
-------------+----------------------------------------------------------------
y2           |
       _cons |   1.398708   .1265397    11.05   0.000     1.150695    1.646722
-------------+----------------------------------------------------------------
y3           |
       _cons |   1.745014   .1646159    10.60   0.000     1.422373    2.067655
-------------+----------------------------------------------------------------
y4           |
       _cons |   2.146055   .1823172    11.77   0.000      1.78872     2.50339
-------------+----------------------------------------------------------------
y5           |
       _cons |   2.141486   .2338656     9.16   0.000     1.683118    2.599854
------------------------------------------------------------------------------
Test of homogeneity: Q_M = chi2(10) = 10.10                 Prob > Q_M = 0.4318
```

The homogeneity test based on the statistic $Q_M = 10.1$ favors the fixed-effects model ($p = 0.4318$). However, we should be careful not to rely solely on this test because it is known to have low power when the number of studies is small (Hedges and Pigott 2001).

◁

Stored results

meta mvregress stores the following in e():

Scalars

e(N)	total number of observations on *depvars*
e(k)	number of parameters
e(k_eq)	number of dependent variables
e(k_f)	number of fixed-effects parameters
e(k_r)	number of random-effects parameters
e(k_rs)	number of variances
e(k_rc)	number of covariances
e(seadj)	standard error adjustment (se() only)
e(ll)	log (restricted) likelihood (mle and reml only)
e(rank)	rank of e(V)
e(ic)	number of iterations (mle and reml only)
e(df_m)	model degrees of freedom
e(chi2)	model χ^2 Wald test statistic
e(df_r)	model denominator degrees of freedom (tdistribution() only)
e(F)	model F statistic (tdistribution() only)
e(p)	p-value for model test
e(Q_M)	multivariate Cochran Q residual homogeneity test statistic
e(df_Q_M)	degrees of freedom for residual homogeneity test
e(p_Q_M)	p-value for residual homogeneity test
e(converged)	1 if converged, 0 otherwise (mle and reml only)
e(s_max)	maximum number of observations per study
e(s_avg)	average number of observations per study
e(s_min)	minimum number of observations per study
e(N_s)	number of studies

Macros

e(cmd)	meta mvregress
e(cmdline)	command as typed
e(model)	multivariate meta-analysis model
e(method)	multivariate meta-analysis estimation method
e(title)	title in estimation output
e(chi2type)	Wald; type of model χ^2 test
e(depvars)	names of dependent variables
e(indepvars)	names of independent variables (moderators)
e(wcovvariables)	variables defining within-study covariance matrix
e(wsevariables)	standard error variables from wsevariables()
e(wcorrelations)	values of the assumed within-study correlations from wcorrelations()
e(redim)	random-effects dimensions
e(vartypes)	variance-structure types
e(seadjtype)	type of standard error adjustment (se() only)
e(technique)	maximization technique (mle and reml only)
e(ml_method)	type of ml method
e(opt)	type of optimization (mle and reml only)
e(optmetric)	matsqrt or matlog; random-effects matrix parameterization (mle and reml only)
e(properties)	b V
e(predict)	program used to implement predict
e(estat_cmd)	program used to implement estat
e(marginsok)	predictions allowed by margins
e(marginsnotok)	predictions disallowed by margins
e(marginsdefault)	default predict() specification for margins
e(asbalanced)	factor variables fvset as asbalanced
e(asobserved)	factor variables fvset as asobserved

Matrices

`e(b)`	coefficient vector
`e(V)`	variance–covariance matrix of the estimators

Functions

`e(sample)`	marks estimation sample

In addition to the above, the following is stored in `r()`:

Matrices

`r(table)`	matrix containing the coefficients with their standard errors, test statistics, p-values, and confidence intervals

Note that results stored in `r()` are updated when the command is replayed and will be replaced when any r-class command is run after the estimation command.

Methods and formulas

Methods and formulas are presented under the following headings:

For an overview of estimation methods used by multivariate meta-regression, see van Houwelingen, Arends, and Stijnen (2002), Jackson, Riley, and White (2011), White (2011), and Sera et al. (2019).

Consider data from K independent studies and d outcomes (effect sizes). Let $\widehat{\theta}_{ij}$ be the estimated effect size reported by study j for outcome i, and let the $d \times 1$ vector $\widehat{\boldsymbol{\theta}}_j = (\widehat{\theta}_{1j}, \widehat{\theta}_{2j}, \ldots, \widehat{\theta}_{dj})'$ be an estimate of the true population effect size $\boldsymbol{\theta}_j$ for study j.

Fixed-effects multivariate meta-regression

A model for the fixed-effects multivariate meta-regression (Raudenbush, Becker, and Kalaian 1988) can be expressed as

$$\widehat{\theta}_{ij} = \beta_{i0} + \beta_{i1}x_{1j} + \cdots + \beta_{i,p-1}x_{p-1,j} + \epsilon_{ij} = \mathbf{x}_j\boldsymbol{\beta}_i + \epsilon_{ij}$$

for outcome $i = 1, \ldots, d$ and study $j = 1, \ldots, K$. Here $\mathbf{x}_j = (1, x_{1j}, \ldots, x_{p-1,j})$ is a $1 \times p$ vector of categorical and continuous moderators (covariates), $\boldsymbol{\beta}_i$ is an outcome-specific $p \times 1$ vector of unknown regression coefficients, and $\boldsymbol{\epsilon}_j = (\epsilon_{1j}, \epsilon_{2j}, \ldots, \epsilon_{dj})'$ is a $d \times 1$ vector of within-study errors that have a d-variate normal distribution with zero mean vector and a $d \times d$ covariance matrix $\text{Var}(\boldsymbol{\epsilon}_j) = \boldsymbol{\Lambda}_j$. The within-study covariance matrices $\boldsymbol{\Lambda}_j$'s are treated as known and do not require estimation. $\boldsymbol{\Lambda}_j$'s reduce to $\widehat{\sigma}_j^2$ in the case of univariate meta-analysis; see *Methods and formulas* of [META] **meta summarize**.

In matrix notation, the above fixed-effects model can be defined as

$$\widehat{\boldsymbol{\theta}}_j = \mathbf{X}_j\boldsymbol{\beta} + \boldsymbol{\epsilon}_j, \quad \boldsymbol{\epsilon}_j \sim N_d\left(\mathbf{0},\ \boldsymbol{\Lambda}_j\right)$$

where $\mathbf{X}_j = \mathbf{x}_j \otimes I_d$ ($\otimes$ is the Kronecker product) is a $d \times dp$ matrix and $\boldsymbol{\beta} = (\boldsymbol{\beta}_1', \boldsymbol{\beta}_2', \ldots, \boldsymbol{\beta}_d')'$ is a $dp \times 1$ vector of all unknown regression coefficients.

Let $\mathbf{W}_j = \boldsymbol{\Lambda}_j^{-1}$, a $d \times d$ matrix. Then the fixed-effects estimator for the regression coefficients is

$$\widehat{\boldsymbol{\beta}} = \left(\sum_{j=1}^{K} \mathbf{X}_j' \mathbf{W}_j \mathbf{X}_j \right)^{-1} \sum_{j=1}^{K} \mathbf{X}_j' \mathbf{W}_j \hat{\boldsymbol{\theta}}_j$$

and the corresponding covariance matrix is

$$\mathrm{Var}(\widehat{\boldsymbol{\beta}}) = \left(\sum_{j=1}^{K} \mathbf{X}_j' \mathbf{W}_j \mathbf{X}_j \right)^{-1} \tag{1}$$

The above fixed-effects regression does not account for residual heterogeneity. This can lead to standard errors of regression coefficients that are too small. Next we present a random-effects multivariate meta-regression model that incorporates residual heterogeneity by including an additive between-study covariance component $\boldsymbol{\Sigma}$.

Random-effects multivariate meta-regression

Consider the following extension of a fixed-effects multivariate meta-regression model (Berkey et al. 1998):

$$\hat{\boldsymbol{\theta}}_j = \mathbf{X}_j \boldsymbol{\beta} + \boldsymbol{\epsilon}_j^*, \quad \text{where } \boldsymbol{\epsilon}_j^* \sim N_d\left(\mathbf{0}, \boldsymbol{\Lambda}_j + \boldsymbol{\Sigma}\right)$$

Alternatively, the above model can be written as

$$\hat{\boldsymbol{\theta}}_j = \mathbf{X}_j \boldsymbol{\beta} + \mathbf{u}_j + \boldsymbol{\epsilon}_j, \quad \boldsymbol{\epsilon}_j \sim N_d\left(\mathbf{0}, \boldsymbol{\Lambda}_j\right)$$

where random effects $\mathbf{u}_j = (u_{1j}, u_{2j}, \dots, u_{dj})' \sim N_d(\mathbf{0}, \boldsymbol{\Sigma})$ $(j = 1, \dots, K)$ account for the additional variation that is not explained by moderators $\mathbf{X}_j$.

The models above define a random-effects multivariate meta-regression.

Let $\widehat{\boldsymbol{\Sigma}}$ be an estimate of the between-study covariance matrix $\boldsymbol{\Sigma}$ (to be discussed later), and let $\mathbf{W}_j^* = (\widehat{\boldsymbol{\Sigma}} + \boldsymbol{\Lambda}_j)^{-1}$. The random-effects estimator for the regression coefficients is

$$\widehat{\boldsymbol{\beta}}^* = \left(\sum_{j=1}^{K} \mathbf{X}_j' \mathbf{W}_j^* \mathbf{X}_j \right)^{-1} \sum_{j=1}^{K} \mathbf{X}_j' \mathbf{W}_j^* \hat{\boldsymbol{\theta}}_j$$

The corresponding covariance matrix is given by

$$\mathrm{Var}(\widehat{\boldsymbol{\beta}}^*) = \left(\sum_{j=1}^{K} \mathbf{X}_j' \mathbf{W}_j^* \mathbf{X}_j \right)^{-1} \tag{2}$$

In the following section, we outline the estimation of the between-study covariance matrix $\boldsymbol{\Sigma}$ for the ML and REML iterative methods. For the noniterative Jackson–White–Riley of estimating $\boldsymbol{\Sigma}$, see *Noniterative method for computing* $\boldsymbol{\Sigma}$.

Iterative methods for computing Σ

The two estimators described below do not have a closed-form solution, and an iterative algorithm is needed to estimate $\boldsymbol{\Sigma}$.

The joint log-likelihood function of $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$ for a random-effects multivariate meta-regression can be expressed as

$$\ln L_{\mathrm{ML}}\left(\boldsymbol{\beta},\, \boldsymbol{\Sigma}\right) = -\frac{1}{2}\left\{ n\ln(2\pi) + \sum_{j=1}^{K} \ln\left|\mathbf{V}_j\right| + \sum_{j=1}^{K}\left(\hat{\boldsymbol{\theta}}_j - \mathbf{X}_j\boldsymbol{\beta}\right)' \mathbf{V}_j^{-1}\left(\hat{\boldsymbol{\theta}}_j - \mathbf{X}_j\boldsymbol{\beta}\right)\right\}$$

where $\mathbf{V}_j = \boldsymbol{\Sigma} + \boldsymbol{\Lambda}_j$, $|\mathbf{V}_j|$ is the determinant of $\mathbf{V}_j$, and n is the total number of observations $\widehat{\theta}_{ij}$ ($n = Kd$ when there are no missing data).

The between-study covariance $\boldsymbol{\Sigma}$ is estimated by maximizing the profile log-likelihood function obtained by treating $\boldsymbol{\beta}$ as known and plugging $\widehat{\boldsymbol{\beta}}^*$ into $\ln L_{\mathrm{ML}}(\boldsymbol{\beta},\, \boldsymbol{\Sigma})$ in place of $\boldsymbol{\beta}$ (Pinheiro and Bates [2000, ch. 2]):

$$\ln L_{\mathrm{ML}}\left(\boldsymbol{\Sigma}\right) = -\frac{1}{2}\left\{ n\ln(2\pi) + \sum_{j=1}^{K} \ln\left|\mathbf{V}_j\right| + \sum_{j=1}^{K}\left(\hat{\boldsymbol{\theta}}_j - \mathbf{X}_j\widehat{\boldsymbol{\beta}}^*\right)' \mathbf{V}_j^{-1}\left(\hat{\boldsymbol{\theta}}_j - \mathbf{X}_j\widehat{\boldsymbol{\beta}}^*\right)\right\}$$

The MLE of $\boldsymbol{\Sigma}$ does not incorporate the uncertainty about the unknown regression coefficients $\boldsymbol{\beta}$ and thus can be negatively biased.

The REML estimator of $\boldsymbol{\Sigma}$ maximizes the restricted log-likelihood function

$$\ln L_{\mathrm{REML}}\left(\boldsymbol{\Sigma}\right) = \ln L_{\mathrm{ML}}\left(\boldsymbol{\Sigma}\right) - \frac{1}{2}\ln\left|\sum_{j=1}^{K}\mathbf{X}_j'\mathbf{V}_j^{-1}\mathbf{X}_j\right| + \frac{dp}{2}\ln(2\pi)$$

The REML method estimates $\boldsymbol{\Sigma}$ by accounting for the uncertainty in the estimation of $\boldsymbol{\beta}$, which leads to a nearly unbiased estimate of $\boldsymbol{\Sigma}$. The optimization of the above log-likelihood functions can be done using the machinery of the mixed-effects models to obtain the estimates $\widehat{\boldsymbol{\beta}}^*$ and $\widehat{\boldsymbol{\Sigma}}$. For details, see Pinheiro and Bates (2000) and *Methods and formulas* of [ME] **mixed**. When $d = 1$, that is, in the context of univariate meta-analysis, the above ML and REML estimators reduce to their univariate counterparts as reported by `meta regress`.

Noniterative method for computing Σ

This section describes a noniterative method to estimate the between-study covariance matrix $\boldsymbol{\Sigma}$, which has a closed-form expression. The formulas in this section are based on Jackson, White, and Riley (2013).

Using the notation for a fixed-effects multivariate meta-regression, define a $d \times d$ matrix

$$\mathbf{Q}_{\mathrm{JWR}} = \sum_{j=1}^{K}\mathbf{W}_j\left(\hat{\boldsymbol{\theta}}_j - \mathbf{X}_j\widehat{\boldsymbol{\beta}}\right)\left(\hat{\boldsymbol{\theta}}_j - \mathbf{X}_j\widehat{\boldsymbol{\beta}}\right)'\mathbf{R}_j$$

where $\mathbf{R}_j$ is a $d \times d$ diagonal matrix with the ith diagonal element equal to 1 if $\widehat{\theta}_{ij}$ is observed and 0 if it is missing.

The role of $\mathbf{R}_j$ is to ensure that missing outcomes do not contribute to the computation of $\mathbf{Q}_{\mathrm{JWR}}$. Let $\mathbf{R} = \oplus_{j=1}^{K}\mathbf{R}_j$ and $\mathbf{W} = \oplus_{j=1}^{K}\mathbf{W}_j$ be $Kd \times Kd$ block-diagonal matrices formed by submatrices $\mathbf{R}_j$ and $\mathbf{W}_j$, respectively; $\oplus$ is the Kronecker sum. In the presence of missing outcome values, the matrix $\mathbf{W}_j = \boldsymbol{\Lambda}_j^{-1}$ is obtained by inverting the submatrix of $\boldsymbol{\Lambda}_j$ corresponding to the observed outcome values and by replacing the remaining elements with zeros.

Let $\mathbf{X}$ denote a $Kd \times p$ matrix constructed by vertically stacking the $d \times p$ matrices $\mathbf{X}_j$, that is, $\mathbf{X} = (\mathbf{X}_1', \mathbf{X}_2', \ldots, \mathbf{X}_K')'$. Define

$$\begin{aligned} \mathbf{P}_{\text{M}} &= (\mathbf{I}_{Kd} - \mathbf{H})' \, \mathbf{W} \\ \mathbf{B} &= (\mathbf{I}_{Kd} - \mathbf{H})' \, \mathbf{R} \end{aligned} \tag{3}$$

where $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}$ and $\mathbf{I}_{Kd}$ is the $Kd \times Kd$ identity matrix. The subscript M in $\mathbf{P}_{\text{M}}$ is used to emphasize that the $Kd \times Kd$ matrix $\mathbf{P}_{\text{M}}$ generalizes the $K \times K$ matrix $\mathbf{P}$, defined by (1) in *Methods and formulas* of [META] **meta regress**, to the multivariate meta-regression setting.

Partition the $Kd \times Kd$ matrices $\mathbf{P}_{\text{M}}$ and $\mathbf{B}$ into K^2 blocks of $d \times d$ matrices, and denote the jth by lth submatrix of $\mathbf{P}_{\text{M}}$ by $(\mathbf{P}_{\text{M}})_{jl}$ and of $\mathbf{B}$ by $(\mathbf{B})_{jl}$, respectively. The method of moments estimator proposed by Jackson, White, and Riley (2013) solves the system of d^2 estimating equations

$$\text{vec}\,(\mathbf{Q}_{\text{JWR}}) = \text{vec}\left\{\sum_{j=1}^{K} (\mathbf{B})_{jj}\right\} + \left\{\sum_{l=1}^{K}\sum_{j=1}^{K} (\mathbf{B})_{jl}' \otimes (\mathbf{P}_{\text{M}})_{lj}\right\} \text{vec}(\widetilde{\boldsymbol{\Sigma}})$$

where vec($\mathbf{A}$) vectorizes $\mathbf{A}$ column by column and $\otimes$ is the Kronecker product. Solving for vec($\widetilde{\boldsymbol{\Sigma}}$) and hence $\widetilde{\boldsymbol{\Sigma}}$, we obtain the JWR estimator of the between-study covariance matrix,

$$\widehat{\boldsymbol{\Sigma}}_{\text{JWR}} = \frac{\widetilde{\boldsymbol{\Sigma}} + \widetilde{\boldsymbol{\Sigma}}'}{2}$$

The estimator $\widehat{\boldsymbol{\Sigma}}_{\text{JWR}}$ is symmetric but not necessarily positive semidefinite. We can obtain a positive semidefinite estimator, $\widehat{\boldsymbol{\Sigma}}^{+}_{\text{JWR}}$, based on spectral decomposition $\widehat{\boldsymbol{\Sigma}}_{\text{JWR}} = \sum_{i=1}^{d} \lambda_i \mathbf{e}_i \mathbf{e}_i'$ as follows,

$$\widehat{\boldsymbol{\Sigma}}^{+}_{\text{JWR}} = \sum_{i=1}^{d} \max\,(0, \lambda_i)\, \mathbf{e}_i \mathbf{e}_i'$$

where λ_is are the eigenvalues of $\widehat{\boldsymbol{\Sigma}}_{\text{JWR}}$ and $\mathbf{e}_i$s are the corresponding orthonormal eigenvectors. $\widehat{\boldsymbol{\Sigma}}^{+}_{\text{JWR}}$ has the same eigenvectors as $\widehat{\boldsymbol{\Sigma}}_{\text{JWR}}$ but with negative eigenvalues truncated at 0.

The JWR estimator can be viewed as an extension of the DerSimonian–Laird estimator from the random-effects meta-regression to multivariate meta-regression. For univariate meta-analysis ($d = 1$), the JWR estimator reduces to the DerSimonian–Laird estimator from `meta regress`. The truncation of $\widehat{\boldsymbol{\Sigma}}_{\text{JWR}}$ to obtain $\widehat{\boldsymbol{\Sigma}}^{+}_{\text{JWR}}$ is equivalent to truncating $\widehat{\tau}^2_{\text{DL}}$ at 0 in univariate meta-regression whenever the estimate is negative.

Random-effects covariance structures

Several covariance structures may be assumed for the between-study covariance matrix $\boldsymbol{\Sigma}$. The default covariance structure is `unstructured`, which is the most general structure in which all elements or, more precisely, $d(d + 1)/2$ variance components are estimated. Other covariance structures are `independent`, `exchangeable`, `identity`, and `fixed(`*matname*`)`. These structures may be useful to provide more stable estimates by reducing the complexity of the model, especially when the number of observations, n, is relatively small.

For example, when $d = 3$, the covariance structures are

$$\texttt{unstructured} \quad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & & \\ \sigma_{21} & \sigma_{22} & \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix}$$

$$\texttt{independent} \quad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & & \\ 0 & \sigma_{22} & \\ 0 & 0 & \sigma_{33} \end{bmatrix}$$

$$\texttt{exchangeable} \quad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & & \\ \sigma_{21} & \sigma_{11} & \\ \sigma_{21} & \sigma_{21} & \sigma_{11} \end{bmatrix}$$

$$\texttt{identity} \quad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & & \\ 0 & \sigma_{11} & \\ 0 & 0 & \sigma_{11} \end{bmatrix}$$

Any of the above covariance structures may be specified with the ML and REML methods. Only the `unstructured` covariance structure is allowed with the JWR method. When covariance structure `fixed(`*matname*`)` is specified, *matname* is assumed to be the known between-study covariance, and thus no iteration is needed.

Jackson–Riley standard-error adjustment

By default, the inference about the regression coefficients and their confidence intervals from meta-regression is based on a normal distribution. The test of the significance of all regression coefficients is based on a χ^2 distribution with $d(p-1)$ degrees of freedom.

Jackson and Riley (2014) proposed an adjustment to the standard errors of the estimated regression coefficients to account for the uncertainty in the estimation of $\boldsymbol{\Sigma}$. They showed that the corresponding tests of individual regression coefficients and their confidence intervals are based on the Student's t distribution with $n - dp$ degrees of freedom and that the overall test of significance is based on an F distribution with $d(p-1)$ numerator and $n - dp$ denominator degrees of freedom.

The Jackson–Riley adjustment first calculates the quadratic form,

$$q_{\rm JR} = \frac{1}{n - dp} \sum_{j=1}^{K} \left(\hat{\boldsymbol{\theta}}_j - \mathbf{X}_j \widehat{\boldsymbol{\beta}}\right)' \mathbf{W}_j^* \left(\hat{\boldsymbol{\theta}}_j - \mathbf{X}_j \widehat{\boldsymbol{\beta}}\right)$$

It then multiplies the regular expressions of the variances of regression coefficients by $q_{\rm JR}$ or, in the case of the truncated Jackson–Riley adjustment, by $\max(1, q_{\rm JR})$. When $d = 1$, the Jackson–Riley adjustment, $q_{\rm JR}$, reduces to the Knapp–Hartung adjustment, $q_{\rm KH}$, from *Knapp–Hartung standard-error adjustment* in *Methods and formulas* in [META] **meta regress**.

Multivariate meta-analysis

The formulas presented so far are derived for the general case of multivariate meta-regression. Methods and formulas for the special case of multivariate meta-analysis (when no moderators are included) can be obtained by taking $\mathbf{x}_j = 1$ and $p = 1$. When $d = 1$, the REML, ML, and JWR estimators reduce to the univariate REML, ML, and DL estimators described in [META] **meta summarize** for constant-only models and in [META] **meta regress** for regression models.

Residual homogeneity test

Consider a test of residual homogeneity, which mathematically translates to H_0: $\boldsymbol{\Sigma} = \mathbf{0}_{d\times d}$ for the random-effects multivariate meta-regression. This test is based on the multivariate residual weighted sum of squares, $Q_{\rm M}$,

$$Q_{\rm M} = \sum_{j=1}^{K} \left(\hat{\boldsymbol{\theta}}_j - \mathbf{X}_j\widehat{\boldsymbol{\beta}}\right)' \mathbf{W}_j \left(\hat{\boldsymbol{\theta}}_j - \mathbf{X}_j\widehat{\boldsymbol{\beta}}\right)$$

where $\widehat{\boldsymbol{\beta}}$ is a fixed-effects estimator of regression coefficients defined for a fixed-effects multivariate meta-regression.

Under the null hypothesis of residual homogeneity, $Q_{\rm M}$ follows a χ^2 distribution with $n - dp$ degrees of freedom (Seber and Lee 2003, sec. 2.4). The $Q_{\rm M}$ statistic reduces to the univariate residual homogeneity test statistic, $Q_{\rm res}$, when $d = 1$ (see *Residual homogeneity test* in *Methods and formulas* in [META] **meta regress**). It also reduces to the univariate homogeneity statistic Q when no moderators are included (see *Homogeneity test* in *Methods and formulas* in [META] **meta summarize**).

References

Antczak-Bouckoms, A., K. Joshipura, E. Burdick, and J. F. Camilla Tulloch. 1993. Meta-analysis of surgical versus non-surgical methods of treatment for periodontal disease. *Journal of Clinical Periodontology* 20: 259–268. https://doi.org/10.1111/j.1600-051X.1993.tb00355.x.

Berkey, C. S., D. C. Hoaglin, F. Mosteller, and G. A. Colditz. 1998. Meta-analysis of multiple outcomes by regression with random effects. *Statistics in Medicine* 17: 2537–2550. https://doi.org/10.1002/(SICI)1097-0258(19981130)17:22%3C2537::AID-SIM953%3E3.0.CO;2-C.

Fiore, M. C., W. C. Bailey, S. J. Cohen, S. F. Dorfman, M. G. Goldstein, E. R. Gritz, R. B. Heyman, J. Holbrook, C. R. Jaen, T. E. Kottke, H. A. Lando, R. Mecklenburg, P. D. Mullen, L. M. Nett, L. Robinson, M. L. Stitzer, A. C. Tommasello, L. Villejo, and M. E. Wewers. 1996. "Smoking Cessation". In *Clinical Practice Guideline No. 18*. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research. AHCPR Publication No. 96-0692.

Glas, A. S., D. Roos, M. Deutekom, A. H. Zwinderman, P. M. M. Bossuyt, and K. H. Kurth. 2003. Tumor markers in the diagnosis of primary bladder cancer. A systematic review. *Journal of Urology* 169: 1975–1982. https://doi.org/10.1097/01.ju.0000067461.30468.6d.

Gleser, L. J., and I. Olkin. 2009. "Stochastically dependent effect sizes". In *The Handbook of Research Synthesis and Meta-Analysis*, edited by H. Cooper, L. V. Hedges, and J. C. Valentine, 357–376. 2nd ed. New York: Russell Sage Foundation.

Harbord, R. M., and J. P. T. Higgins. 2016. "Meta-regression in Stata". In *Meta-Analysis in Stata: An Updated Collection from the Stata Journal*, edited by T. M. Palmer and J. A. C. Sterne, 85–111. 2nd ed. College Station, TX: Stata Press.

Hartung, J., G. Knapp, and B. K. Sinha. 2008. *Statistical Meta-Analysis with Applications*. Hoboken, NJ: Wiley. https://doi.org/10.1002/9780470386347.

Hedges, L. V., and T. D. Pigott. 2001. The power of statistical tests in meta-analysis. *Psychological Methods* 6: 203–217. https://doi.org/10.1037/1082-989X.6.3.203.

Higgins, J. P. T., and S. G. Thompson. 2004. Controlling the risk of spurious findings from meta-regression. *Statistics in Medicine* 23: 1663–1682. https://doi.org/10.1002/sim.1752.

Higgins, J. P. T., S. G. Thompson, J. J. Deeks, and D. G. Altman. 2003. Measuring inconsistency in meta-analyses. *BMJ* 327: 557–560. https://doi.org/10.1136/bmj.327.7414.557.

Jackson, D., and R. D. Riley. 2014. A refined method for multivariate meta-analysis and meta-regression. *Statistics in Medicine* 33: 541–554. https://doi.org/10.1002/sim.5957.

Jackson, D., R. D. Riley, and I. R. White. 2011. Multivariate meta-analysis: Potential and promise. *Statistics in Medicine* 30: 2481–2498. https://doi.org/10.1002/sim.4172.

Jackson, D., I. R. White, and R. D. Riley. 2012. Quantifying the impact of between-study heterogeneity in multivariate meta-analyses. *Statistics in Medicine* 31: 3805–3820. https://doi.org/10.1002/sim.5453.

———. 2013. A matrix-based method of moments for fitting the multivariate random effects model for meta-analysis and meta-regression. *Biometrical Journal* 55: 231–245. https://doi.org/10.1002/bimj.201200152.

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.

Lu, G., and A. E. Ades. 2006. Assessing evidence inconsistency in mixed treatment comparisons. *Journal of the American Statistical Association* 101: 447–459. https://doi.org/10.1198/016214505000001302.

Orsini, N. 2021. Weighted mixed-effects dose–response models for tables of correlated contrasts. *Stata Journal* 21: 320–347.

Pinheiro, J. C., and D. M. Bates. 2000. *Mixed-Effects Models in S and S-PLUS*. New York: Springer. https://doi.org/10.1007/b98882.

Raudenbush, S. W., B. J. Becker, and H. Kalaian. 1988. Modeling multivariate effect sizes. *Psychological Bulletin* 103: 111–120. https://doi.org/10.1037/0033-2909.103.1.111.

Riley, R. D., K. R. Abrams, A. J. Sutton, P. C. Lambert, and J. R. Thompson. 2007. Bivariate random-effects meta-analysis and the estimation of between-study correlation. *BMC Medical Research Methodology* 7: art. 3. https://doi.org/10.1186/1471-2288-7-3.

Rose, C. J. 2024. Multivariate random-effects meta-analysis for sparse data using smvmeta. *Stata Journal* 24: 301–328.

Rosenthal, R., and D. B. Rubin. 1986. Meta-analytic procedures for combining studies with multiple effect sizes. *Psychological Bulletin* 99: 400–406. https://doi.org/10.1037/0033-2909.99.3.400.

Seber, G. A. F., and A. J. Lee. 2003. *Linear Regression Analysis*. 2nd ed. Hoboken, NJ: Wiley. https://doi.org/10.1002/9780471722199.

Sera, F., B. Armstrong, M. Blangiardo, and A. Gasparrini. 2019. An extended mixed-effects framework for meta-analysis. *Statistics in Medicine* 38: 5429–5444. https://doi.org/10.1002/sim.8362.

Thompson, S. G., and S. J. Sharp. 1999. Explaining heterogeneity in meta-analysis: A comparison of methods. *Statistics in Medicine* 18: 2693–2708. https://doi.org/10.1002/(sici)1097-0258(19991030)18:20<2693::aid-sim235>3.0.co;2-v.

van Houwelingen, H. C., L. R. Arends, and T. Stijnen. 2002. Advanced methods in meta-analysis: Multivariate approach and meta-regression. *Statistics in Medicine* 21: 589–624. https://doi.org/10.1002/sim.1040.

White, I. R. 2011. Multivariate random-effects meta-regression: Updates to mvmeta. *Stata Journal* 11: 240–254.

———. 2016. "Multivariate random-effects meta-analysis". In *Meta-Analysis in Stata: An Updated Collection from the Stata Journal*, edited by T. M. Palmer and J. A. C. Sterne, 232–248. 2nd ed. College Station, TX: Stata Press.

Also see

[META] **meta mvregress postestimation** — Postestimation tools for meta mvregress

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

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

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

[META] **Glossary**

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

[U] **20 Estimation and postestimation commands**