A Note on Cohen’s d From a Partitioned Linear Regression Model

Jürgen Groß · Annette Möller

Accepted: 17 January 2023 / Published online: 8 February 2023
© The Author(s) 2023

Abstract
In this note, we introduce a generalized formula for Cohen’s $d$ under the presence of additional independent variables, providing a measure for the size of a possible effect concerning the size of a difference location effect of a variable in two groups. This is done by employing the so-called Frisch–Waugh–Lovell theorem in a partitioned linear regression model. The generalization is motivated by demonstrating the relationship to appropriate $t$ and $F$ statistics. Our discussion is further illustrated by inference about a publicly available data set.

Keywords Hypothesis testing · Effect size · Cohen’s $d$ · Partitioned linear regression · Frisch–Waugh–Lovell theorem · Multivariate normal distribution

Mathematics Subject Classification 62J20 · 62F03 · 91C99

1 Introduction

When applying statistical testing of hypotheses to data, it is often recommended not only to report the corresponding p value, but in addition to provide a measure for the effect associated with a possible rejection of the null hypothesis, see e.g. Wilkinson [18]. Such a measure may be useful when sample sizes are to be fixed during the planning phase of a study, or when it is desired to assess the relevance of an actual...
rejection when given sample sizes are large. Effect size measures are strongly related to power analysis as carried out in the seminal book by Cohen [3].

A widely used measure is the so-called Cohen’s $d$, see also Hedges [12] and Kraemer [13], which is an effect size measure for the two-sample $t$ test with equal variances. Consider independent samples of sizes $n_1$ and $n_2$ of a statistical variable $y$ in two groups such that $y$ follows a normal distribution with expectation $\mu_1$ and variance $\sigma^2$ in group 1 and expectation $\mu_2$ and the same variance $\sigma^2$ in group 2. Let $t$ denote the usual two-sample test statistic for the null hypotheses $H_0 : \mu_1 = \mu_2$ versus the alternative $H_1 : \mu_1 \neq \mu_2$. As a measure for the size of an effect, Cohen [3, p. 66ff] considers the absolute value of

$$d = \left| \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{s_1^2 + s_2^2}{n_1 + n_2 - 2}}} \right| ,$$

where $\bar{y}_j$ is the sample mean in group $j$, $j = 1, 2$, and $s_j^2 = \sum_i (y_i - \bar{y}_j)^2$, where summation is carried out with respect to all observations from group $j$. The effect size $d$ is related to the test statistic $t$ by the formula

$$d = t \sqrt{\frac{n_1 + n_2}{n_1 n_2}} ,$$

see (2.5.3) in Cohen [3]. According to Cohen, values $|d| = 0.2$, $|d| = 0.5$ and $|d| = 0.8$ indicate a small, medium and large effect, respectively.

It may also be of interest to have a corresponding measure when the variable $y$ depends on further independent variables. In his Chapter 9, Cohen [3] deals with such a multiple regression situation and discusses the effect size measure $f^2$ at length, as will further be explicated in our Section 4.

However, an analogous measure to $d$ is rare to find; see Wilson [19, Sect. 3.14], Lipsey and Wilson [14] for such a proposal. Nonetheless, it may be of particular interest to have comparable measures of an effect size for the very same grouped variable $y$ but additionally depending on different sets of independent variables. This is exemplarily carried out in our Section 5. In the following, we introduce such a measure as a generalization to $d$ by considering a linear regression model

$$y = \beta_0 + \beta_1 z + \beta_2 x_1 + \cdots + \beta_{w+1} x_w + \varepsilon ,$$

where $z$ takes the value $z_i = 0$ if the corresponding observation $y_i$ of the dependent variable $y$ belongs to group 1 and $z_i = 1$ if $y_i$ belongs to group 2, $i = 1, \ldots, n_1 + n_2$. It is assumed that there are $w$ independent variables $x_1, \ldots, x_w$. The error variable $\varepsilon$ is assumed to follow a normal distribution with expectation 0 and variance $\sigma^2$.

As will be shown in the following Sections 2, 3, and 4, a natural generalization of Cohen’s $d$ is given by
\[
d_* = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{s^2_1 + s^2_2 / n_1 + n_2 - 2}}, \quad s^2_j = \sum_i (y_{*i} - \bar{y}_j)^2, \quad j = 1, 2, \tag{4}
\]

where \( y_* = y - \hat{\beta}_2 x_1 - \cdots - \hat{\beta}_{w+1} x_w \) is the dependent variable adjusted for the independent variables. The \( \hat{\beta}_k \) are the ordinary least squares estimates of the regression coefficients \( \beta_k, k = 2, \ldots, w+1 \) in model (3). In case \( w = 0 \), the adjusted \( y_* \) coincides with the original \( y \), so that (4) reduces to (1) and therefore can be seen as a natural generalization of Cohen’s \( d \).

## 2 Partitioned Linear Regression

Let \( n = n_1 + n_2 \) be the total sample size. The above model (3) may also be written in vector-matrix notation as

\[
y = X_1 \delta_1 + X_2 \delta_2 + \varepsilon, \tag{5}
\]

where now \( y \) represents the \( n \times 1 \) vector of observations of the dependent variable. Without loss of generality, it is assumed that the first \( n_1 \) observations belong to group 1, while the last \( n_2 \) observations belong to group 2. By introducing the notation \( 1_m \) for an \( m \times 1 \) vectors of ones, the \( n \times 2 \) matrix \( X_1 \) and the corresponding \( 2 \times 1 \) parameter vector \( \delta_1 \) may be written as

\[
X_1 = \begin{pmatrix} 1_{n_1} & 0 \\ 1_{n_2} & 1_{n_2} \end{pmatrix} \quad \text{and} \quad \delta_1 = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}. \tag{6}
\]

The \( n \times w \) matrix \( X_2 \) contains the observations of the independent variables with corresponding regression coefficients \( \delta^T_2 = (\beta_2, \ldots, \beta_{w+1}) \), where the \( T \) superscript denotes transposition. The \( n \times 1 \) random vector \( \varepsilon \) is assumed to follow a multivariate normal distribution with expectation vector 0 and variance-covariance matrix \( \sigma^2 I_n \), where \( I_n \) stands for the \( n \times n \) identity matrix. It is assumed that the \( n \times (2 + w) \) model matrix \( (X_1, X_2) \) has full column rank \( 2 + w \). Equation (5) represents a partitioned linear regression model as considered e.g. in Fiebig et al [7]. Generalizations and further properties are investigated by Puntanen [16], Groß and Puntanen [10, 11], Ding [5], among others.

Under model (5), the ordinary least squares estimator for the parameter vector \( (\delta^T_1, \delta^T_2) \) is given by

\[
\begin{pmatrix} \hat{\delta}_1 \\ \hat{\delta}_2 \end{pmatrix} = (X^T X)^{-1} X^T y, \quad X = (X_1, X_2). \tag{7}
\]

The Frisch–Waugh–Lovell theorem, see Fiebig et al [7], Lovell [15], Frisch and Waugh [8], states that

\[
\hat{\delta}_2 = (X^T M_1 X_2)^{-1} X^T M_1 y, \quad M_1 = I_n - X_1 (X^T_1 X_1)^{-1} X^T_1. \tag{8}
\]
For the specific choice (6), the matrix $M_1$ becomes

$$M_1 = \begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix}, \quad C_j = I_n - n_j^{-1} 1_n 1_j^T, \quad j = 1, 2. \quad (9)$$

The following result is not restricted to the case (6) but remains valid in situations where the matrix $X_1$ corresponds to an arbitrary set of $v$ independent variables such that the assumptions of (5) are satisfied.

**Theorem 1** Under the partitioned linear regression model (5),

$$\hat{\delta}_1 = (X_1^T X_1)^{-1} X_1^T (y - X_2 \hat{\delta}_2) \quad (10)$$

is the ordinary least squares estimator of $\delta_1$.

A proof is given in the appendix. Theorem 1 means that if $\hat{\delta}_2$ is known (e.g. computed by (8)), then the remaining parameters $\delta_1$ can be estimated by regressing the adjusted

$$y_\ast = y - X_2 \hat{\delta}_2 \quad (11)$$
on the remaining $X_1$ and this procedure just yields the identical estimate of $\delta_1$ from (7).

**Theorem 2** Under the partitioned linear regression model (5) and (6),

$$\hat{\sigma}^2 = (y - X_2 \hat{\delta}_2)^T M_1 (y - X_2 \hat{\delta}_2)/(n - 2 - w) \quad (12)$$

$$= (s_{\ast 1}^2 + s_{\ast 2}^2)/(n - 2 - w) \quad (13)$$

is an unbiased estimator for $\sigma^2$.

As a matter of fact, $\hat{\sigma}^2$ coincides with the usual estimator for $\sigma^2$ in model (5). Identity (13) follows immediately from (9) when the above $y_\ast$ is partitioned into two vectors of length $n_1$ and $n_2$, respectively.

### 3 Testing for a group effect

From Theorem 1 with $X_1$ from (6), it follows that

$$\hat{\delta}_1 = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} n_1^{-1} 1_{n_1} & 0 \\ -n_1^{-1} 1_{n_1} & n_2^{-1} 1_{n_2} \end{pmatrix} y_\ast = \begin{pmatrix} \bar{y}_{\ast 1} \\ \bar{y}_{\ast 2} - \bar{y}_{\ast 1} \end{pmatrix}. \quad (14)$$

Hence, it is seen that $|d_\ast|$ from (4) is identical to

$$|d_\ast| = \frac{\| \hat{\beta}_1 \|}{\hat{\sigma}} \quad (15)$$
with $\hat{\sigma}$ being the square root of $\hat{\sigma}^2$ from Theorem 2. The statistic $d_*$ is closely related to the test statistic $t_*$ (defined below) for the null hypothesis $H_0 : \beta_1 = 0$ in model (5).

**Theorem 3** Under the partitioned linear regression model (5) and (6) let $M_2 = I_n - P_2$, $P_2 = X_2(X_2^TX_2)^{-1}X_2^T$, and let $\gamma$ be the lower-right element of the $2 \times 2$ matrix $(X_1^TM_2X_1)^{-1}$. Then, the statistic

$$t_* = d_*/\sqrt{\gamma}$$

follows a central $t$ distribution with $n - 2 - w$ degrees of freedom, provided $\beta_1 = 0$.

In the above theorem, $\gamma$ is the scaled variance of $\hat{\beta}_1$, i.e. $\text{Var}(\hat{\beta}_1) = \hat{\sigma}^2 \gamma$. See the proof of Theorem 3 in the appendix. The standard error of $\hat{\beta}_1$ is thus $\text{se}(\hat{\beta}_1) = \hat{\sigma}\sqrt{\gamma}$ with $\hat{\sigma}$ being the square root of $\hat{\sigma}^2$ from Theorem 2.

Note that if $w = 0$, then $M_2 = I_n$, and one gets

$$(X_1^TX_1)^{-1} = \begin{pmatrix} n_1^{-1} & -n_1^{-1} \\ -n_1^{-1} & \gamma \end{pmatrix}, \quad \gamma = \frac{n_1 + n_2}{n_1n_2},$$

and hence

$$d_* = t_* \sqrt{\frac{n_1 + n_2}{n_1n_2}},$$

which is just a reformulation of (2). These considerations show that $d_*$ is a natural extension of Cohen’s $d$ in the context of additional independent variables.

**4 Effect Size in Multiple Regression**

In his Chapter 9, Cohen [3] discusses the effect size measure $f^2$ based on the $F$ test of a linear hypothesis. It may be applied when $X_1$ does not only comprise intercept and one dummy as under model (5), but a total of $u$ independent variables of arbitrary type. Then, it might be of interest to measure the effect size of the set of variables in $X_1$ given the set in $X_2$, which is Cohen’s case 1. Cohen suggests values $f^2 = 0.02$, $f^2 = 0.15$ and $f^2 = 0.35$ for a small, medium and large effect, respectively. Since the measure $d_*$ refers to one dummy ($u = 1$), one might expect a relationship between $d_*$ and the corresponding $f^2$. Actually, as noted in our Remark below, such a relationship can be specified.

The measure $f^2$ for Cohen’s case 1 is given by

$$f^2 = F\frac{u}{v}, \quad v = n - u - w - 1,$$
where under model (5) $F$ is the $F$ statistic for testing the null hypothesis $H_0 : \beta_1 = 0$. From (9.2.3) in Cohen [3],

$$f^2 = \frac{R^2 - R_0^2}{1 - R^2},$$

(20)

where $R^2$ is the coefficient of determination from model (5) and $R_0^2$ is the coefficient of determination in the reduced model with $\beta_1 = 0$, admitting model matrix $X_0 = (1_n, X_2)$. If $P$ denotes the orthogonal projector onto the column space of the model matrix of a regression model with intercept, the coefficient of determination is given by

$$R^2 = 1 - \frac{y^T (I_n - P)y}{y^T Cy}$$

(21)

with $C = I_n - n^{-1}1_n 1_n^T$ being the so-called centering matrix, e.g. see [9, Sect. 6.2]. From this, (20) becomes

$$f^2 = \frac{y^T (P - P_0)y}{y^T (I_n - P)y}$$

(22)

with $P = X(X^T X)^{-1}X^T$, $X = (X_1, X_2)$, and $P_0 = X_0(X_0^T X_0)^{-1}X_0^T$. In view of $\text{rank}(P - P_0) = 1$ and $\text{rank}(I_n - P) = n - (2 + w)$, the corresponding $F$ statistic reads

$$F = \frac{y^T (P - P_0)y/\text{rank}(P - P_0)}{y^T (I_n - P)y/\text{rank}(I_n - P)}.$$ 

(23)

Then, from Theorem 3.2.1 (ii) in Christensen [2], $F$ follows a central $F$ distribution with 1 and $n - 2 - w$ degrees of freedom, provided $\beta_1 = 0$.

Now, it is well known and readily verified that the squared $t$ statistic for the null hypothesis $H_0 : \beta_1 = 0$ is identical to the test statistic of the $F$ test for the very same hypothesis. Thus, by combining (16) and (19), the following is true.

**Remark** The identity

$$f^2 = \frac{d^2/\gamma}{n - 2 - w}$$

(24)

specifies the exact relationship between the effect size measures $f^2$ and $d_*$ from above.

Note that in case $w = 0$, the above identity reads

$$f^2 = d^2 \frac{n_1 n_2}{n(n - 2)},$$

(25)
which slightly differs from formula (9.3.5) in Cohen [3] and lacks some of its beauty. Since (25) is expected to coincide with (9.3.5), this reveals a fallacy in the latter formula. Formula (25) may also be verified independently by directly assuming model (5) without any additional independent variables $X_2$. In most cases, actual computations of the two formulas in question only differ in a later digit after the decimal point (say the fifth or sixth), so the difference usually has no practical meaning. The correctness of (24) and (25) is additionally confirmed by applications to real data.

5 Data Example

To give a possible outline for applications and an illustration of the previous formulas, we employ a data set available from the UCI machine learning repository, see Dua and Graff [6]. It contains student achievement in secondary education of two Portuguese schools, see Cortez and Silva [4]. In the following, computations are carried out with the statistical software R [17].

As the dependent variable $y$, we consider the final grade with integer values ranging between 0 and 20 in Portuguese language (variable G3) of $n_1 = 383$ female and $n_2 = 266$ male students. The dummy variable $z$ takes values 0 for female and 1 for male. As also indicated by Figure 1, female students perform better with an average of $\bar{y}_1 = 12.25326$ compared to $\bar{y}_2 = 11.40602$ for male students. The corresponding equal variances two-sample test statistic admits $t = 3.310938$ with p value $0.0009815287$. Although this implies strong significance, the corresponding effect size from (2) reads $d = 0.264261$, thereby indicating only a slightly more than low effect. This value may also be obtained by function cohen_s_d from the R package effectsize, see Ben-Shachar et al [1].

As additional independent variables, we consider the education of the father (Fedu) $x_1$ and the travel time from home to school (traveltime) $x_2$. Both variables are measured on an ordinal scale with integer values ranging from 0 to 4 and 1 to 4, respectively, and are included as quantitative variables in our regression approach, implying $w = 2$.  

![Fig. 1 Frequency of final results (variable G3) in Portuguese language of $n = 649$ students](image)
Table 1  Extract from R output by fitting the complete model (5) with function lm, where variable sex is included as a factor, implying that $ZM$ represents the dummy $z$

|             | Estimate | Std. Error | t value | Pr(>|t|)     |
|-------------|----------|------------|---------|-------------|
| (Intercept) | 11.4138  | 0.4306     | 26.507  | < 2e-16***  |
| ZM          | -0.9406  | 0.2503     | -3.759  | 0.000186*** |
| X1          | 0.6096   | 0.1144     | 5.329   | 1.37e-07*** |
| X2          | -0.3369  | 0.1676     | -2.010  | 0.044826*   |

Least squares estimates of the coefficients with corresponding $t$ statistic values are given in Table 1. The intercept estimate $\hat{\beta}_0 = 11.41385$ is the average of the adjusted final grade $y_*$ of females, while the dummy variable estimate $\hat{\beta}_1 = -0.9406209$ is the difference between the average of $y_*$ in the male group minus the average in the female group as given in 14. As it is also seen, better father’s education comes along with better grades (positive $\hat{\beta}_2$) while longer travel times to school come along with lower grades (negative $\hat{\beta}_3$) when the other variables are held constant, respectively. The coefficient of determination from this model reads $R^2 = 0.07238847$, while for the reduced model (omitting sex) it is $R^2_0 = 0.05207054$. Then, $f^2$ from (20) is $f^2 = 0.0219035$, implying a slightly more than small effect concerning the difference between female and male final grades, ceteris paribus.

The relationship (24) may also be used to infer Cohen’s $d_*$ from $f^2$. For this

$$ (X_1^T M_2 X_1)^{-1} = \begin{pmatrix} 0.019062813 & -0.001591796 \\ -0.001591796 & 0.006438624 \end{pmatrix}, \quad \text{(26)}$$

the lower-right element being the scaled variance $\gamma$ of $\hat{\beta}_1$. With $\hat{\sigma} = 3.118756$ it follows $d_* = 0.3016013$. This indicates a slightly stronger effect when variables $x_1$ and $x_2$ are held constant than seen before from $d = 0.264261$ not considering any additional independent variables at all. Alternatively, $d_*$ may be computed by either of the formulas (15) or (4) yielding the very same absolute value.

Acknowledgements  Annette Möller acknowledges support by the Helmholtz Association’s pilot project “Uncertainty Quantification”. We are thankful to two anonymous referees for helpful comments which improved the presentation.

Funding  Open Access funding enabled and organized by Projekt DEAL.

Declaration

Conflict of interest  The authors declare that they have no conflict of interest.

Open Access  This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted
6 Appendix A Proofs of Theorems

In this section, we give short proofs of the three stated theorems. Although most of the supporting formulas and derivations may already be found throughout the literature, we present them here in order to provide a unified and self-contained treatment of the topic.

Proof of Theorem 1 From (7), it follows that

\[ y - X_1 \hat{\delta}_1 - X_2 \hat{\delta} = (I_n - P)y, \tag{A1} \]

where \( P = X(X^T X)^{-1}X^T \) is the orthogonal projector onto the column space of the model matrix \( X = (X_1, X_2) \). Since the column space of \( X_1 \) is contained in the column space of \( X \), it follows

\[ PP_1 = P_1 = P_1 P \tag{A2} \]

for \( P_1 = X_1(X_1^T X_1)^{-1}X_1^T \). Hence,

\[ X_1 \hat{\delta}_1 = P_1 y = P_1 (y - X_2 \hat{\delta}_2). \tag{A3} \]

Since \( X_1 \) is of full column rank, Theorem 1 follows. \( \square \)

Proof of Theorem 2 From (8)

\[ y_* = y - X_2 \hat{\delta}_2 = (I_n - X_2(X_2^T M_1 X_2)^{-1}X_2^T M_1)y. \tag{A4} \]

Then,

\[ (y - X_2 \hat{\delta}_2)^T M_1 (y - X_2 \hat{\delta}_2) = y^T Ly, \quad L = M_1 - M_1 X_2(X_2^T M_1 X_2)^{-1}X_2^T M_1. \tag{A5} \]

From Theorem 1.3.2 in Christensen [2], the expectation of the quadratic form \( y^T Ly \) is given by

\[ \mathbb{E}(y^T Ly) = \sigma^2 \text{trace}(L) + \mu^T A \mu \tag{A6} \]

with \( \mu = X_1 \delta_1 + X_2 \delta_2 \). From (9) it follows \( \text{trace}(M_1) = n_1 - 1 + n_2 - 1 \) and therefore \( \text{trace}(L) = n_1 + n_2 - 2 - w \). In addition \( LX_1 = 0 \) and \( LX_2 = 0 \), implying \( \mathbb{E}(y^T Ly) = \sigma^2(n_1 + n_2 - 2 - w) \). This gives Theorem 2. \( \square \)
Proof of Theorem 3  Similarly to (8), the Frisch–Waugh–Lovell theorem states that
\[
\hat{\delta}_1 = (X_1^T M_2 X_1)^{-1} X_1^T M_2 y, \quad M_2 = I_n - X_2 (X_2^T X_2)^{-1} X_2^T.
\]
(A7)

Hence, it is seen that \( \hat{\delta}_1 \) follows a multivariate normal distribution with expectation vector \( \delta_1 \) and variance-covariance matrix \( \sigma^2 (X_1^T M_2 X_1)^{-1} \), e.g. by Exercise 1.8 in Christensen [2]. Then, \( \hat{\beta}_1 \) follows a univariate normal distribution with expectation \( \beta_1 \) and variance \( \sigma^2 \gamma \), where \( \gamma \) is the lower-right element of the \( 2 \times 2 \) matrix \( (X_1^T M_2 X_1)^{-1} \).

If \( \beta_1 = 0 \), then
\[
U = \hat{\beta}_1 / (\sigma \sqrt{\gamma})
\]
(A8)

follows a standard normal distribution. Let
\[
V = y^T L y / \sigma^2
\]
(A9)

with \( L \) from (A5). It is readily verified that \( L \) is symmetric and idempotent, so that Theorem 1.3.3 in Christensen [2] implies that \( V \) follows a central \( \chi^2 \) distribution with \( \text{rank}(L) = \text{trace}(L) = n - 2 - w \) degrees of freedom. From
\[
\hat{\delta}_1 = F y, \quad F = (X_1^T X_1)^{-1} X_1^T (I_n - X_2 (X_2^T M_1 X_2)^{-1} X_2^T M_1)
\]
(A10)

it is seen that \( FL = 0 \), implying that \( F y \) ad \( y^T L y \) are independent, see Theorem 1.3.7 in Christensen [2]. Then, also \( U \) and \( V \) are independent and
\[
t^* = \frac{U}{\sqrt{V/(n - 2 - w)}} = \frac{\hat{\beta}_1 / \sqrt{\gamma}}{\hat{\sigma}}
\]
(A11)

follows a central \( t \) distribution with \( n - 2 - w \) degrees of freedom when \( \beta_1 = 0 \). \( \square \)

References

1. Ben-Shachar MS, Lüdecke D, Makowski D (2020) Effectsize: estimation of effect size indices and standardized parameters. J Open Source Soft 5:2815. https://doi.org/10.21105/joss.02815
2. Christensen R (2020) Plane answers to complex questions, 5th edn. Springer, Germany
3. Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd edn. Lawrence Erlbaum Associates, New Jersey
4. Cortez P, Silva AMG (2008) Using data mining to predict secondary school student performance. In: Brito A, Teixeira J (eds) EUROYSIS, pp 5–22
5. Ding P (2021) The Frisch-Waugh-Lovell theorem for standard errors. Stat Probabil Lett 168(108):945
6. Dua D, Graff C (2017) UCI machine learning repository. http://archive.ics.uci.edu/ml
7. Fiebig DG, Bartels R, Krämer W (1996) The Frisch-Waugh theorem and generalized least squares. Economet Rev 15:431–443
8. Frisch R, Waugh FV (1933) Partial time regressions as compared with individual trends. Econometrica: J Econ Soc 1:387–401
9. Groß J (2003) Linear Regression. Springer. Lecture Notes in Statistics 175
10. Groß J, Puntenan S (2000) Estimation under a general partitioned linear model. Linear Algebra Appl 321:131–144

Springer
11. Groß J, Puntanen S (2005) Extensions of the Frisch-Waugh-Lovell theorem. Discuss Math Probabil Stat 25:39–49
12. Hedges LV (1981) Distribution theory for Glass’s estimator of effect size and related estimators. J Educ Stat 6:107–128
13. Kraemer HC (1983) Theory of estimation and testing of effect sizes: Use in meta-analysis. J Educ Stat 8:93–101
14. Lipsey MW, Wilson DB (2001) Practical meta-analysis. SAGE publications, Inc, California
15. Lovell MC (1963) Seasonal adjustment of economic time series and multiple regression analysis. J Am Stat Assoc 58(304):993–1010
16. Puntanen S (1996) Some matrix results related to a partitioned singular linear model. Commun Stat-Theory Methods 25:269–279
17. R Core Team (2022) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/
18. Wilkinson L (1999) Statistical methods in psychology journals: Guidelines and explanations. Am Psychol 54:594–604
19. Wilson DB (2016) Formulas used by the “practical meta-analysis effect size calculator”. Practical meta-analysis https://mason.gmu.edu/~dwilsoneb/

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.