Web-based Supporting Materials for "Improving likelihood-based inference in control rate regression" by Annamaria Guolo

Web Appendix A: Derivation of Skovgaard's statistic

Given the framework described in Section 2 of the paper, the log-likelihood function $\ell(\theta)$ for the whole parameter vector θ is

$$\ell(\theta) \propto -\frac{1}{2}\sum_{i=1}^{n}\log|V_i| - \frac{1}{2}\sum_{i=1}^{n}(y_i - f_i)^\top V_i^{-1}(y_i - f_i),$$

where $y_i = (\hat{\eta}_i, \hat{\xi}_i)^\top$ is the observed value of the random vector Y_i with mean vector f_i and variance/covariance matrix V_i, following the notation in Section 3.1 of the paper. The score vector

$$\ell_\theta(\theta) = \begin{bmatrix} \ell_{\beta_0}(\theta) \\ \ell_{\beta_1}(\theta) \\ \ell_{\mu}(\theta) \\ \ell_{\tau^2}(\theta) \\ \ell_{\sigma^2}(\theta) \end{bmatrix}$$

has components

$$\begin{aligned} \ell_{\beta_j}(\theta) &= -\frac{1}{2}\sum_{i=1}^{n}\text{trace}\left(V_i^{-1}V_{i,\beta_j}\right) \\ &\quad -\frac{1}{2}\sum_{i=1}^{n}\left(y_i^\top V_{i,\beta_j}^{-1}y_i - 2f_{i,\beta_j}^\top V_i^{-1}y_i - 2f_i^\top V_{i,\beta_j}^{-1}y_i + 2f_{i,\beta_j}V_i^{-1}f_i + f_i^\top V_{i,\beta_j}^{-1}f_i\right), j = 0, 1, \end{aligned}$$

$$\ell_{\mu}(\theta) = \sum_{i=1}^{n} f_{i,\mu}^{\top} V_i^{-1}(y_i - f_i)$$

and

$$\ell_{\psi_j}(\theta) = -\frac{1}{2}\sum_{i=1}^{n} \text{trace}\left(V_i^{-1} V_{i,\psi_j}\right) - \frac{1}{2}\sum_{i=1}^{n}\left(y_i^{\top} V_{i,\psi_j}^{-1} y_i - 2 f_i^{\top} V_{i,\psi_j}^{-1} y_i + f_i^{\top} V_{i,\psi_j}^{-1} f_i\right),\ \psi_j \in \{\tau^2, \sigma^2\}.$$

The expected information matrix

$$i(\theta) = \begin{bmatrix} i_{\beta_0\beta_0}(\theta) & i_{\beta_0\beta_1}(\theta) & i_{\beta_0\mu}(\theta) & i_{\beta_0\tau^2}(\theta) & i_{\beta_0\sigma^2}(\theta) \\ i_{\beta_0\beta_1}(\theta) & i_{\beta_1\beta_1}(\theta) & i_{\beta_1\mu}(\theta) & i_{\beta_1\tau^2}(\theta) & i_{\beta_1\sigma^2}(\theta) \\ i_{\beta_0\mu}(\theta) & i_{\beta_1\mu}(\theta) & i_{\mu\mu}(\theta) & i_{\mu\tau^2}(\theta) & i_{\mu\sigma^2}(\theta) \\ i_{\beta_0\tau^2}(\theta) & i_{\beta_1\tau^2}(\theta) & i_{\mu\tau^2}(\theta) & i_{\tau^2\tau^2}(\theta) & i_{\tau^2\sigma^2}(\theta) \\ i_{\beta_0\sigma^2}(\theta) & i_{\beta_1\sigma^2}(\theta) & i_{\mu\sigma^2}(\theta) & i_{\tau^2\sigma^2}(\theta) & i_{\sigma^2\sigma^2}(\theta) \end{bmatrix}$$

has generic component

$$i_{\theta_j\theta_k} = \frac{1}{2}\sum_{i=1}^{n} \text{trace}\left(V_{i,\theta_k}^{-1} V_{i,\theta_j} + V_i^{-1} V_{i,\theta_j\theta_k} - V_{i,\theta_j}^{-1} V_{i,\theta_j\theta_k} V_{i,\theta_j}^{-1} V_i\right) + \sum_{i=1}^{n} f_{i,\theta_j} V_i^{-1} f_{i,\theta_k},\ \theta_j, \theta_k \in \theta,$$

where $V_{i,\theta_j\theta_k}$ denotes the second derivative of V_i with respect to $\theta_j, \theta_k \in \theta$.
In order to derive the components of S and q, consider that

$$\text{cov}\left(Y_i^{\top} \hat{V}_{i,\theta_j}^{-1} Y_i, Y_i^{\top} \tilde{V}_{i,\theta_k} Y_i\right) = \text{trace}\left(\hat{V}_{i,\theta_j}^{-1} \hat{V}_i \tilde{V}_{i,\theta_k}^{-1} \hat{V}_i\right) + 4 \hat{f}_i^{\top} \hat{V}_{i,\theta_j}^{-1} \hat{V}_i \tilde{V}_{i,\theta_k}^{-1} \hat{f}_i,$$

$$\text{cov}\left(Y_i^{\top} \hat{V}_{i,\theta_j}^{-1} Y_i, Y_i^{\top} \tilde{V}_{i,\theta_k} Y_i\right) = 2 f_i^{\top} \hat{V}_{i,\theta_j}^{-1} \hat{V}_i \tilde{V}_i^{-1} \tilde{f}_{i,\theta_j}$$

and

$$\text{cov}\left(\hat{f}_{i,\theta_j} \hat{V}_i^{-1} Y_i, \tilde{f}_i^{\top} \tilde{V}_{i,\theta_k}^{-1} Y_i\right) = \hat{f}_{i,\theta_j}^{\top} \hat{V}_i^{-1} \hat{V}_i \tilde{V}_{i,\theta_k}^{-1} \tilde{f}_i,$$

for $\theta_j, \theta_k \in \theta$.

Then,

$$\begin{aligned} S_{\beta_j,\beta_k}(\theta) &= \text{cov}\left\{\ell_{\beta_j}(\theta_1), \ell_{\beta_k}(\theta_2)\right\}\big|_{\theta_1=\hat{\theta},\theta_2=\tilde{\theta}} \\ &= \frac{1}{4}\text{cov}\sum_{i=1}^{n}\Big(Y_i \hat{V}_{i,\beta_j}^{-1} Y_i - 2\hat{f}_{i,\beta_j}^{\top} \hat{V}_i^{-1} Y_i - 2\hat{f}_i \hat{V}_{i,\beta_j}^{-1} Y_i, Y_i^{\top} \tilde{V}_{i,\beta_k}^{-1} Y_i - 2\tilde{f}_{i,\beta_k}^{\top} \tilde{V}_i^{-1} Y_i \\ &\quad -2\tilde{f}_i^{\top} \tilde{V}_{i,\beta_k}^{-1} Y_i\Big) \\ &= \sum_{i=1}^{n}\left\{\frac{1}{2}\text{trace}\left(\hat{V}_{i,\beta_j}^{-1} \hat{V}_i \tilde{V}_{i,\beta_k}^{-1} \hat{V}_i\right) + \hat{f}_{i,\beta_j}^{\top} \tilde{V}_{i,\beta_k}^{-1}\left(\tilde{f}_i - \hat{f}_i\right) + \hat{f}_{i,\beta_k} \tilde{V}_i^{-1} \tilde{f}_{i,\beta_k}\right\},\ j,k = 0,1 \end{aligned}$$

$$
\begin{aligned}
S_{\beta_j,\mu}(\theta) &= \operatorname{cov}\left\{\ell_{\beta_j}(\theta_1), \ell_\mu(\theta_2)\right\}\big|_{\theta_1=\hat{\theta},\theta_2=\tilde{\theta}} \\
&= \frac{1}{2}\operatorname{cov}\sum_{i=1}^{n}\left(Y_i^\top \hat{V}_{i,\beta_j}^{-1} Y_i - \hat{f}_{i,\beta_j}^\top \hat{V}_i^{-1} Y_i - 2\hat{f}_i \hat{V}_{i,\beta_j}^{-1} Y_i, \tilde{f}_{i,\mu}^\top \tilde{V}_i^{-1} Y_i\right) \\
&= \sum_{i=1}^{n} \hat{f}_{i,\beta_j}^\top \tilde{V}_i^{-1} \tilde{f}_{i,\mu},\ \ j = 0, 1
\end{aligned}
$$

$$
\begin{aligned}
S_{\mu,\mu}(\theta) &= \operatorname{cov}\left\{\ell_\mu(\theta_1), \ell_\mu(\theta_2)\right\}\big|_{\theta_1=\hat{\theta},\theta_2=\tilde{\theta}} \\
&= \operatorname{cov}\sum_{i=1}^{n}\left(\hat{f}_{i,\mu}^\top \hat{V}_i^{-1} Y_i, \tilde{f}_{i,\mu}^\top \tilde{V}_i^{-1} Y_i\right) \\
&= \sum_{i=1}^{n} \hat{f}_{i,\mu}^\top \tilde{V}_i^{-1} \tilde{f}_{i,\mu}
\end{aligned}
$$

$$
\begin{aligned}
S_{\beta_j,\psi_k}(\theta) &= \operatorname{cov}\left\{\ell_{\beta_j}(\theta_1), \ell_{\psi_k}(\theta_2)\right\}\big|_{\theta_1=\hat{\theta},\theta_2=\tilde{\theta}} \\
&= \frac{1}{2}\operatorname{cov}\sum_{i=1}^{n}\left(Y_i^\top \hat{V}_{i,\beta_j}^{-1} Y_i - 2\hat{f}_{i,\beta_j}^\top \hat{V}_i^{-1} Y_i - 2\hat{f}_i^\top \hat{V}_{i,\beta_j}^{-1} Y_i, Y_i^\top \tilde{V}_{i,\psi_k}^{-1} Y_i - 2\tilde{f}_i^\top \tilde{V}_{i,\psi_k}^{-1} Y_i\right) \\
&= \sum_{i=1}^{n}\left\{\frac{1}{2}\operatorname{trace}\left(\hat{V}_{i,\beta_j}^{-1} \hat{V}_i \tilde{V}_{i,\psi_k}^{-1} \hat{V}_i\right) + \hat{f}_{i,\beta_j}^\top \tilde{V}_{i,\psi_k}^{-1}\left(\tilde{f}_i - \hat{f}_i\right)\right\},\ \ j = 0, 1,\ \ \psi_k \in \{\tau^2, \sigma^2\}
\end{aligned}
$$

$$
\begin{aligned}
S_{\mu,\psi_k}(\theta) &= \operatorname{cov}\left\{\ell_\mu(\theta_1), \ell_{\psi_k}(\theta_2)\right\}\big|_{\theta_1=\hat{\theta},\theta_2=\tilde{\theta}} \\
&= -\frac{1}{2}\operatorname{cov}\sum_{i=1}^{n}\left(\hat{f}_{i,\mu} \hat{V}_i^{-1} Y_i, Y_i^\top \tilde{V}_i^{-1} Y_i - 2\tilde{f}_i^\top \tilde{V}_{i,\psi_k}^{-1} Y_i\right) \\
&= \sum_{i=1}^{n}\left\{\frac{1}{2}\operatorname{trace}\left(\hat{V}_{i,\mu}^{-1} \hat{V}_i \tilde{V}_{i,\psi_k}^{-1} \hat{V}_i\right) + \hat{f}_{i,\mu}^\top \tilde{V}_{i,\psi_k}^{-1}\left(\tilde{f}_i - \hat{f}_i\right)\right\},\ \ \psi_k \in \{\tau^2, \sigma^2\}
\end{aligned}
$$

$$
\begin{aligned}
S_{\psi_j,\psi_k}(\theta) &= \operatorname{cov}\left\{\ell_{\psi_j}(\theta_1), \ell_{\psi_k}(\theta_2)\right\}\big|_{\theta_1=\hat{\theta},\theta_2=\tilde{\theta}} \\
&= \frac{1}{4}\operatorname{cov}\sum_{i=1}^{n}\left(Y_i^\top \hat{V}_{i,\psi_j}^{-1} Y_i - 2\hat{f}_i^\top \hat{V}_{i,\psi_j}^{-1} Y_i, Y_i^\top \tilde{V}_{i,\psi_k}^{-1} Y_i - 2\tilde{f}_i^\top \tilde{V}_{i,\psi_k}^{-1} Y_i\right) \\
&= \frac{1}{2}\sum_{i=1}^{n}\operatorname{trace}\left(\hat{V}_{i,\psi_j}^{-1} \hat{V}_i \tilde{V}_{i,\psi_k}^{-1} \hat{V}_i\right),\ \ \psi_j, \psi_k \in \{\tau^2, \sigma^2\}
\end{aligned}
$$

$$
\begin{aligned}
S_{\mu,\beta_j}(\theta) &= \operatorname{cov}\{\ell_\mu(\theta_1), \ell_{\beta_k}(\theta_2)\}\big|_{\theta_1=\hat\theta,\theta_2=\tilde\theta} \\
&= -\frac{1}{2}\operatorname{cov}\sum_{i=1}^n \left(\hat f_{i,\mu}^\top \hat V_i^{-1} Y_i, Y_i^\top \tilde V_{i,\beta_j}^{-1} Y_i - 2\tilde f_{i,\beta_j}^\top \tilde V_i^{-1} Y_i - 2\tilde f_i \tilde V_{i,\beta_j}^{-1} Y_i\right) \\
&= \sum_{i=1}^n \left(\hat f_{i,\mu}^\top \tilde V_i^{-1} \tilde f_{i,\beta_j} + \hat f_{i,\mu}^\top \tilde V_{i,\beta_j}^{-1} \tilde f_i - \hat f_i^\top \tilde V_{i,\beta_j}^{-1} \hat f_{i,\mu}\right),\ j=0,1
\end{aligned}
$$

$$
\begin{aligned}
S_{\psi_j,\beta_k}(\theta) &= \operatorname{cov}\left\{\ell_{\psi_j}(\theta_1), \ell_{\beta_k}(\theta_2)\right\}\Big|_{\theta_1=\hat\theta,\theta_2=\tilde\theta} \\
&= \frac{1}{4}\operatorname{cov}\sum_{i=1}^n \left(Y_i^\top \hat V_{i,\psi_j}^{-1} Y_i - 2\hat f_i^\top \hat V_{i,\psi_j}^{-1} Y_i, Y_i^\top \tilde V_{i,\beta_k}^{-1} Y_i - 2\tilde f_{i,\beta_k}^\top \tilde V_i^{-1} Y_i - 2\tilde f_i^\top \tilde V_{i,\beta_k}^{-1} Y_i\right) \\
&= \frac{1}{2}\sum_{i=1}^n \operatorname{trace}\left(\hat V_{i,\psi_j}^{-1} \hat V_i \tilde V_{i,\beta_k}^{-1} \hat V_i\right), \psi_j \in \{\tau^2, \sigma^2\}, k=0,1
\end{aligned}
$$

$$
\begin{aligned}
S_{\psi_j,\mu}(\theta) &= \operatorname{cov}\left\{\ell_{\psi_j}(\theta_1), \ell_\mu(\theta_2)\right\}\Big|_{\theta_1=\hat\theta,\theta_2=\tilde\theta} \\
&= -\frac{1}{2}\operatorname{cov}\sum_{i=1}^n \left(Y_i^\top \hat V_{i,\psi_k}^{-1} Y_i - 2\hat f_i^\top \hat V_{i,\psi_k}^{-1} Y_i, \tilde f_{i,\mu}^\top \tilde V_i^{-1} Y_i\right) \\
&= 0,\ \psi_j \in \{\tau^2, \sigma^2\}
\end{aligned}
$$

$$
\begin{aligned}
q_{\beta_j}(\theta) &= \operatorname{cov}\left\{\ell_{\beta_\mathrm{j}}(\theta_1), \ell(\theta_1) - \ell(\theta_2)\right\}\Big|_{\theta_1=\hat\theta,\theta_2=\tilde\theta} \\
&= \frac{1}{4}\operatorname{cov}\sum_{i=1}^n \Big(Y_i^\top \hat V_{i,\beta_j}^{-1} Y_i - 2\hat f_{i,\beta_j}^\top \hat V_i^{-1} Y_i - 2\hat f_i^\top \hat V_{i,\beta_j}^{-1} Y_i, Y_i^\top \hat V_i^{-1} Y_i \\
&\quad -2 f_i^\top \hat V_i^{-1} Y_i - Y_i^\top \tilde V_i^{-1} Y_i + 2\tilde f_i \tilde V_i^{-1} Y_i\Big) \\
&= \sum_{i=1}^n \left[\frac{1}{2}\operatorname{trace}\left\{\hat V_{i,\beta_j}^{-1} \hat V_i \left(\hat V_i^{-1} - \tilde V_i^{-1}\right) \hat V_i\right\} + \hat f_{i,\beta_j}^\top \tilde V_i^{-1}\left(\hat f_i - \tilde f_i\right)\right],\ j=0,1
\end{aligned}
$$

$$
q_\mu(\theta) = \operatorname{cov}\{\ell_\mu(\theta_1), \ell(\theta_1) - \ell(\theta_2)\}\big|_{\theta_1=\hat\theta,\theta_2=\tilde\theta}
$$

$$
\begin{aligned}
&= \frac{1}{4}\mathrm{cov}\sum_{i=1}^{n}\Big(Y_i^\top \hat{V}_{i,\mu}^{-1}Y_i - 2\hat{f}_{i,\mu}^\top \hat{V}_i^{-1}Y_i - 2\hat{f}_i^\top \hat{V}_{i,\mu}^{-1}Y_i, Y_i^\top \hat{V}_i^{-1}Y_i \\
&\quad -2f_i^\top \hat{V}_i^{-1}Y_i - Y_i^\top \tilde{V}_i^{-1}Y_i + 2\tilde{f}_i\tilde{V}_i^{-1}Y_i\Big) \\
&= \sum_{i=1}^{n}\left[\frac{1}{2}\mathrm{trace}\left\{\hat{V}_{i,\mu}^{-1}\hat{V}_i\left(\hat{V}_i^{-1}-\tilde{V}_i^{-1}\right)\hat{V}_i\right\} + \hat{f}_{i,\mu}^\top \tilde{V}_i^{-1}\left(\hat{f}_i - \tilde{f}_i\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
q_{\psi_j}(\theta) &= \mathrm{cov}\left\{\ell_{\psi_\mathrm{j}}(\theta_1), \ell(\theta_1)-\ell(\theta_2)\right\}\Big|_{\theta_1=\hat{\theta},\theta_2=\tilde{\theta}} \\
&= \frac{1}{4}\mathrm{cov}\sum_{i=1}^{n}\Big(Y_i^\top \hat{V}_{i,\psi_j}^{-1}Y_i - 2\hat{f}_{i,\psi_j}^\top \hat{V}_i^{-1}Y_i - 2\hat{f}_i^\top \hat{V}_{i,\mu}^{-1}Y_i, Y_i^\top \hat{V}_i^{-1}Y_i \\
&\quad -2f_i^\top \hat{V}_i^{-1}Y_i - Y_i^\top \tilde{V}_i^{-1}Y_i + 2\tilde{f}_i\tilde{V}_i^{-1}Y_i\Big) \\
&= \frac{1}{2}\sum_{i=1}^{n}\left\{\mathrm{trace}\left(\hat{V}_{\psi_j}^{-1}\hat{V}_i\right) - \mathrm{trace}\left(\hat{V}_{\psi_j}^{-1}\hat{V}_i\tilde{V}_i^{-1}\hat{V}_i\right)\right\},\ \psi_j \in \{\tau^2,\sigma^2\}
\end{aligned}
$$

Web Appendix B: Simulation results

This web appendix reports a portion of the results of the simulation study carried out to evaluate the performance of Skovgaard's statistic against competing approaches, as described in Section 4 of the main manuscript.

Simulations refer to different scenarios with decreasing log event rate in the treatment group corresponding to different values for $(\beta_0, \beta_1, \mu)^\top$. Different values for the variance components τ^2 and σ^2 are considered as well. The examined situations and the corresponding simulation results in terms of empirical coverage probabilities of the nominally 95% confidence interval for β_1 are listed below.

- Scenario with $(\beta_0, \beta_1, \mu)^\top = (-1.5, 1, -0.5)^\top$, called *scenario 2* in the main text, with σ^2 equal to 1: Figure B1

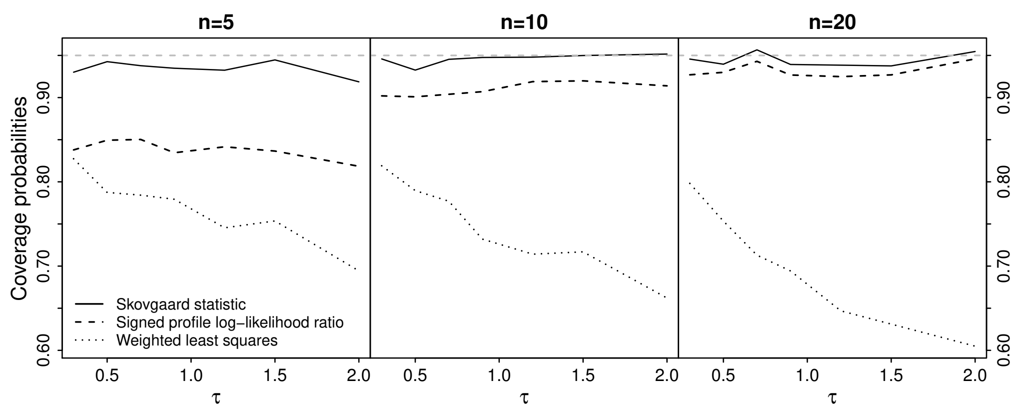

Figure B1: Empirical coverage probabilities of the nominally 95% confidence interval for β_1 when $(\beta_0, \beta_1, \mu)^\top = (-1.5, 1, -0.5)^\top$, under increasing sample size n and square root τ of the variance component τ^2. Variance component $\sigma^2 = 1$. The plotted curves correspond to Skovgaard's statistic (solid), the signed profile log-likelihood ratio statistic (dashed), the weighted least squares approach (dotted). The dashed, grey horizontal line is the nominal level.

- Scenario with $(\beta_0, \beta_1, \mu)^\top = (-1.5, 1, -2.5)^\top$, called *scenario 3* in the main text, with σ^2 equal to 1: Figure B2

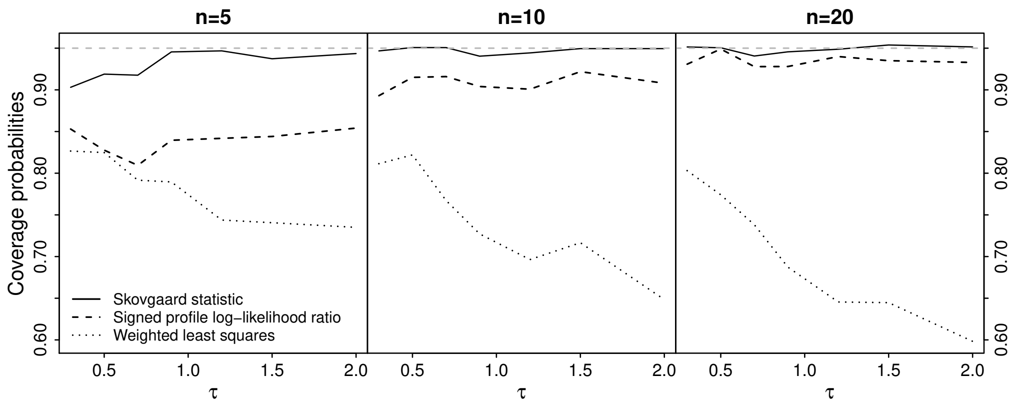

Figure B2: Empirical coverage probabilities of the nominally 95% confidence interval for β_1 when $(\beta_0, \beta_1, \mu)^\top = (-1.5, 1, -2.5)^\top$, under increasing sample size n and square root τ of the variance component τ^2. Variance component $\sigma^2 = 1$. The plotted curves correspond to Skovgaard's statistic (solid), the signed profile log-likelihood ratio statistic (dashed), the weighted least squares approach (dotted). The dashed, grey horizontal line is the nominal level.

- Scenario with $(\beta_0, \beta_1, \mu)^\top = (-3, 1, -2)^\top$, called *scenario 4* in the main text, with σ^2 equal to 1: Figure B3

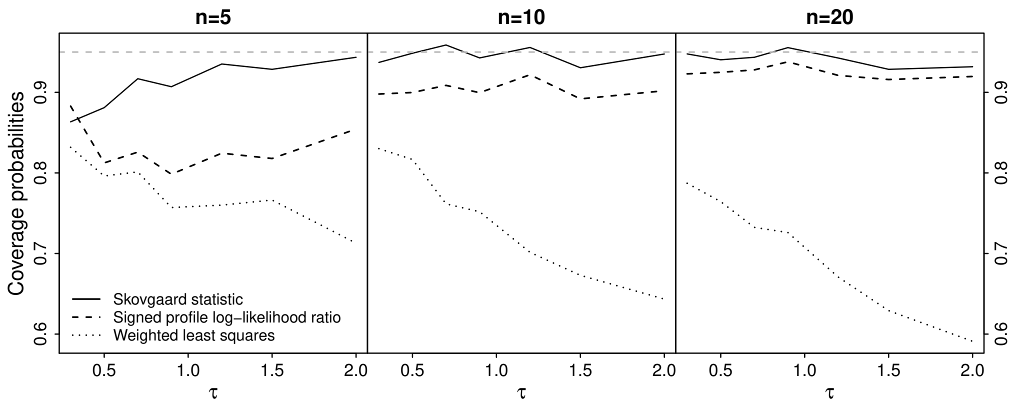

Figure B3: Empirical coverage probabilities of the nominally 95% confidence interval for β_1 when $(\beta_0, \beta_1, \mu)^\top = (-3, 1, -2)^\top$, under increasing sample size n and square root τ of the variance component τ^2. Variance component $\sigma^2 = 1$. The plotted curves correspond to Skovgaard's statistic (solid), the signed profile log-likelihood ratio statistic (dashed), the weighted least squares approach (dotted). The dashed, grey horizontal line is the nominal level.

- Scenario with $(\beta_0, \beta_1, \mu)^\top = (0, 1, 1)^\top$, called *scenario 1* in the main text, with τ^2 equal to 1.2: Figure B4

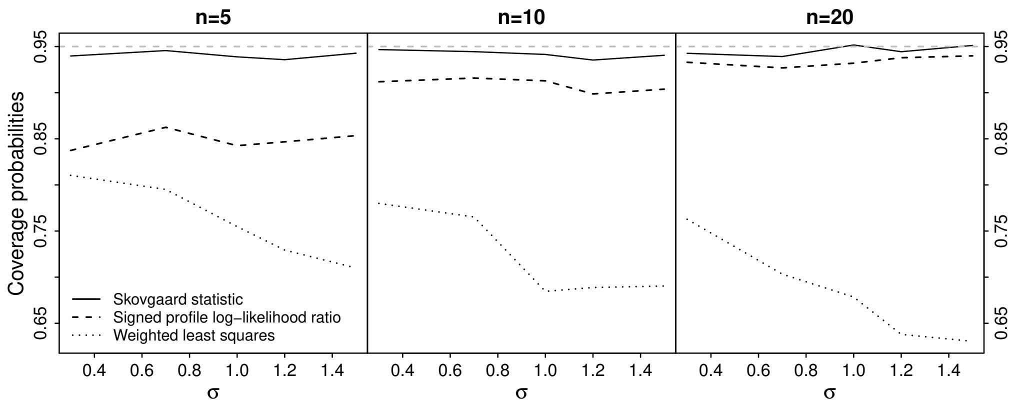

Figure B4: Empirical coverage probabilities of the nominally 95% confidence interval for β_1 when $(\beta_0, \beta_1, \mu)^\top = (0, 1, 1)^\top$, under increasing sample size n and square root σ of the variance σ^2 in the control group. Variance component $\tau^2 = 1.2$. The plotted curves correspond to Skovgaard's statistic (solid), the signed profile log-likelihood ratio statistic (dashed), the weighted least squares approach (dotted). The dashed, grey horizontal line is the nominal level.

- Scenario with $(\beta_0, \beta_1, \mu)^\top = (-1.5, 1, -0.5)^\top$, called *scenario 2* in the main text, with τ^2 equal to 1.2: Figure B5

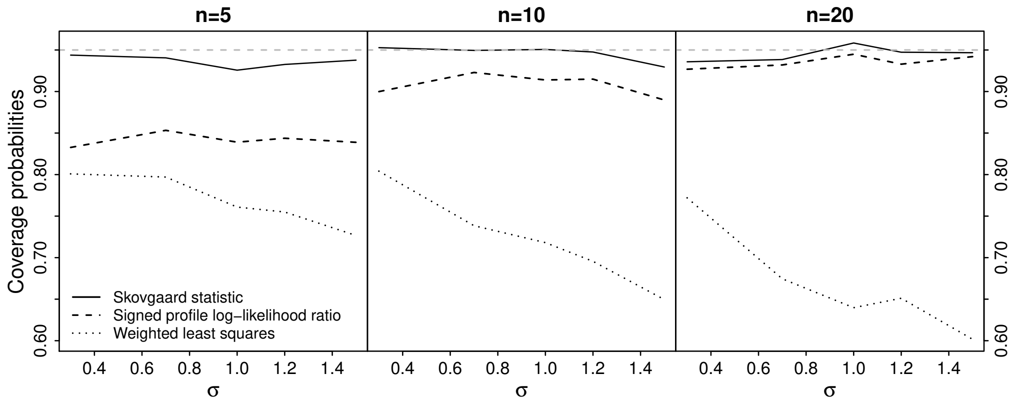

Figure B5: Empirical coverage probabilities of the nominally 95% confidence interval for β_1 when $(\beta_0, \beta_1, \mu)^\top = (-1.5, 1, -0.5)^\top$, under increasing sample size n and square root σ of the variance σ^2 in the control group. Variance component $\tau^2 = 1.2$. The plotted curves correspond to Skovgaard's statistic (solid), the signed profile log-likelihood ratio statistic (dashed), the weighted least squares approach (dotted). The dashed, grey horizontal line is the nominal level.

- Scenario with $(\beta_0, \beta_1, \mu)^\top = (-1.5, 1, -2.5)^\top$, called *scenario 3* in the main text, with τ^2 equal to 1.2: Figure B6

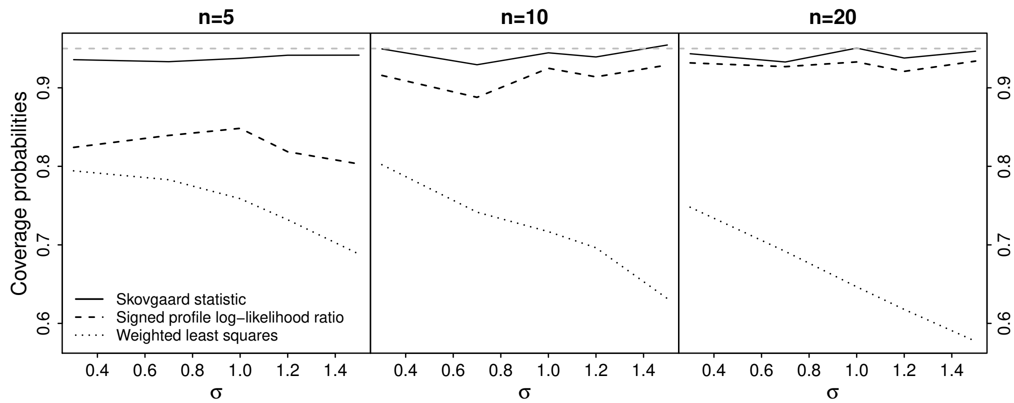

Figure B6: Empirical coverage probabilities of the nominally 95% confidence interval for β_1 when $(\beta_0, \beta_1, \mu)^\top = (-1.5, 1, -2.5)^\top$, under increasing sample size n and square root σ of the variance σ^2 in the control group. Variance component $\tau^2 = 1.2$. The plotted curves correspond to Skovgaard's statistic (solid), the signed profile log-likelihood ratio statistic (dashed), the weighted least squares approach (dotted). The dashed, grey horizontal line is the nominal level.

- Scenario with $(\beta_0, \beta_1, \mu)^\top = (-3, 1, -2)^\top$, called *scenario 4* in the main text, with τ^2 equal to 1.2: Figure B7

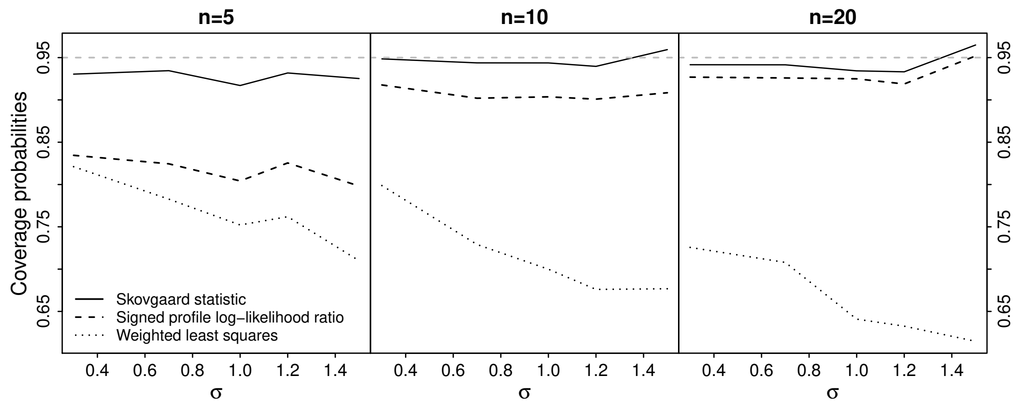

Figure B7: Empirical coverage probabilities of the nominally 95% confidence interval for β_1 when $(\beta_0, \beta_1, \mu)^\top = (-3, 1, -2)^\top$, under increasing sample size n and square root σ of the variance σ^2 in the control group. Variance component $\tau^2 = 1.2$. The plotted curves correspond to Skovgaard's statistic (solid), the signed profile log-likelihood ratio statistic (dashed), the weighted least squares approach (dotted). The dashed, grey horizontal line is the nominal level.

Web Appendix C: Data analysis

This appendix shows how to evaluate Skovgaard's statistic for inference on the slope of the control rate regression in the `R` programming language. The illustration is based on the data of Hoes et al. [1] reported in Table 1 of the paper. Functions needed to implement Skovgaard's statistic are obtained as supplementary material and they can be loaded as follows

```
R> source("control_rate_regression_LRTs.R")
```

Consider the hypothesis test $\beta_1 = 1$ against the two-sided alternative. Wald statistic, the signed profile log-likelihood ratio statistic and Skovgaard's statistic are obtained by applying function `crr.test` (control rate regression test)

```
crr.test(data, beta1.null, alternative = c("two.sided", "less",
    "greater"), maxit = 1000)
```

with arguments

- `data`: the dataset
- `beta1.null`: the value of β_1 under the null hypothesis
- `alternative`: a character string specifying the alternative hypothesis, chosen between "two.sided" (default), "greater" or "less"; just the initial letter can be specified
- `maxit`: the maximum number of iterations for the Nelder and Mead [2] optimization algorithm; default value 1,000

The dataset is composed by n rows corresponding to the studies recruited in the meta-analysis and 6 columns including the values of $\hat{\eta}_i$, $\hat{\xi}_i$, and the elements of the variance/covariance matrix Γ_i inserted by row, namely, $\mathrm{var}(\hat{\eta}_i)$, $\mathrm{cov}(\hat{\eta}_i, \hat{\xi}_i)$, $\mathrm{cov}(\hat{\eta}_i, \hat{\xi}_i)$, $\mathrm{var}(\hat{\xi}_i)$. For the analysis of Hoes et al. [1] data, the object to be passed to function `crr.test` can be constructed as follows

```
R> deaths.treated <- c(10, 2, 54, 47, 53, 10, 25, 47, 43, 25, 157, 92)
R> ## number of person-years for the cases
R> py.treated <- c(595.2, 762, 5635, 5135, 3760, 2233, 7056.1, 8099,
R+                 5810, 5397, 22162.7, 20885)
R> deaths.controls <- c(21, 0, 70, 63, 62, 9, 35, 31, 39, 45, 182, 72)
```

```
R> deaths.controls[2] <- 0.5
R> ## number of person-years for the controls
R> py.controls <- c(640.2, 756, 5600, 4960, 4210, 2084.5, 6824, 8267,
R+                 5922, 5173, 22172.5, 20645)
R> py.controls[2] <- py.controls[2]+0.5
R> hoes.data.original <- data.frame(deaths.treated, py.treated,
R+                                  deaths.controls, py.controls)
## estimated log event rate for the controls
R> xi.obs <- log(hoes.data.original$deaths.treated /
R+               hoes.data.original$py.treated)
## estimated log event rate for the treated
R> eta.obs <- log(hoes.data.original$deaths.controls /
R+                hoes.data.original$py.controls)
R> n <- length(hoes.data.original$deaths.treated)
## variance/covariance matrix
R> gamma.matrix <- matrix(0.0, ncol=4, nrow=n)
R> for(i in 1:n)
R+   gamma.matrix[i,] <- c(1/hoes.data.original$deaths.treated[i], 0,
R+                         0, 1/hoes.data.original$deaths.controls[i])
R> hoes.data <- data.frame(eta.obs, xi.obs, gamma.matrix)
R> colnames(hoes.data) <- c('eta.obs', 'xi.obs', 'var.eta', 'cov.etaxi',
R+                          'cov.etaxi', 'var.xi')
```

Function `crr.test`

```
R> crr.test(data=hoes.data, beta1.null=1, alternative='two.sided')

Estimate of beta1:
     Estimate  Std.Err.
WLS  0.60973   0.10892
MLE  0.68917   0.08124

Hypothesis test for beta1:
                                               Value       P-value
Wald statistic                                 -3.5830787   0.0003396
Signed profile log-likelihood ratio statistic  -2.3447177   0.0190415
Skovgaard statistic                            -1.2709290   0.2037539

alternative hypothesis: parameter is different from 1
```

provides the following information:

- the weighted least squares estimate and the maximum likelihood estimate of β_1;
- the associated standard error;
- the value of Wald statistic, the value of the signed profile log-likelihood ratio statistic r_P and the value of Skovgaard's statistic $\overline{r}_P$ under the null hypothesis;
- the p-value of the test based on the three statistics for the specified alternative hypothesis.

References

[1] Hoes AW, Grobbee DE, Lubsen J. Does drug treatment improve survival? Reconciling the trials in mild-to-moderate hypertension. *Journal of Hypertension* 1995; **13**: 805–811.

[2] Nelder JA, Mead R. A simplex algorithm for function minimization. *Scandinavian Journal of Statistics* 1965; **7**: 308–313.